On combinatorial approximation algorithms in geometry

Bruno Jartoux

Université Paris-Est
École doctorale MSTIC

On Combinatorial Approximation Algorithms in Geometry

Sur les algorithmes d'approximation combinatoires en géométrie

Bruno Jartoux

Thèse de doctorat en informatique

soutenue le 12 septembre 2018.

Composition du jury :

Lilian Buzer	ESIEE Paris	
Jean Cardinal	Université libre de Bruxelles	*président du jury*
Guilherme Dias da Fonseca	Université Clermont Auvergne	*rapporteur*
Jesús A. de Loera	University of California, Davis	*rapporteur*
Frédéric Meunier	École nationale des ponts et chaussées	
Nabil H. Mustafa	ESIEE Paris	*directeur*
Vera Sacristán	Universitat Politècnica de Catalunya	
Kasturi R. Varadarajan	The University of Iowa	*rapporteur*

Last revised 16th December 2018.

Thèse préparée au laboratoire d'informatique Gaspard-Monge (LIGM), équipe A3SI, dans les locaux d'ESIEE Paris.

LIGM UMR 8049
Cité Descartes, bâtiment Copernic
5, boulevard Descartes
Champs-sur-Marne
77454 Marne-la-Vallée Cedex 2

ESIEE Paris
Département IT
Cité Descartes
2, boulevard Blaise-Pascal
93162 Noisy-le-Grand Cedex

If I have seen further it is by standing on the shoulders of Giants.

Newton, citing Bernard of Chartres

If I have not seen as far as others, it is because giants were standing on my shoulders.

Hal Abelson

Summary

The analysis of approximation techniques is a key topic in computational geometry, both for practical and theoretical reasons. In this thesis we discuss sampling tools for geometric structures and geometric approximation algorithms in combinatorial optimisation.

Part I focuses on the combinatorics of geometric set systems. We start by discussing packing problems in set systems, including extensions of a lemma of Haussler, mainly the so-called *shallow packing lemma*. For said lemma we also give an optimal lower bound that had been conjectured but not established in previous work on the topic. Then we use this lemma, together with the recently introduced *polynomial partitioning* technique, to study a combinatorial analogue of the Macbeath regions from convex geometry: Mnets, for which we unify previous existence results and upper bounds, and also give some lower bounds. We highlight their connection with epsilon-nets, staples of computational and combinatorial geometry, for example by observing that the unweighted epsilon-net bound of Chan et al. (SODA 2012) or Varadarajan (STOC 2010) follows directly from our results on Mnets.

Part II deals with local-search techniques applied to geometric restrictions of classical combinatorial optimisation problems. Over the last ten years such techniques have produced the first polynomial-time approximation schemes for various problems, such as that of computing a minimum-sized hitting set for a collection of input disks from a set of input points. In fact, it was shown that for many of these problems, local search with radius $\Theta(1/\epsilon^2)$ gives a $(1+\epsilon)$-approximation with running time $n^{O(1/\epsilon^2)}$. However the question of whether the exponent of n could be decreased to $o(1/\epsilon^2)$ was left open. We answer it in the negative: the approximation guarantee of local search cannot be improved for any of these problems.

Keywords computational geometry, combinatorial optimisation, epsilon-nets, approximation algorithms, local search

Résumé

L'analyse des techniques d'approximation est centrale en géométrie algorithmique, pour des raisons pratiques comme théoriques. Dans cette thèse nous traitons de l'échantillonage des structures géométriques et des algorithmes d'approximation géométriques en optimisation combinatoire.

La première partie est consacrée à la combinatoire des hypergraphes. Nous débutons par les problèmes de *packing*, dont des extensions d'un lemme de Haussler, particulièrement le lemme dit de *shallow packing*, pour lequel nous donnons aussi un minorant optimal, conjecturé mais pas établi dans les travaux antérieurs. Puis nous appliquons ledit lemme, avec la méthode de *partition polynomiale* récemment introduite, à l'étude d'un analogue combinatoire des régions de Macbeath de la géométrie convexe : les M-réseaux, pour lesquels nous unifions les résultats d'existence et majorations existants, et donnons aussi quelques minorants. Nous illustrons leur relation aux epsilon-réseaux, structures incontournables en géométrie combinatoire et algorithmique, notamment en observant que les majorants de Chan et al. (SODA 2012) ou Varadarajan (STOC 2010) pour les epsilon-réseaux (uniformes) découlent directement de nos résultats sur les M-réseaux.

La deuxième partie traite des techniques de recherche locale appliquées aux restrictions géométriques de problèmes classiques d'optimisation combinatoire. En dix ans, ces techniques ont produit les premiers schémas d'approximation en temps polynomial pour divers problèmes tels que celui de calculer un plus petit ensemble intersectant pour une ensemble de disques donnés en entrée parmi un ensemble de points donnés en entrée. En fait, il a été montré que pour de nombreux tels problèmes, la recherche locale de rayon $\Theta(1/\epsilon^2)$ donne une $(1+\epsilon)$-approximation en temps $n^{O(1/\epsilon^2)}$. Savoir si l'exposant de n pouvait être ramené à $o(1/\epsilon^2)$ demeurait une question ouverte. Nous répondons par la négative : la garantie d'approximation de la recherche locale n'est améliorable pour aucun desdits problèmes.

Mots-clefs géométrie algorithmique, optimisation combinatoire, algorithmes d'approximation, epsilon-réseaux, recherche locale

Introduction à l'usage du lecteur francophone

Le lecteur de langue française pourra regretter que l'emploi de l'anglais, *lingua franca* de la recherche scientifique, crée entre lui-même et l'auteur un obstacle superflu outre ceux déjà introduits par l'écrit. Ces quelques pages lui sont dédiées à titre d'excuse. Elles suivent à peu près la structure du chapitre d'exposition rédigé en anglais.

Ces travaux de thèse ressortissent surtout à la géométrie algorithmique, discipline assez jeune née dans les années 1970 de la spécialisation de l'algorithmique aux objets géométriques.

❧

❧ ❧

L'objectif unissant les travaux présentés dans ce mémoire est l'analyse d'algorithmes d'approximation dans le contexte spécifique de l'optimisation combinatoire géométrique. Sous l'hypothèse, largement considérée comme plausible, que la classe **P** des problèmes admettant un algorithme polynomial soit strictement incluse dans la classe **NP**, plusieurs problèmes d'importance capitale ne sauraient être résolus par des algorithmes polynomiaux.

L'algorithméticien est contraint d'envisager une notion de résolution plus large pour contourner ces limitations théoriques. Pour les problèmes d'optimisation, s'il n'est pas envisageable de calculer efficacement une solution optimale, on essaie d'en proposer une dont la valeur soit garantie *assez proche* dudit optimum ; dans la pratique, il sera souvent préférable d'être à quatre-vingt-quinze pourcent de l'optimum en temps raisonnable, plutôt que de l'atteindre après de longs calculs.

Outre ces aspects pratiques, l'étude des propriétés d'approximabilité de tels problèmes est en soi une contribution à la théorie de la complexité. Par exemple, le problème du voyageur de commerce est également **NP**-complet dans sa formulation générale et dans un monde euclidien, mais le fait qu'il admette un *schéma d'approximation en temps polynomial* dans ce second cas (travaux pour lesquels S. Arora et J. Mitchell ont partagé le prix Gödel en 2010) suggère une différence de difficulté entre les deux versions.

La première partie de ce manuscrit est dédiée à certains problèmes combinatoires dans les hypergraphes. Eu égard à une famille de parties tirées d'un même ensemble *(univers)* à n éléments (une telle famille et ses éléments sont un *hypergraphe* et ses *arêtes* dans la terminologie classique de Berge [21], ou un *set system* et ses *ranges* en anglais) et à un paramètre strictement positif ϵ, on se propose de prélever dans l'univers un échantillon de façon à intersecter chaque arête contenant au moins ϵn éléments. Un tel échantillon est appelé ϵ-réseau *(ϵ-net)*.

Le succès de cet objet en géométrie algorithmique est attribuable à deux facteurs principaux. D'une part, cette notion capture bien certaines propriétés souvent attendues d'un échantillon statistique : l'ϵ-réseau intersecte toutes les arêtes les plus grandes et le sur-échantillonage est permis puisqu'un sur-ensemble d'un ϵ-réseau demeure un ϵ-réseau. (Par comparaison, d'autres notions d'approximation plus contraignantes exigent non seulement que les arêtes de cardinalité ϵn soient atteintes, mais aussi qu'elles le soient en proportion de leur taille. L'inconvénient en est que cela peut nécessiter de prélever bien plus d'éléments.) D'autre part, Haussler et Welzl [60] ont montré que les hypergraphes de dimension de Vapnik–Chervonenkis bornée[1] admettent toujours des ϵ-réseaux de taille

$$O\left(\frac{1}{\epsilon}\log\frac{1}{\epsilon}\right),$$

ceci indépendemment de la cardinalité de l'univers. Les hypergraphes définis par intersection de n points avec des régions géométriques « simples » (demi-plans, triangles, coniques, etc.) ont habituellement cette propriété. Cette quantité bornée indépendamment de n est à comparer avec une méthode d'échantillonage qui prélèverait une fraction constante de l'univers. En outre, connaissant la taille du plus petit ϵ-réseau, un tirage aléatoire d'un échantillon de cette cardinalité produit un tel réseau avec probabilité constante.

L'introduction ultérieure de mesures plus fines de la complexité des hypergraphes, telle que la complexité en cellules peu profondes (qui est à peu de chose près la fonction de répartition des cardinalités des arêtes), a permis ces dernières années de préciser les résultats classiques sur la taille minimale de ces ϵ-réseaux [32, 99]. Lorsque pour tous entiers m et l tout sous-hypergraphe induit sur m éléments a au plus $\tilde{\varphi}(m)\cdot l^{O(1)}$ arêtes de cardinalité inférieure à l avec $\tilde{\varphi}$ majorée par un polynôme, la taille des plus petits ϵ-réseaux est

$$O\left(\frac{1}{\epsilon}\log\tilde{\varphi}\left(\frac{1}{\epsilon}\right)\right),$$

ce qui étend le majorant précédent, et l'améliore dans certains cas : par exemple pour les disques de $\mathbb{R}^2$, qui ont $\tilde{\varphi}(m) = O(1)$.

Nous contribuons à cette étude de l'échantillonnage des hypergraphes par des résultats sur une généralisation du réseau récemment introduite : le M-réseau (*Mnet*).

1. Pour une définition de cette propriété purement combinatoire, voir *Definition 1.3*.

En géométrie convexe, toute partie compacte et convexe C d'un espace euclidien inclut pour tout petit paramètre $\epsilon > 0$ un petit nombre de « régions de Macbeath » convexes, toutes de volume proportionnel à $\epsilon \cdot \operatorname{vol} C$ et telles que tout demi-espace affine qui inclue une fraction ϵ de C contienne une de ces régions [23]. Le M-réseau, introduit par Mustafa et Ray [86], en est l'analogue combinatoire : un ϵ-M-réseau d'un hypergraphe à n éléments consiste en une famille de parties de l'univers (mais pas nécessairement d'arêtes), toutes de cardinalité $\Omega(\epsilon n)$, telles que toute arête à plus d'ϵn éléments en inclue au moins une.

La principale contribution de cette première partie est l'unification de précédents résultats sur l'existence et la taille d'un plus petit M-réseau dans les hypergraphes géométriques, résultats initialement obtenus par des méthodes disparates. Notre approche combine deux outils : d'une part, le point de vue algébrique pour la manipulation des incidences entre points et régions géométriques (plus particulièrement la technique de partition polynomiale introduite par Guth et Katz [54]), d'autre part, un lemme de Haussler [59] sur l'entropie métrique (la taille maximale d'une famille d'arêtes (δ/n)-discernable, c'est-à-dire dont les différences symétriques deux à deux comprennent toutes au moins δ points, ce qu'en anglais on appelle un problème de *packing*) des hypergraphes dont la dimension de Vapnik–Chervonenkis est bornée :

Théorème (Haussler)**.** *Dans un hypergraphe sur n éléments de dimension de Vapnik–Chervonenkis $d \in \mathbb{N}$, toute famille (δ/n)-discernable a $O((n/\delta)^d)$ arêtes.*

Si le résultat final est une majoration de la cardinalité des plus petits M-réseaux, dont nous montrons aussi qu'elle est presque optimale, nos outils intermédiaires incluent de nouveaux résultats sur deux variantes du lemme de Haussler exploitant le rang de l'hypergraphe, c'est-à-dire la cardinalité maximale de ses arêtes. En particulier nous donnons une borne inférieure conjecturée mais pas obtenue dans des travaux antérieurs, et ce par la construction explicite d'un hypergraphe. De plus, nous observons qu'il est possible d'obtenir un ϵ-réseau par une méthode probabiliste à partir d'un ϵ-M-réseau du même hypergraphe. Nous dérivons par cette nouvelle méthode les théorèmes usuels sur la cardinalité des ϵ-réseaux, illustrant de ce fait les liens étroits entre les deux structures.

Dans un second temps, nous étudions les algorithmes de recherche locale. Là où l'exploration totale de l'espace des solutions d'un problème n'est pas envisageable en raison de sa taille, les méthodes de ce type limitent la recherche à un voisinage de la solution initiale, pourvu bien sûr qu'une notion de voisinage ait été définie sur l'espace des solutions.

Le problème archétypal que nous considérons est une restriction géométrique de celui de Plus Petit Ensemble Intersectant *(Minimum Hitting Set)*. Étant donnée une instance composée d'ensembles finis $\mathcal{P}$ et $\mathcal{D}$ de points et de disques de $\mathbb{R}^2$, l'objectif est de calculer une plus petite partie de $\mathcal{P}$ qui intersecte tous les disques de $\mathcal{D}$.

BRÖNNIMANN et GOODRICH [24] ont donné un algorithme qui, exploitant l'existence d'ϵ-réseaux de cardinalité $O(1/\epsilon)$ pour l'hypergraphe défini sur $\mathcal{P}$ par $\mathcal{D}$, retourne en temps polynomial une solution dont la cardinalité est à un facteur multiplicatif constant de l'optimum, soit une $O(1)$-approximation.

Ce résultat peut sembler surprenant puisque la version non géométrique du problème, c'est-à-dire la recherche d'un plus petit transversal dans un hypergraphe quelconque à n sommets, est au mieux $\log(n)$-approximable. Toutefois cette méthode ne saurait fournir une $(1+\epsilon)$-approximation dès lors qu'ϵ est petit. Un autre outil permet d'attaquer ce problème : la recherche locale.

Son principe est d'améliorer itérativement une solution initiale. Spécifiquement, à chaque étape on s'autorise à retirer $t \leq \lambda$ éléments de la solution courante (ici, des points de $\mathcal{P}$) et à les remplacer par au plus $t-1$ autres, pourvu que l'ensemble résultant soit toujours solution. L'entier λ est le *pas*, ou *rayon*, de la recherche locale. On définit ainsi une règle de parcours sur l'espace des solutions, qui reçoit une structure de graphe orienté acyclique. Les solutions que l'on ne peut améliorer de cette manière sont dites *localement optimales*. Puisque l'exploration du voisinage à chaque étape nécessite de considérer $n^{O(\lambda)}$ solutions candidates, ce parcours de l'espace des solutions doit s'achever en un optimum local en un temps $n^{O(\lambda)}$.

Si l'idée d'amélioration par recherche limitée à un petit voisinage (ici une boule de rayon 2λ pour la distance de Hamming qui à deux ensembles A et B associe la cardinalité $|(A \cup B) \setminus (A \cap B)|$) est commune à tous les algorithmes de recherche locale, qu'on les applique à des problèmes discrets ou continus, la singularité de cette approche en optimisation combinatoire géométrique est l'existence de garanties liant les optima locaux à l'optimum global. Plus précisément, AGARWAL et MUSTAFA [3] observent que la recherche locale fournit une $O(1)$-approximation pour un certain problème géométrique[2], puis CHAN et HAR-PELED [33] et MUSTAFA et RAY [87] montrent simultanément que le ratio d'approximation peut être pris aussi proche de 1 que voulu lorsque le pas augmente, à savoir $1+\epsilon$ pour $\lambda = \Theta(1/\epsilon^2)$.

Cette technique a été étendue par de nombreux auteurs à de multiples problèmes géométriques d'optimisation combinatoire, fournissant pour tous une $(1+\epsilon)$-approximation en temps $n^{O(1/\epsilon^2)}$. Dans la deuxième partie de cette thèse, nous étudions une question posée dès les premières applications [52, 87], à savoir s'il est possible par une analyse plus rigoureuse de diminuer cet exposant : par exemple on voudrait savoir si l'algorithme termine en fait en $n^{O(1/\epsilon)}$, voire même en $n^{o(1/\epsilon)}$.

La clef de l'analyse de ces algorithmes est un théorème sur les graphes bipartis. On appelle un tel graphe k-*expanseur* lorsque chaque ensemble de t sommets à gauche a au moins t voisins à droite, ce pour tout $t \leq k$. Par le théorème de HALL, cela revient aussi à exiger que chaque ensemble d'au plus k sommets à gauche participe à un couplage.

2. À savoir PLUS GRAND ENSEMBLE INDÉPENDANT dans un graphe d'intersection de rectangles.

Théorème. *Si un graphe planaire biparti, d'ensembles de sommets gauche et droit G et D, est k-expanseur, $k \geq 3$, alors*

$$|G| \leq \left(1 + \frac{c}{\sqrt{k}}\right) |D|$$

pour une certaine constante $c > 0$.

L'analyse commune aux divers algorithmes géométriques de recherche locale de pas λ, telle qu'elle apparaît par exemple chez Chan et Grant [31], Chan et Har-Peled [33], Gibson et Pirwani [52] et Mustafa et Ray [87], passe par la construction d'un *graphe d'échange* planaire, biparti et λ-expanseur, dont les ensembles de sommets sont deux solutions, l'une étant un optimum local. On conclut alors que le rapport de leurs cardinalités doit être au plus $1 + c/\sqrt{\lambda}$ (par analogie avec la programmation linéaire, nous appelons cette quantité *écart de localité* du problème), et donc qu'un pas $\lambda = \Theta(1/\epsilon^2)$ correspond à une garantie d'approximation de $1 + \epsilon$.

Nous prouvons que l'inégalité sur les graphes bipartis est, à un facteur multiplicatif près, optimale ; ceci par la description explicite d'une certaine famille de graphes planaires et k-expanseurs sur des ensembles de sommets (G_n, D_n) vérifiant

$$\begin{aligned} &|G_n|, |D_n| = \Theta(n), \\ \text{et } &|G_n| \sim \left(1 + \frac{c'}{\sqrt{k}}\right) |D_n| \text{ lorsque } n \to \infty. \end{aligned}$$

À partir de la famille de graphes saturant l'inégalité, nous construisons pour divers problèmes géométriques des instances dans lesquelles l'écart de localité est effectivement de $1 + \Omega(1/\sqrt{\lambda})$, y compris pour celui du Plus Petit Ensemble Intersectant dans un hypergraphe défini par des disques. Par conséquent, nous pouvons répondre par la négative à la question posée précédemment : l'exposant $O(1/\epsilon^2)$ du temps d'exécution de la recherche locale n'est pas contingent, mais nécessaire si l'on ne connaît de la structure des graphes d'échange que leur planarité.

Enfin, l'analyse des algorithmes de recherche locale s'étend à des graphes d'échange non nécessairement planaires, mais possédant toujours des séparateurs de taille souslinéaire, et notre construction aussi : pour les graphes k-expanseurs dont tout sous-graphe d'ordre m possède un séparateur équilibré[3] de taille $O(m^{1-1/d})$, avec $d \in \mathbb{N}^*$, le théorème ci-dessus s'applique toujours, mais avec un facteur de la forme $1 + O(k^{-1/d})$, tandis que notre construction atteint un ratio $1 + \Omega(k^{-1/d})$.

3. C'est-à-dire un ensemble de sommets dont la suppression ne laisse pas plus de $2m/3$ sommets dans une même composante connexe.

Acknowledgements

As I finalise this manuscript, listing each and every person to whom I am indebted is quite simply impossible. As it turns out, it also takes a village to get a PhD. Still, let us try to come up with a fair, representative *sample*. (There's your real-life motivation for the results in Part I. Now to work local search heuristics into this.)

First and foremost my thanks go to my doctoral advisor. Over three years, Nabil has done everything to impart to me his encyclopædic knowledge of computational geometry—always with patience and pedagogy.

I am also honoured by the members of this jury, who took the time to read my dissertation and to sit on its defence.

Within the doctoral school, Sylvie Cach was very helpful when dealing with administrative or regulatory questions—that is, fairly often.

The broader computational geometry community has been nothing but welcoming and I have met many of its members with great pleasure, be they senior researchers or fellow PhD candidates. Out of all those researchers I am especially grateful to Arijit Ghosh and Kunal Dutta ; if fun were publishable we'd all hold chairs by now.

Jean-Daniel Boissonnat kindly invited me to visit his Datashape research team in Nice. Mathijs Wintraecken made it much more fun to get betrayed by Google Maps in foreign suburbia on the other side of the earth. Frédéric Meunier essentially made this PhD happen.

It was a pleasure to share some courseload with Stéphane Breuils, or at least a shared suffering when it came to grading. Thanks also to Vincent Nozick for entrusting us both with a full course in his beloved IMAC department. Kacper Pluta's company is much appreciated as we both enjoy complaining about software. (This in spite of his too-often objectively incorrect taste in those matters). Matthieu Pluntz kindly helped with the French in this dissertation. Globish and TCS jargon have crept into mine over the last few years.

Thanks to the friends and fellow PhD candidates who toiled and moiled alongside : Clara, Daniel, Deise, Diane, Edward, Élodie, Éloise, Julie, Julien, Kacper, Karla, Ketan, Martin, Pavel, Raphaël, Siddarth, Stéphane, Thibault and many others still.

As I reread the thesis of Shamos (1978), typeset on a PDP-11 in the PUB language with some 150 stylus-drawn pictures, I am also grateful to the oh-so-many developers (of GNU and Linux, of TeX, LaTeX, LuaLaTeX and such packages as `TikZ`, `biblatex` and `memoir`) who together will have enabled me to write this in three years rather than five.

I will close by noting that my friends and family have always been kind, patient and supportive. To a PhD candidate in their chrysalis stage—when arXiv binge-reading and thesis proofreading supplant human interaction—that is both a miracle and a blessing.

Paris, 7th September 2018

Notations and Terminology

In this dissertation, the letters $\mathbb{N}$ and $\mathbb{R}$ stand for the sets of natural numbers (including 0) and real numbers respectively. The (possibly infinite) cardinality of a set X is denoted by $|X|$, its powerset is $\mathcal{P}(X)$, its $0-1$ valued indicator function is $\mathbb{1}_X$ and the symmetric difference of two set sets X and Y, denoted by $X \,\Delta\, Y$, is $X \cup Y \setminus (X \cap Y)$. For $n \in \mathbb{N}$, the set $\binom{X}{n}$ consists of all subsets of X of cardinality n.

For any $(m, n) \in \mathbb{N}^2$ we write $[m\,..\,n]$ for the integer interval $\{i \in \mathbb{N} : m \leq i \leq n\}$ as opposed to $[m, n]$ and (m, n) for the corresponding closed and open real intervals.

For a real x, we write $\lceil x \rceil$ for the smallest integer greater than $x - 1$ and $\lfloor x \rfloor$ for the largest integer smaller than $x + 1$.

A graph is a pair (V, E) where V is a (possibly infinite) set and $E \subseteq \binom{V}{2}$. The elements of V are called vertices and those of E edges. For any subset X of V, the set $N(V)$ consists of the neighbours of X, that is, the vertices in $V \setminus X$ that share an edge with a vertex of X.

Wherever they appear, the functions φ and $\tilde{\varphi}$ correspond to the shallow-cell complexity of a set system as defined in Chapter 1. They will never have another meaning.

The function log is the natural logarithm. The function $\log^*$ is the iterated logarithm, an integer-valued slow-growing function defined by

$$\log^* : x \in (0, +\infty) \mapsto \min \left\{ n \in \mathbb{N} : \underbrace{(\log \circ \cdots \circ \log)}_{n \text{ times}}(x) \leq 1 \right\}.$$

Concepts are **highlighted** where they are first defined. Finally, our contributions are numbered whereas theorems from other authors are labeled with letters.

Contents

List of Figures

List of Tables

Chapter Zero

A Detailed Overview

THIS INTRODUCTORY CHAPTER lays out in broad strokes the main ideas of this work. With technical minutiae and the machinery of proofs committed to later chapters, we want to give a landscape view of the results, their context and their interplay. Appropriate references are given so that the details of proofs can be found in the remainder of the thesis.

Our core problem is the study of approximation in geometric settings. This is motivated both by practical applications (a key topic in computational geometry is to approximate a complicated object with a simpler one that has roughly the same properties of interest: examples include polytope approximation, surface simplification, Vapnik-Chervonenkis theory, coresets) and by theoretical consequences, as showing that some problems whose exact resolution is challenging admit efficient approximation algorithms—or that others do not—helps chart the 'invisible electric fence'[1] between 'easy' and 'hard' problems.

In this chapter, original contributions are highlighted in the margin.

Combinatorics of Geometric Set Systems

The first part of this thesis is dedicated to the study of *set systems* or *range spaces*: those are just a family of subsets (*ranges*) of a common universe (the *ground set*). As geometers we are particularly interested with such systems built from geometric objects, for example when the ground set is a subset of $\mathbb{R}^2$ and the ranges are all its intersections with disks.

1. The simile is S. Aaronson's [1].

Nets: sampling set systems

Given a set system $(X, \mathcal{R})$ with $\mathcal{R} \subseteq \mathcal{P}(X)$, a simple and natural problem is to compute a small set of elements of X that together intersect all nonempty ranges in $\mathcal{R}$. For example, let X be a subset of $\mathbb{R}^2$ and let the ranges of $\mathcal{R}$ be of the form $X \cap D$ with D a disk (we say that $\mathcal{R}$ is defined on X by disks). If we allow D to be any disk, there will be $|X|$ singleton ranges and only the whole universe X will be a solution. This is not too interesting, but we might do better if we introduce the condition that all ranges contain at least a fraction $\epsilon \in (0, 1)$ of X. Formally, given a set system $(X, \mathcal{R})$ with a finite subset Y of X (very often we will take $Y = X$) and a positive parameter ϵ, an *ϵ-net* for Y is any $N \subseteq Y$ such that

$$\forall R \in \mathcal{R},\ |R \cap Y| \geq \epsilon\,|Y| \implies N \cap R \neq \emptyset.$$

Applications

Nets share a property with random samples: they hit the most significant ranges. (They are not however required to hit the ranges in proportion to their cardinalities, a stronger constraint that defines *approximations*. Approximations are nets, but the smallest nets can be much smaller than approximations.)

This is sufficient for nets to replace random samples in various algorithmic applications that rely on this property, and constructive methods to build nets are used to derandomise sampling-based algorithms. For example nets have been employed in deterministic algorithms for linear programming since the nineties, see Chan [30] for a recent reference. In this setting the ground set's elements are the linear constraints (forbidden half-spaces) and each range consists of all constraints violated by a same point, that is, all half-spaces that contain this point. Then an ϵ-net is a subset N of constraints such that any point which violates a fraction ϵ of all constraints must violate at least one of those in N. Conversely solutions that satisfy all constraints in N will satisfy at least a fraction $1 - \epsilon$ of all constraints. Another use, which we will discuss in a few pages, is in optimisation problems that involve hitting sets or set covers.

Finally there is also a natural interpretation in statistical learning. There the ranges in $\mathcal{R}$ are called *classifiers*. In realisable classification problems, a specific classifier R^* is distinguished in $\mathcal{R}$, but is not known to the learner. Without loss of generality we can assume that R^* is the empty set: translate all ranges by $R \mapsto R \,\Delta\, R^*$. A random sample $x_1, \ldots, x_m$ is drawn from X and one is told for each i whether $x_i \in R^*$. The goal is to identify R^* based on this knowledge. If the sample is an ϵ-net for $(X, \mathcal{R})$, this information is sufficient to reject classifiers that disagree with R^* on more than $\epsilon\,|X|$ elements. The question is then to determine how large m must be for the sample to be a net with high probability. See the recent paper by Kupavskii and Zhivotovskiy [70] for such parallels between computational geometry and learning theory.

The main question about nets is to determine how small they can be depending on the combinatorics of the set system, that is, the cardinality of a smallest net. For our example with disks in $\mathbb{R}^2$, we can build an X for which every ϵ-net must have at least $1/\epsilon$ points: consider $1/\epsilon$ disjoint disks containing ϵn points each and let X be the union of these sets of points, then an ϵ-net for X (with respect to disks) must contain at least one point from each disk. (For simplicity assume that $1/\epsilon$ and ϵn are integers.) The same construction applies also to ranges defined by half-planes, or rectangles, or other geometric regions, so that in general the size of the smallest ϵ-nets should be at least $1/\epsilon$.

On the other hand, the system $([n], \mathcal{P}([n]))$ whose ranges are all subsets of $[n]$ cannot have ϵ-nets with fewer than $(1-\epsilon)n$ elements, i.e. almost all of $[n]$ must be taken in any net. Note that this system is geometric: it is (up to isomorphism) the system defined on n points in convex position by all convex regions. Thus any good upper bound result on the size of nets (such as (1) below) will require additional constraints on $(X, \mathcal{R})$.

Geometric Set Systems

Although the definition of nets is purely combinatorial, their study has been grounded in geometric intuition since its inception.

For example, consider a finite set of points X in $\mathbb{R}^2$. Remember that the ϵ-net problem for half-planes asks how small $N \subseteq X$ can be if every half-plane containing $\lceil \epsilon \left|X\right| \rceil$ points of X contains at least a point of N. The answer in this case relies on and highlights the fact that half-plane arrangements are simpler than abstract set systems. It follows from a combinatorial property enjoyed by many natural geometric set systems (defined by half-planes in the plane, by half-spaces in higher-dimensional Euclidean space, by disks and other semi-algebraic regions, etc.): finite VC dimension.

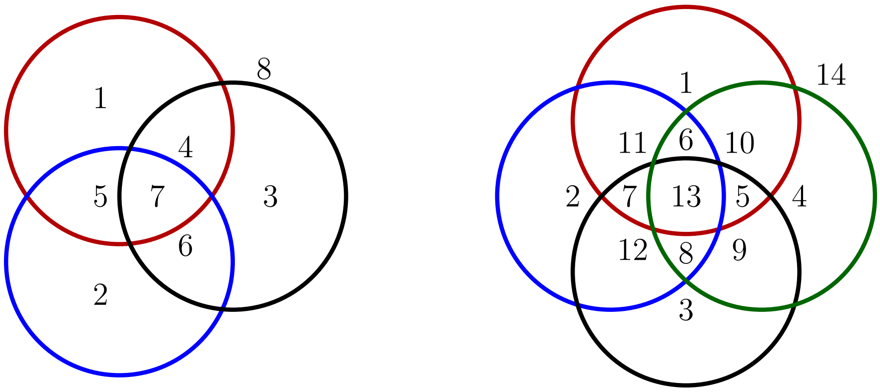
Broadly, for any such set system, there is an integer d (the *VC dimension*) such that every set of $d+1$ points has a subset that is not obtainable as the intersection of the points with a range. On Figure 0.1a, one sees that this is true of the system defined by disks in $\mathbb{R}^2$, with $d = 3$. Intervals of $\mathbb{R}$ and half-spaces of $\mathbb{R}^d$, with VC dimensions 2 and $d+1$, are popular examples in learning theory.

Building on ideas of Vapnik and Chervonenkis [97], Haussler and Welzl [60] in the seminal paper that introduced nets proved that any set system with finite VC dimension d has small ϵ-nets. After some later simplifications [22], their bound is

$$O\left(\frac{d}{\epsilon}\log\frac{1}{\epsilon}\right). \tag{1}$$

Their proof is probabilistic: it is based on a modification of the *ghost sample* technique of Vapnik and Chervonenkis [97]. They show that taking a random sample of the

(a) Disks have VC dimension 3. Given any four points of $\mathbb{R}^2$, some subset cannot be cut out by a disk. (There are in essence two configurations to examine. In both any disk that contains the red squares must contain one black point.)



(b) Disks also have dual VC dimension 3. Three disks of $\mathbb{R}^2$ may define 2^3 full-dimensional cells (sign patterns) but no arrangement of four disks has 2^4 cells.

Figure 0.1: Disks have both *primal* and *dual* VC dimension 3. See Definition 1.2 for the definition of set system duality.

ground set of this size yields an ϵ-net with constant probability. The most important fact here is that the size of this sample does not depend on the universe's cardinality.

Shallow-cell complexity

This bound of Haussler and Welzl must be close to optimal, since—as we noted previously—a lower bound of $\Omega(1/\epsilon)$ on the size of ϵ-nets is easily obtained for natural geometric systems. In terms of VC dimension only it is tight: for every integer $d \geq 2$ there are systems with VC dimension at most d whose ϵ-nets all have $\Omega((d/\epsilon)\log(1/\epsilon))$ elements [68]. However it was observed via various specialised constructions that some set systems admit smaller nets: see Table 0.1.

It was believed for some time that $O(1/\epsilon)$ should be the correct bound for all set systems that beat (1). However in 2010 Alon [6] gave a lower bound on the size of

Objects	**Space**	Upper bound on the smallest ϵ-nets
Intervals	$\mathbb{R}$	$1/\epsilon$
Half-spaces	$\mathbb{R}^2, \mathbb{R}^3$	$O(1/\epsilon)$
Disks	$\mathbb{R}^2$	$13.4/\epsilon$
Axis-parallel rectangles	$\mathbb{R}^2$	$O((1/\epsilon)\log\log(1/\epsilon))$

Table 0.1: Some systems with nets smaller than Haussler and Welzl's upper bound (1).

nets for the very simple example of lines in $\mathbb{R}^2$ that is (barely) super-linear in $1/\epsilon$. Alon's lower bound has been improved recently but is still roughly a logarithmic factor away from the upper bound. This was followed by some more strongly super-linear lower bounds of Pach and Tardos [91], such as $\Omega((1/\epsilon)\log\log(1/\epsilon))$ for axis-parallel rectangles (rectangles with their sides horizontal and vertical) in $\mathbb{R}^2$.

To understand this behaviour one has to introduce finer measures of complexity, usually rooted in geometric properties.

A first example is the *union complexity* of a class of geometric regions, defined as the maximum number of faces in the boundary of the arrangement of any n members, as a function of n. See the 'State of the Union' survey of Agarwal, Pach, and Sharir [4] for further details and the papers by Clarkson and Varadarajan [39] and Varadarajan [98] that highlight a connection between union complexity and size of nets.

Shallow cell complexity extends the idea of union complexity and is the most successful complexity measure to date. Again, consider disks in $\mathbb{R}^2$. From any set of n points in general position, $\Theta(n^3)$ different subsets can be cut out by a disk, but we also have some information on their cardinalities: most are large. More precisely, only $O(nl^2)$ of them have fewer than l elements for each $l \in [1 \mathinner{..} n]$. For various geometric set systems, similar bounds on the distribution function of cardinalities are known. Generally speaking, we say that the system has shallow cell complexity $\tilde{\varphi}\colon \mathbb{N} \mapsto \mathbb{R}_+$ when the number of ranges of size at most l induced on n elements is $n \cdot \tilde{\varphi}(n) \cdot l^{O(1)}$. Thus $\tilde{\varphi}(n) = O(1)$ for disks. When the VC dimension is bounded—or equivalently $\tilde{\varphi}(n) = O(n^d)$ for some constant d—the system admits ϵ-nets as small as

$$O\left(\frac{1}{\epsilon}\log\tilde{\varphi}\left(\frac{1}{\epsilon}\right)\right). \tag{2}$$

(This result as it appears in both Chan et al. [32] and Varadarajan [99] actually extends to a more general setting, where the universe's elements have positive weights and (2) is the fraction of the total weight carried by the lightest ϵ-net.) In particular (2) refines the Haussler–Welzl bound (1), which is obtained from it by injecting $\tilde{\varphi}(n) = O(n^d)$. Observe that (2) is more precise in some cases such as that of disks, for which it yields $O(1/\epsilon)$. See Chapter 1 for further information about the combinatorics of set systems and nets.

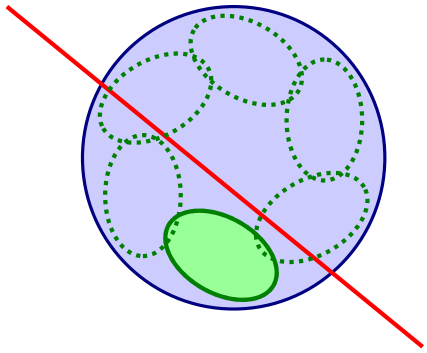


Figure 0.2: Macbeath regions of an Euclidean ball. Each cap containing a fraction ϵ of the ball's volume includes one.

Combinatorial Macbeath regions

The object that we investigate is an extension of nets, adding a thickness constraint to the intersection requirement: rather than to consist of separate elements (or singletons), our cover must consist of large subsets of the ground set. Formally, given a set system $(X, \mathcal{R})$ with a distinguished finite subset $Y \subseteq X$, consider a collection $\mathcal{M} \subseteq \mathcal{P}(Y)$ of subsets of Y that are relatively large:

$$\forall M \in \mathcal{M},\ |M| \geq C\epsilon\,|Y|\,,$$

with C a positive constant, and that 'hit' every large range:

$$\forall R \in \mathcal{R},\ |R \cap Y| \geq \epsilon\,|Y| \implies \exists M \in \mathcal{M} : M \subseteq R.$$

Yet again the intuition for this structure comes from geometry—and so does its name. Brönnimann, Chazelle, and Pach [23], based on a classical result of Macbeath [74], show how to obtain a small family of relatively large convex regions of any given convex body such that every large cap of the body must include one of the regions (see Figure 0.2, note that we call *Macbeath regions* any objects with these properties regardless of their construction). Recently these Macbeath regions have been used in the design of algorithms for polytope approximation [11]. By analogy, the discrete object $\mathcal{M}$ is called a collection of combinatorial Macbeath regions or an Mnet.

In their paper introducing Mnets, Mustafa and Ray [86] gave a few bounds (see Table 3.1 on page 61), using different techniques for various kinds of geometric set systems. For example, they showed that the primal system defined by lines in $\mathbb{R}^2$ had ϵ-Mnets of cardinality $O((1/\epsilon)^2 \cdot \log^2(1/\epsilon))$, whereas that defined by half-planes has ϵ-Mnets as small as $O(1/\epsilon)$. They observed an apparent relation between the upper bounds on nets and Mnets. Notably, systems with linear-sized—i.e. $O(1/\epsilon)$—ϵ-nets also had linear-sized ϵ-Mnets. Generally speaking their bounds tended to be $O((1/\epsilon) \cdot$

Objects of $\mathbb{R}^2$	$\tilde{\varphi}(n)$	Smallest ϵ-nets	Smallest ϵ-Mnets
Lines	$O(n)$	$(2/\epsilon) \cdot \log(1/\epsilon)$	$O((1/\epsilon)^2 \cdot \log^2(1/\epsilon))$
Disks	$O(1)$	$O(1/\epsilon)$	$O(1/\epsilon)$
Half-Planes	$O(1)$	$O(1/\epsilon)$	$O(1/\epsilon)$

Table 0.2: Net–Mnet correlation observed by Mustafa and Ray: some examples. In particular, linear-sized nets correspond to linear-sized Mnets.

$\tilde{\varphi}(1/\epsilon))$ for set systems of shallow-cell complexity $\tilde{\varphi}$—sometimes up to polylogarithmic factors in $1/\epsilon$.

We are able to confirm this intuition with the following theorem (Theorem 3), which extends these results.

Theorem (Mnets, informal)**.** *Set systems defined on an Euclidean space by some constant number of bounded-degree polynomial constraints with VC dimension d and shallow cell complexity $\tilde{\varphi}$ a non-decreasing function have ϵ-Mnets of size*

$$O\left(\frac{d}{\epsilon} \cdot \tilde{\varphi}\left(\frac{8d}{\epsilon}\right)\right). \tag{3}$$

This theorem unifies previous bounds on Mnets, in some cases with a polylogarithmic improvement, such as that of lines in $\mathbb{R}^2$ for which the new bound is $O(1/\epsilon^2)$.

Algebraic tools.

In the statement of this theorem, we restrict ourselves to *semi-algebraic objects*, such as polytopes with a bounded number of sides or zero-sets of bounded-degree polynomials. This is because our proof—like many recent results in incidence geometry—relies on algebraic techniques, namely a recursive form of the celebrated polynomial partitioning approach introduced by Guth and Katz [54]. At the same time, and as in other results derived with such tools, there is no clear reason why these restrictions should be necessary.

Conjecture. The same bound (3) holds without the algebraicity hypothesis.

Packing numbers.

Besides polynomial partitioning, the proof of the Mnet bound (3) also involves results about packings in set systems (i.e. metric entropy with respect to Hamming distance).

We gather them in Chapter 2, whereas the discussion on Mnets proper is in Chapter 3. Here it may help to view a set system on n elements as a subset of the Boolean hypercube $\{0, 1\}^n$, with the ranges as its elements. The original *packing lemma* of Haussler [59] is as follows.

Theorem (packing lemma, informal). *In a set system on n elements with VC dimension d, a δ-packing of ranges (a family of ranges of which any two differ on at least δ elements) has $O((n/\delta)^d)$ ranges. This bound is tight.*

Without the VC dimension hypothesis, a δ-packing on n elements could have $2^{n/\delta}$ ranges; with the introduction of said hypothesis the bound goes from exponential to polynomial in (n/δ). If we additionally require that all ranges have at most k elements, Haussler's bound can be further strengthened [44, 84].

Theorem (shallow packing lemma). *In a set system on n elements with VC dimension d that for every l has $s(n, l)$ ranges of size at most l, every δ-packing of ranges all of size at most k has at most*

$$6 \cdot s\left(4d \cdot \frac{n}{\delta}, 12d \cdot \frac{k}{\delta}\right) \tag{4}$$

ranges. In particular when $s(n, l) = O(n^{d_1} l^{d_2})$ this bound is

$$O\left(\frac{n^{d_1} k^{d_2}}{\delta^{d_1+d_2}}\right). \tag{5}$$

This is the version that is used in the proof of our main theorem on Mnets. Note that s in the statement of this theorem is almost the shallow-cell complexity: $s(n, l) = n \cdot \tilde{\varphi}(n) \cdot l^{O(1)}$.

In passing, we observe that this variant of Haussler's lemma is still asymptotically tight and give an explicit construction of set systems that achieve the bound (5) (Theorem 1).

Sketching the proof of the Mnet upper bound.

Now that we have introduced algebraicity and packing lemmas, here are the main ideas of the proof for the Mnet theorem. Suppose that all ranges have cardinality ϵn and take an inclusion-maximal $(\epsilon n/2)$-packing $\mathcal{P}$ among them. Its cardinality $|\mathcal{P}|$ is controlled by the shallow packing lemma. Then it can be seen that it suffices to build a $(1/2)$-Mnet $\mathcal{M}_R$ for every range $R \in \mathcal{P}$ with respect to the system induced by all other ranges on R. To do so we use polynomial partitioning. The cardinality of the resulting Mnet is then at most $|\mathcal{P}| \cdot \sup\{|\mathcal{M}_R| : R \in \mathcal{P}\}$. Of course, the actual proof requires some additional work to iron out all difficulties.

Tightness of the Mnet bound.

Our upper bound (3) on the cardinality of a smallest Mnet is tight in terms of VC dimension alone, when it is only known that $\tilde{\varphi}(n) = O(n^{d-1})$. Specifically, it is tight for the set system defined by hyperplanes in $\mathbb{R}^d$ (which has VC dimension d): its ϵ-Mnets must have cardinality $\Omega(\epsilon^{-d})$ (Theorem 4) while the upper bound gives an ϵ-Mnet of cardinality $O(\epsilon^{-d})$. We also have a close, albeit not tight, lower bound in terms of shallow-cell complexity (Theorem 5).

A return to nets.

As a testimony of the deep connection between Mnets and epsilon-nets, the usual bounds on nets (2) are obtained again by a simple probabilistic argument. Consider a set system with an ϵ-Mnet $\mathcal{M}$. Picking at least one element from each $M \in \mathcal{M}$ yields an ϵ-net, because every range containing a fraction ϵ of all elements includes some $M \in \mathcal{M}$. We observe (Corollary 3.1) that this can be done with

$$O\left(\frac{\log(\epsilon\mathcal{M})}{\epsilon}\right) + \frac{1}{\epsilon}$$

elements in total, by random sampling, so that in particular the systems to which (3) applies admit ϵ-nets of cardinality

$$O\left(\frac{1}{\epsilon}\log\tilde{\varphi}\left(\frac{8d}{\epsilon}\right)\right),$$

with a constant depending on d and the algebraic parameters. This is a new derivation of the standard bound (2) on nets which does not rely on the results of Haussler and Welzl.

The reference for the results in this first part is

[45] Kunal Dutta, Arijit Ghosh, Bruno Jartoux, and Nabil H. Mustafa. 'Shallow Packings, Semialgebraic Set Systems, Macbeath Regions, and Polynomial Partitioning'. In: *33rd International Symposium on Computational Geometry (SoCG 2017).* Vol. 77. 2017, 38:1–38:15.

Interlude: Nets and Minimum Hitting Set

Minimum Hitting Set is a combinatorial optimisation problem in which one is given a set system $(X, \mathcal{R})$ with $\emptyset \notin \mathcal{R}$ and $|X|, |\mathcal{R}| \leq n$ and must compute a smallest $H \subseteq X$ that intersects all $R \in \mathcal{R}$. This is **NP**-hard. When all ranges have cardinality 2, the set system is a graph and the problem becomes Minimum Vertex Cover, which is still **NP**-hard. (The decision versions of both are among Karp's 21 **NP**-complete problems [65].)

The greedy algorithm to solve this problem selects at each stage an element of X that hits the most uncovered ranges. It achieves an approximation ratio (slightly better than) $1 + \log n$, meaning that it always returns a solution at most $1 + \log n$ times larger than the optimum.

On the other hand, unless some unexpected things happen in computational complexity—unless $\mathbf{P} = \mathbf{NP}$—there cannot be a polynomial-time algorithm that decreases this ratio to $(1 - \delta) \log n$ for any positive δ [42]. That is, for every $\delta > 0$, it is **NP**-hard to approximate Minimum Hitting Set within a factor $(1 - \delta) \log n$. Thus the question of designing efficient approximations for the general Minimum Hitting Set problem is almost optimally resolved.

However, Brönnimann and Goodrich [24] observed that this lower bound can be beaten on set systems with small ϵ-nets. (Recall that by (1) this is the case in set systems whose VC dimension is bounded). They showed that efficient computation of small nets is sufficient to achieve a polynomial-time approximation for Minimum Hitting Set. More precisely, where ϵ-nets of cardinality $1/\epsilon \cdot f(1/\epsilon)$ can be computed their algorithm achieves an $O(f(\text{OPT}))$-approximation for Minimum Hitting Set, i.e. it returns a set of at most $\text{OPT} \cdot f(\text{OPT})$ elements where OPT would have been the size of a smallest solution. (This requires f to be sub-linear and nondecreasing, which it is in practice.)

For example, consider the geometric restriction of the problem where $(X, \mathcal{R})$ is induced by disks on $X \subset \mathbb{R}^2$. Disks have ϵ-nets of size $O(1/\epsilon)$, so the algorithm of Brönnimann and Goodrich achieves a $O(1)$-approximation, i.e. within a constant factor of the optimum. Agarwal and Pan [5] even showed that this could be made to run in near-linear time. Thus the geometric restriction of Hitting Set to points and disks, while still **NP**-hard [62] even when all disks have the same radius, has much better approximability properties than the general problem.

The constant approximation factor depends on the constant in the size of optimal nets. There's the rub—only recently has the best bound for disks been brought down to $13.4/\epsilon$ [26], which is still a large constant. In any case this technique cannot achieve arbitrarily small approximation factors. To overcome this limitation one has to turn to another algorithmic tool: local search.

Local Search Techniques for Approximation Algorithms

The second part of this thesis deals with the use of local search (a class of optimisation algorithms) in geometric combinatorial optimisation.

Approximation algorithms

We focus on the geometric restriction of Minimum Hitting Set described previously, and similar geometric versions of classic optimisation problems (Independent Set, Dominating Set, etc.). In spite of the geometric restrictions these problems are still **NP**-hard, so that they are unlikely to admit an efficient exact solution (under standard complexity-theoretic assumptions).

Furthermore, some results from the field of parametrised complexity also suggest that there is no f for which an approximation ratio of $1 + \epsilon$ could be obtained in time $f(1/\epsilon) \cdot n^{O(1)}$ for all $\epsilon > 0$. Specifically we observe (page 76) that the decision versions of these problems are **W[1]**-hard and thus are not fixed-parameter tractable unless **W[1]** $=$ **FPT**. Then a result of Bazgan [19, see Theorem I] implies that—again unless **W[1]** $=$ **FPT**—the problems cannot have 'efficient PTASs'. That is, as $\epsilon \downarrow 0$ the exponent of n has to depend on $1/\epsilon$. Typically we could hope for a family of algorithms running in $n^{O(1/\epsilon)}$ for every small positive ϵ. Details on approximation algorithms and hardness results are gathered in Chapter 4.

Principle of local search

Given a combinatorial optimisation problem—such as Minimum Hitting Set on a finite set system—local search is a class of iterative procedures that step-wise modify an initial solution, always maintaining feasibility. At each step, either a new feasible solution with a better objective value can be found in the neighbourhood of the current solution, or the procedure terminates.

In all problems studied here, we are given a finite ground set X and feasible solutions form a subset $\mathcal{S}$ of $\mathcal{P}(X)$. The value of a feasible solution is always its cardinality, which we either want to minimise or maximise depending on the problem. Thus it is natural to equip the solution space with the Hamming distance

$$(s_1, s_2) \in \mathcal{S}^2 \mapsto |s_1 \,\Delta\, s_2|\,,$$

and to say that $s \in \mathcal{S}$ is λ-locally optimal when the Hamming ball of of radius 2λ centred at s contains no better solution.

To fix ideas, let us assume that we are dealing with a minimisation problem. The specific kind of local search that we are studying consists only in finding a swap that

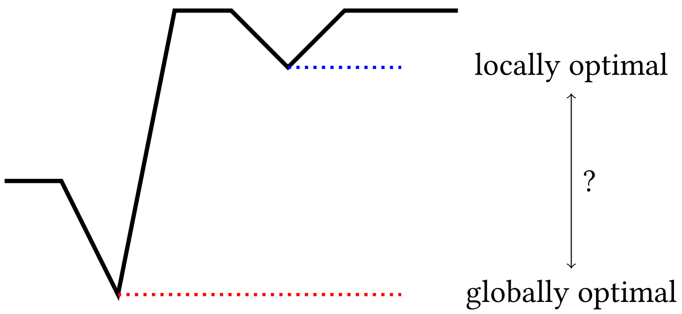


Figure 0.3: The *locality gap* of optimisation problems may be arbitrarily large.

involves up to $2\lambda - 1$ elements of X and strictly decreases the solution's cardinality, that is, replace $t \in [1 .. \lambda]$ elements from the current solution with up to $t - 1$ that did not belong to it.

Since each such step strictly improves the current solution and there are only $n = |X|$ possible solution values, this procedure terminates in at most n steps. At each of them one has to enumerate $\binom{n}{2\lambda}$ candidate swaps, check their feasibility and compute their value, all of which is done in time $n^{O(\lambda)}$. Thus the total cost of finding a λ-local optimum is $n^{O(\lambda)}$.

Combinatorics of local search

In general, local search does not come with any guarantees on the value of locally optimal solutions compared with the global optimum—deep valleys in the Himalaya are still much higher than in the Mariana Trench (Figure 0.3). This lack of guarantees on the 'locality gap' makes it a *heuristic* rather than an approximation algorithm and is the reason for the development of *meta-heuristics* such as tabu search or simulated annealing. Rather surprisingly, it was observed by Agarwal and Mustafa [3] that local search *does* achieve constant-factor approximation for some problems on geometric inputs, and then simultaneously by Chan and Har-Peled [33] and Mustafa and Ray [87] that this constant factor could be made arbitrary small, i.e. $1 + \epsilon$ for any $\epsilon > 0$.

This has since been extended by other authors to various geometric optimisation problems [non-exhaustive list: 15, 28, 31, 46, 52, 53, 57]. The common idea of all these papers is the construction of a planar (or otherwise sparse) 'exchange graph' that encapsulates the interplay of any two feasible solutions. Roughly speaking, it encodes a rule by which elements from the first solution may be replaced with others from the second in such a way that this modified first solution remains feasible. See for example the construction for geometric Minimum Hitting Set in Figure 0.4.

When one of the solutions is locally optimal, the fact that it cannot admit 'small' improvements translates into a specific combinatorial property of the graph. Say that a

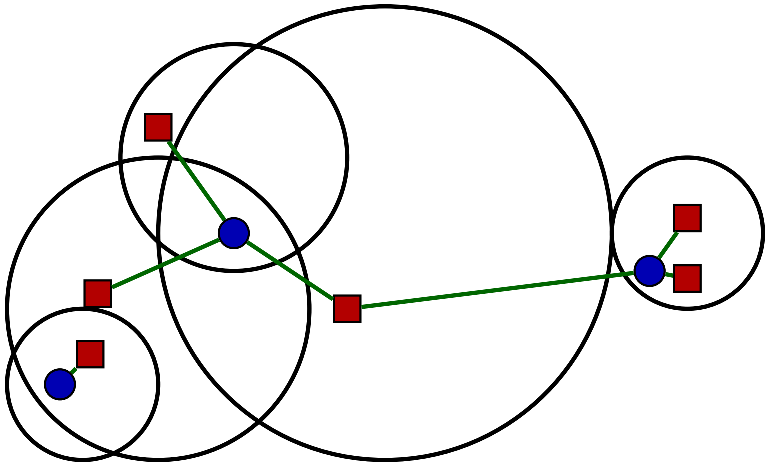


Figure 0.4: Exchange graph construction for MINIMUM HITTING SET of disks. Given two feasible solutions, this *bipartite Delaunay graph* has an edge between two vertices from different solutions if and only if some disk of $\mathbb{R}^2$ contains them and no other vertex from either solution. In this way any disk of the input, being hit by both solutions, must include at least one edge of the exchange graph.

bipartite graph is k-expanding if each subset of up to k vertices on the left participates in a matching. By Hall's marriage theorem (see Appendix A) an equivalent property is that every $t \leq k$ left vertices have at least t neighbours. Such a graph, if it is also planar, cannot be too unbalanced.

Theorem. *If a bipartite planar graph on left and right vertex sets L and R is k-expanding, $k \geq 3$, then*

$$|L| \leq \left(1 + \frac{c}{\sqrt{k}}\right) |R| \tag{6}$$

with c a positive constant.

A simple consequence for MINIMUM HITTING SET for disks of $\mathbb{R}^2$ (and similar problems) is that any $\Theta(\epsilon^{-2})$-locally optimal solution is at most $(1+\epsilon)$ times larger than the global optimum, because Mustafa and Ray [87] construct a $\Theta(\epsilon^{-2})$-expanding planar bipartite *exchange graph* with these two solutions as vertex sets (see Figure 0.4). In other words, local search is a *polynomial-time approximation scheme* that computes a $(1+\epsilon)$-approximate solution in time $n^{O(\epsilon^{-2})}$. This has been applied to other problems, usually with exchange graphs based on a modified Delaunay property.

The 2 in the exponent of ϵ comes from the square root on k in (6), which itself is a consequence of the planar separator theorem of Lipton and Tarjan [72].

Theorem (planar separators). *In every planar graph of order $n \in \mathbb{N}$ there is a set of $O(\sqrt{n})$ vertices whose removal only leaves connected components of order at most $2n/3$.*

Briefly, recursive application of this separator theorem splits the graph into $O(n/k)$ blocks, each of cardinality k with $O(\sqrt{k})$ shared boundary vertices between them. (This construction follows an idea of Frederickson [50].) Then in each block the k-expansion property indicates that there must be at least as many vertices of R as of L. This is then summed over the whole graph with a corrective term that accounts for the multiple counting of boundary vertices and yields (6).

We show that the inequality (6) is tight: up to the value of the constant c the asymptotic rate of convergence in $1/\sqrt{k}$ cannot be improved (Theorem 6).

Theorem. *For any large integer k, there is a family of k-expanding bipartite planar graphs on vertex sets B_n and R_n where $n \in \mathbb{N}$ such that*

$$|B_n|, |R_n| = \Theta(n) \text{ and } |B_n| \sim \left(1 + \frac{c'}{\sqrt{k}}\right) |R_n| \text{ as } n \to +\infty,$$

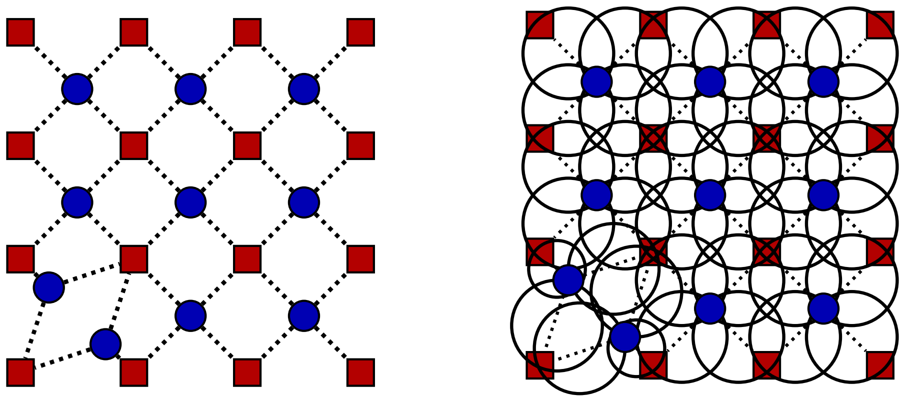
with $c' > 0$ an absolute constant.

Here is the idea of the proof. A square grid is seen as a bipartite planar graph with the bipartition defined by the parities of the coordinates. Its vertex sets (say blue and red) have approximately the same cardinality, even more so when the grid is large. It has good expansion properties, i.e. any subset of blue vertices far from the grid's boundaries can be extended into a matching. Periodically duplicating roughly one in $\sqrt{k}$ blue vertices (as in Figure 0.5a) restricts these expansion properties, but the graph remains $\Theta(k)$-expanding. This is the crux of the proof; it is shown by combining a *charging* argument with an isoperimetric one.

Finally observe that there are in the resulting graph $1 + \Theta(k^{-1/2})$ times as many blue vertices as there are red ones.

Practical consequences for existing algorithms

As the construction of these graphs is explicit, they can be transformed into instances of various combinatorial optimisation problems (such as Minimum Hitting Set, Minimum Dominating Set or Maximum Independent Set for arrangements of disks). For example, we give an instance of Minimum Hitting Set with $\Theta(n)$ points and $\Theta(n)$ disks that has a λ-locally optimal solution $1 + \Theta(1/\sqrt{\lambda})$ times larger than its optimal solution. This shows that the *locality gap* of these problems is $1 + \Theta(1/\sqrt{\lambda})$ and that the ϵ^{-2} in the running time exponent is an inherent limitation of the algorithm rather than an artefact of the analysis, answering an open question from [52, 87].



(a) Detail of the locally-expanding graph that attains the bound.

(b) Part of a 'tight' instance for the hitting set problem based on the lower-bound graph's Delaunay structure.

Figure 0.5: Lower bound construction and algorithmic consequences

Small radii

Note that the inequality (6) can be improved for small values of k. The sharpest possible specialised versions are $|L| \leq 8\,|R|$ for $k = 3$ [8] and $|L| \leq 4\,|R|$ for $k = 4$ [25]. Like ours, the lower bounds are obtained by graph constructions and translate into sharp upper bounds on the approximation ratio of local search—with radii 3 and 4 rather than some large λ. See for example the construction in Figure 0.6.

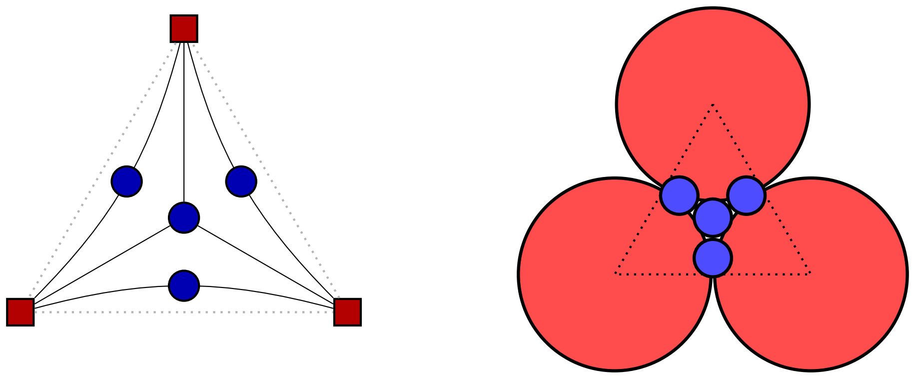


Figure 0.6: Lower bound construction for radius-3 local search on Independent Set of disks based on the graph in [25]. The graphs that asymptotically attain $|L| = 4\,|R|$ are obtained by tiling the plane with the triangular patch on the left (L is blue and R is red). Instances are built by realising these graphs as intersection graphs of disks.

Non-planar extensions.

Several authors have noticed that the inequality (6) could be extended to bipartite graphs that, while not necessarily planar, still possess strong separator properties. Indeed the analysis only requires that any subgraph on $m \in \mathbb{N}$ vertices has a separator of cardinality $O(\sqrt{n})$. Thus the result also applies to graphs of bounded genus and more generally to graphs excluding a fixed minor [28]. (Planar graphs are exactly the graphs of genus 0, and also the graphs that exclude the 5-clique and the $(3,3)$-biclique as minors.) Finally, if one considers instead graphs with separators in $O(m^{1-s})$ for some fixed $s \in (0,1)$ then the analysis still goes through [15, 57] and the inequality becomes

$$|L| \leq \left(1 + \frac{c_s}{k^s}\right) |R| \,, \tag{7}$$

with c_s a constant depending on s.

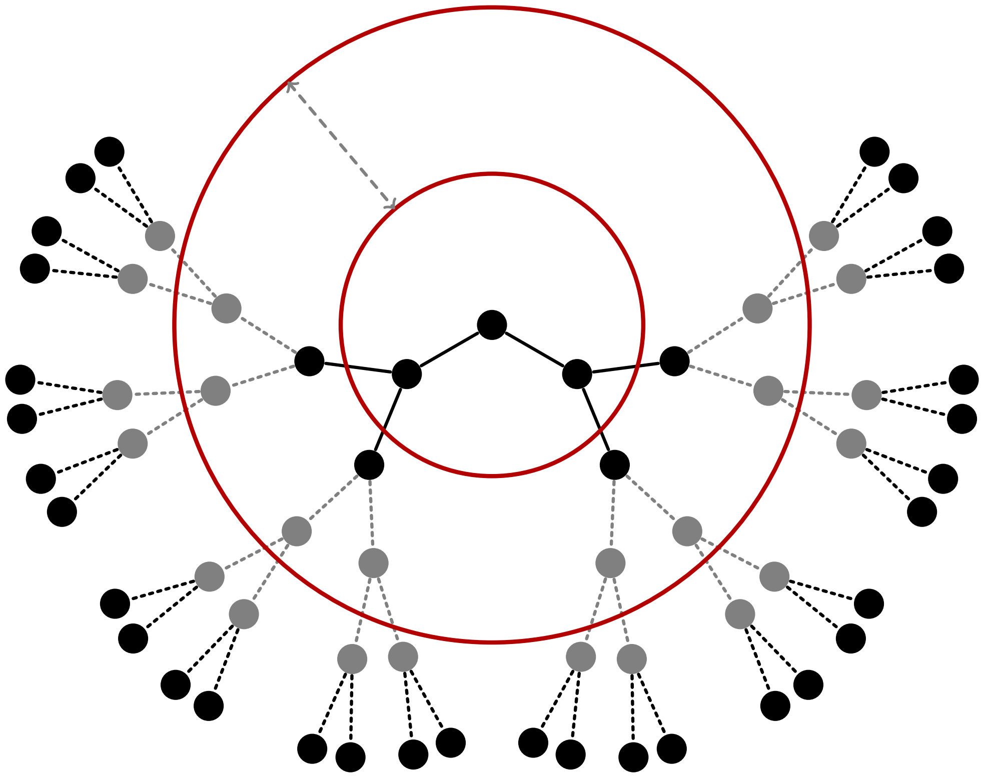
Conversely our lower bound construction extends to dimension $d > 2$ by perturbing a d-dimensional grid rather than a 2-dimensional one, meaning that (7) is also tight at least for $s \in \{1/d : d \in \mathbb{N}^*\}$.

The reference for the results in this second part is

[64] Bruno Jartoux and Nabil H. Mustafa. 'Optimality of Geometric Local Search'. In: *34th International Symposium on Computational Geometry (SoCG 2018)*. 2018. Forthcoming.

Part I

Sampling Geometric Set Systems



Chapter One

Set Systems and Geometry

> He didn't see the expression of the shopman's face, but to smooth over the awkwardness of the position a little he felt called upon to make some purchase. But what should he buy? He looked round the walls of the shop to pick out something inexpensive, and his eyes rested on a green net hanging near the door.
>
> Chekhov, *An Avenger* (trans. Garnett).

BASIC NOTIONS that will be required afterwards are introduced in this chapter. We discuss set systems and several descriptions of their combinatorial complexities, some finer than others, then we define basic operations on them. Finally we consider two simple structures related to the problem of sampling these systems: transversals and nets. Only Definition 1.13 is non-standard.

1 Set Systems

1.1 Definition

The basic framework within which the following work takes place is that of a *set system*. As its definition is quite elementary, this object appears under various names in many different fields in the computational sciences.

Definition 1.1. A **set system** is a pair $(X, \mathcal{R})$ where X is a set and $\mathcal{R} \subseteq \mathcal{P}(X)$, that is, $\mathcal{R}$ is a set of subsets of X.

Borrowing from the vocabulary of range searching, we call $\mathcal{R}$ the **range set** and its elements **ranges**, while X is the **ground set**. We say that $(X, \mathcal{R})$ is **finite** when X is, in which case there are at most $2^{|X|}$ ranges.

Some other names for set systems include range spaces, hypergraphs, incidence structures or Lévy graphs. Objects such as (abstract) simplicial complexes, binary block codes, bipartite graphs with a fixed two-colouring and classifiers in learning theory all have a natural set-system structure.

1.2 Duality

Definition 1.2. The system whose ground set is $\mathcal{R}$ and whose ranges are the pencils defined by points of X, that is, with range set

$$\{\{R \in \mathcal{R} : x \in R\} : x \in X\}$$

is the **dual set system** of $(X, \mathcal{R})$, which is then the corresponding **primal system**.

This involutive duality is similar to the point–line duality of projective planes—as a matter of fact, projective planes are examples of set systems.

2 Geometric Set Systems

The most natural way for a geometer to obtain a set system is to consider the incidences of two families of geometric objects, especially of a space with some of its regions. For example, there is a set system whose ground set is the Euclidean plane and whose ranges are disks. It is tersely referred to as the primal set system defined by disks, omitting a mention of its ground set. In the same way one defines the primal set system of half-planes (or half-spaces in higher dimension), of triangles (simplices), of lines (affine subspaces), etc. Each has its corresponding dual system, whose ground set consists of shapes in Euclidean space and whose ranges are pencils of shapes with a common intersection.

Set systems help formulate many geometric problems from range reporting, where one designs space- and time-efficient data structures that answer range queries on a stored point set, to some variants of facility location problems, where one chooses an economical cover of an input point set from a set of candidate disks.

❧

❧ ❧

3 Combinatorial Complexity of Set Systems

Although all elements of $\mathcal{P}(X)$ could be ranges in a set system on X, the examples that arise in geometry often exhibit sparsity properties—a behaviour that extends for example that of planar graphs. We present several descriptions of those properties.

3.1 Vapnik–Chervonenkis Dimension

In any set of three non-collinear points in $\mathbb{R}^2$, all 2^3 subsets can be realised as intersections with a half-plane. However, in every set of four points of $\mathbb{R}^2$ some subset cannot be cut out in this way: either one point is a convex combination of the others and is thus contained in any half-plane that contains them, or all four points are in convex position and the diagonal pairs cannot be realised. This means that the set system defined by half-planes on a ground set $X \subseteq \mathbb{R}^2$ of four points or more cannot have all of $\mathcal{P}(X)$ as ranges.

Definition 1.3. For a set system $(X, \mathcal{R})$, the **trace** of the range set on any $Y \subseteq X$ is the set

$$\mathcal{R}_{|Y} = \{R \cap Y : R \in \mathcal{R}\} \subseteq \mathcal{P}(Y)$$

and Y is **shattered** (by $\mathcal{R}$) if $\mathcal{R}_{|Y}$ equals $\mathcal{P}(Y)$. The **VC dimension** of $(X, \mathcal{R})$ is the supremum of the cardinalities of shattered subsets of X.

The concept of trace extends that of induced subgraph: the trace is the range set induced by $\mathcal{R}$ on a subset of X. In keeping with this analogy, a shattered subset is one whose induced subsystem is 'full', not unlike cliques in graphs. VC dimension was introduced by Vapnik and Chervonenkis [97] to characterise the 'expressiveness' of a set system. In learning theory, ranges represent classifier sets. A shattered set is one for which all classifications are possible.

Example 1.1. The primal set system of half-planes has VC dimension 3. Every triple of points (not on a same line) in general position can be shattered. On the other hand, consider four distinct points in the plane. Either one of them lies inside the convex hull of the others and thus the corresponding singleton cannot be realised with a half-space, or the four points are in convex position and two of the six pairs of points cannot be realised either.

Example 1.2. The VC dimension of the primal set system of convex sets in $\mathbb{R}^2$ is $+\infty$. All subsets of a circle can be realised by their convex hulls, see Figure 1.1.

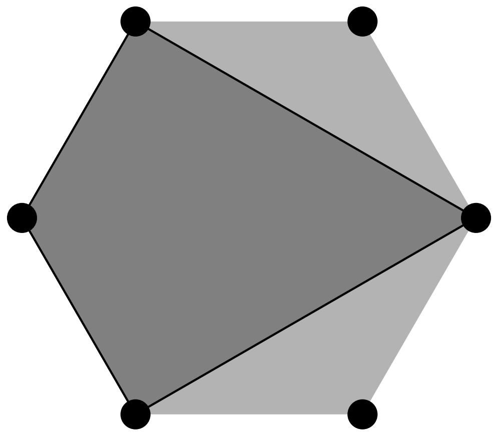


Figure 1.1: The VC dimension of the primal set system of convex sets in $\mathbb{R}^2$ is $+\infty$ as each subset of a set of points in convex position may be cut out by some convex shape.

3.2 Shatter Function

Now we observe the behaviour of the trace of the range set on Y as a function of Y's cardinality.

Definition 1.4. The **shatter function** of $(X, \mathcal{R})$ is the function

$$s_{\mathcal{R}}\colon m \in \mathbb{N} \mapsto \max\left\{\left|\mathcal{R}_{|Y}\right| : Y \subseteq X \text{ and } |Y| \leq m\right\}.$$

Remark 1.1. The shatter function is non-decreasing. Additionally it does not depend on X—the subset Y can always be taken in the union $\cup\mathcal{R}$ of all ranges. So if $\cup\mathcal{R}$ has more than m elements—say it is infinite—the maximum is simply over $Y \in \binom{\cup\mathcal{R}}{m}$.

The quantity $s_{\mathcal{R}}(m)$, or simply $s(m)$ when $\mathcal{R}$ is clear from the context, is sometimes known as the m-th shatter coefficient and is always upper-bounded by 2^m since $\mathcal{R}_{|Y}$ is a subset of $\mathcal{P}(Y)$. In particular shattered sets are those Y for which this bound is attained (recall Definition 1.4) and the VC dimension is the supremum of all $m \in \mathbb{N}$ for which $s_{\mathcal{R}}(m) = 2^m$.

Example 1.3. Consider the primal set system of half-planes in $\mathbb{R}^2$, and let Y consist of m points of $\mathbb{R}^2$. Any range with at least two elements can be realised by a half-plane whose boundary contains (at least) two points of Y. (This can be seen by rotating and translating any half-plane while preserving its intersection with Y.) Thus the trace $\mathcal{R}_{|Y}$ contains at most $\binom{m}{2} + m + 1$ ranges and the system has $s(m) = O(m^2)$.

A more precise relation between shatter function and VC dimension is given by the following lemma. It was independently established by Norbert Sauer, by Saharon Shelah (crediting Micha Perles), and by Vladimir Vapnik and Alexey Chervonenkis [93, 95, 97].

Lemma 1.1. *For a set system whose VC dimension is $d \in \mathbb{N}$,*

$$\forall m \in \mathbb{N}, \quad s(m) \le \sum_{i=0}^{d} \binom{m}{i}.$$

The inequality is tight. For example, the system on $\mathbb{N}$ with range set

$$\bigcup_{i=0}^{d} \binom{\mathbb{N}}{i}$$

has VC dimension d since every set of d integers is shattered and no set of $d+1$ can be, and it attains this bound. See Example 1.4 for a geometric system that also saturates it. The take-away of this lemma is that the right-hand side is less than $(em/d)^d$ for $m, d \ge 1$ and thus the shatter function is polynomially bounded, in sharp contrast with the general bound of 2^m.

Example 1.4. Consider the primal set system of lines in $\mathbb{R}^2$. The VC dimension of this system is 2, so it follows from Lemma 1.1 that $s(m) \le \binom{m}{2} + m + 1$. This bound is sharp: any set of m points in the plane, no three of them collinear, attains it.

3.3 Shallow-Cell Complexity

Common geometric set systems have finite VC dimension (convex sets being the main counterexample) and thus satisfy Lemma 1.1: their shatter function grows polynomially rather than exponentially. However, these systems often also satisfy a finer property—not only is the size of $\mathcal{R}_{|Y}$ polynomially bounded, but also the number of *small* sets in $\mathcal{R}_{|Y}$ is lower.

Example 1.5. Let $(\mathbb{R}^2, \mathcal{R})$ be the primal set system induced by disks. For any finite $Y \subset \mathbb{R}^2$ and $l \in \mathbb{N}$, the number of sets in $\mathcal{R}_{|Y}$ of size at most l is $O(|Y| \cdot l^2)$. For small values of l, this contrasts with the total cardinality of $\mathcal{R}_{|Y}$, which can be $\Theta(|Y|^3)$.

This motivates a finer classification of set systems, which was gradually introduced over several papers [32, 44, 47, 99]. Given a system $(X, \mathcal{R})$, define size-sensitive versions of its trace and shatter function:

$$\mathcal{R}_{|Y,\le l} = \{S \in \mathcal{R}_{|Y} : |S| \le l\} = \{R \cap Y : R \in \mathcal{R} \text{ and } |R \cap Y| \le l\},$$
$$s_{\mathcal{R}} \colon (m, l) \in \mathbb{N}^2 \mapsto \max \left\{ \left|\mathcal{R}_{|Y,\le l}\right| : Y \subseteq X \text{ and } |Y| \le m \right\}.$$

This extends Definition 1.3 and Definition 1.4: clearly $\mathcal{R}_{|Y,\le|Y|} = \mathcal{R}_{|Y}$ for every subset Y and $s_{\mathcal{R}}(m) = s_{\mathcal{R}}(m, m)$ for every natural number m. Again, we write s for $s_{\mathcal{R}}$ when this does not introduce any ambiguity.

Example 1.6. In [44, 47], a set system was said to satisfy the **Clarkson–Shor property** with parameters $(d, d_1) \in \mathbb{N}^2$ when it had $s(m, l) = O(m^{d_1} l^{d-d_1})$. For example the primal system for disks in $\mathbb{R}^2$ is $(3, 1)$-Clarkson–Shor as it has $s(m, l) = O(ml^2)$. For most natural geometric set systems, the best known upper bounds on s are of this form.

Definition 1.5. The **shallow-cell complexity** (SCC) of a set system $(X, \mathcal{R})$ is the function

$$\begin{aligned} \varphi_{\mathcal{R}} \colon \mathbb{N} \times \mathbb{N} &\to \mathbb{R} \\ (m, l) &\mapsto \begin{cases} 1 & m = 0 \\ \frac{1}{m} \cdot s(m, l) & m \in \mathbb{N}^* \end{cases} . \end{aligned}$$

As for the shatter function, the subscript will be dropped when $\mathcal{R}$ is clear from the context.

The reason for which we normalise by m instead of working directly with $s_{\mathcal{R}}$ is that this makes later results on nets simpler to state. For reasonable geometric set systems $s(m, l)$ is linear or super-linear in m. For example, in the primal system of disks in $\mathbb{R}^2$, some subset Y of m points of $\mathbb{R}^2$ induces $s(m, l)$ ranges of size at most l and $Y \cup \tau(Y)$ will determine at least $2 \cdot s(m, l)$ such ranges for a large enough translation τ: it follows that for disks $s(2m, l) \geq 2 \cdot s(m, l)$.

Often the precise dependency of $\varphi(m, l)$ on l is less important, so we introduce a one-variable SCC.

Definition 1.6. The **one-variable shallow-cell complexity** of a set system $(X, \mathcal{R})$ is the least $\tilde{\varphi} \colon \mathbb{N} \to \mathbb{R}$ such that, for some fixed nonnegative constant c,

$$\forall (m, l) \in \mathbb{N} \times \mathbb{N}, \quad \varphi(m, l) \leq \tilde{\varphi}(m) \cdot l^c.$$

The VC dimensions and (upper bounds on the) SCCs of several basic geometric set systems are gathered in Table 1.1.

3.4 Union Complexity

Union complexity is another measure of complexity specific to geometric set systems and an historical precursor of SCC. It is well-defined as soon as there is a notion of face and cell in a geometric arrangement of the considered objects, hence in particular when these objects are semi-algebraic (see the next subsection for a discussion of semi-algebraicity).

Definition 1.7. The **union complexity** of a family $\mathcal{O}$ of geometric objects is the function that maps each positive integer m to the maximum number of faces of all dimensions in the boundary of the arrangement of any m members of $\mathcal{O}$.

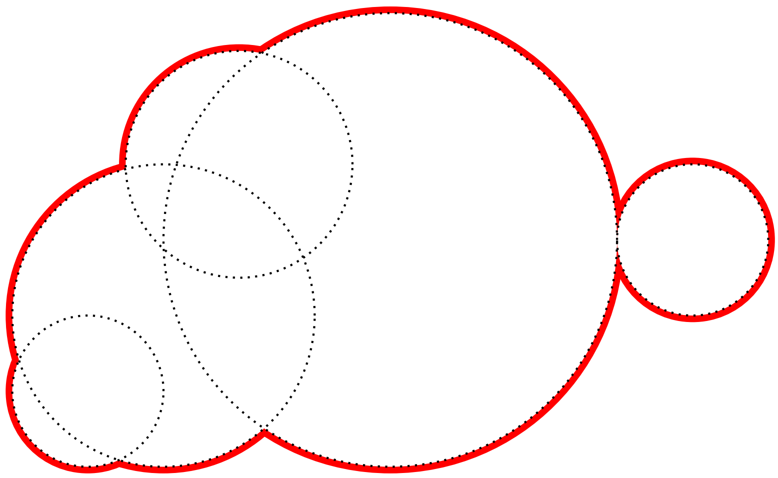


Figure 1.2: Union complexity of disks in $\mathbb{R}^2$. The combinatorial complexity of the boundary (total number of arcs and vertices) in an arrangement of n disks is linear in n.

There is an immediate relation between union and shallow-cell complexities: the dual set system induced by a family of objects with union complexity κ has SCC $\tilde{\varphi}(m) = O(\kappa(m)/m)$ [4, 94].

Example 1.7. Disks in $\mathbb{R}^2$ have linear union complexity: in an arrangement of $n \geq 3$ disks, the boundary is composed of at most $6n - 12$ circular arcs, so that the total number of 1-faces (arcs) and 0-faces (endpoints of said arcs) is linear in n [4, 66], see Figure 1.2. As a consequence the dual system of disks has SCC $O(1)$.

Remark 1.2. The **dual VC dimension**, **dual shatter function** or **dual shallow-cell complexity** of a set system simply refer to the corresponding quantities in its dual.

3.5 Semi-Algebraic Set Systems

Quite frequently, the geometric set systems that we manipulate are associated with regions of the Euclidean space defined by polynomial constraints: this is the case of half-spaces, of balls, of polyhedra, etc.

Definition 1.8. **(Real) semi-algebraic sets** are subsets of $\mathbb{R}^d$ obtained by taking Boolean operations (unions, complements, and intersections) on a finite family of sets of the form $\{x \in \mathbb{R}^d \colon g(x) \geq 0\}$, where g is a d-variate real polynomial.

Example 1.8. The cube $[-1, 1]^d$ of $\mathbb{R}^d$ is the intersection of the slabs $\{1 - x_i^2 \geq 0\}$ over $i \in [1 \,..\, d]$. The unit ball of $\mathbb{R}^d$ is $\{1 - \sum x_i^2 \geq 0\}$.

Semi-algebraic sets can be classified according to the complexity of the formula that defines them. For $(d, \Delta, s) \in \mathbb{N}^3$ we denote by $\Gamma_{d,\Delta,s}$ the family of all semi-algebraic sets in $\mathbb{R}^d$ obtained by taking Boolean operations on at most s polynomial inequalities, each of degree at most Δ.

Definition 1.9. A set system is semi-algebraic with parameters $(d, \Delta, s) \in \mathbb{N}^3$ if its groundset X is a subset of $\mathbb{R}^d$ and its ranges are all of the form $X \cap S$ where $S \in \Gamma_{(d,\Delta,s)}$.

The 'semi-algebraic parameters' of a system control its VC dimension.

Lemma 1.2. *The VC dimension of $\Gamma_{d,\Delta,s}$ as a set system on $\mathbb{R}^d$ is a finite function of (d, Δ, s).*

A precise, almost-tight bound can be found in [20]: at least $\frac{s}{2}\binom{\Delta+d}{\Delta}$ and at most $2s\binom{\Delta+d}{\Delta}\log_2\left(s(s+1)\binom{d+\Delta}{\Delta}\right)$.

Furthermore, the same parameters also control the behaviour of the dual set system.

Lemma 1.3. *There is a function $f : \mathbb{N}^3 \to \mathbb{N}^3$ such that the dual of a (d, Δ, s)-semi-algebraic set system is itself $f(d, \Delta, s)$-semi-algebraic.*

Proof. Let $\mathcal{P}$ consist of all d-variate real polynomials of degree at most Δ. As a vector space, $\mathcal{P}^s \simeq \mathbb{R}^{Ns}$ where $N = \binom{\Delta+d}{d}$. This isomorphism sends a tuple of polynomials $p = (p_1, \ldots, p_s) \in \mathcal{P}^s$ to the vector $\tilde{p} \in \mathbb{R}^{Ns}$ of its coefficients.

For every s-variate Boolean function Φ we let

$$E_\Phi = \{(x, \tilde{p}) \in \mathbb{R}^d \times \mathbb{R}^{Ns} : \Phi((p_i(x) \geq 0)_{1 \leq i \leq s})\}.$$

Observe that for every given x and $i \in [1 \mathinner{..} s]$ the condition $p_i(x) \geq 0$ is linear in the coefficients of $\tilde{p}$. It suffices to consider the 'full' system $(\mathbb{R}^d, \Gamma_{d,\Delta,s})$ whose ranges are all sets of the form $\{x \in \mathbb{R}^d : (x, \tilde{p}) \in E_\Phi\}$ for every choice of Φ and $p \in \mathcal{P}^s$. If we let $i \in [1 \mathinner{..} 2^{2^s}] \mapsto \Phi_i$ be any permutation of the 2^{2^s} Boolean functions of s variables, the range of its dual corresponding to a given choice of x is

$$\left\{ (t, \tilde{p}) \in \mathbb{R} \times \mathbb{R}^{Ns} : \bigvee_{i=1}^{2^{2^s}} (t = i \wedge (x, \tilde{p}) \in E_{\Phi_i}) \right\}. \qquad \square$$

4 Hitting and Sampling Set Systems

Once we have defined set systems, sampling those structures becomes a natural problem. Indeed the ground set might hold more elements than is computationally feasible to manipulate in practical applications; it might even have infinitely many. Moreover many natural problems can be formulated as the choice of a system of representatives for the ranges: for example the design of surveys.

Objects	**Space**		VC dim.	$s(m)$	$\tilde{\varphi}(m)$
Intervals	$\mathbb{R}$	P	2	$m(m+1)/2+1$	$O(1)$
Intervals	$\mathbb{R}$	D	2	$O(m^2)$	$O(1)$
Lines	$\mathbb{R}^2$	P	2	$m(m+1)/2+1$	$O(m)$
Lines	$\mathbb{R}^2$	D	2	$O(m^2)$	$O(m)$
Triangles	$\mathbb{R}^2$	D	7	$O(m^7)$	$O(m)$
Pseudodisks	$\mathbb{R}^2$	P	3	$O(m^3)$	$O(1)$
Pseudodisks	$\mathbb{R}^2$	D	$O(1)$	$O(m^2)$	$O(1)$
Half-spaces	$\mathbb{R}^d$	P/D	$d+1$	$O(m^d)$	$O(m^{\lfloor d/2\rfloor-1})$
Balls	$\mathbb{R}^d$	P/D	$d+1$	$O(m^{d+1})$	$O(m^{\lceil d/2\rceil-1})$
Convex sets	$\mathbb{R}^d$	P	$+\infty$	2^m	$O(2^m/m)$

Table 1.1: Complexities of some geometric set systems (**P**rimal and **D**ual). (A **family of pseudodisks** is a set of bounded planar regions whose boundaries are Jordan curves and such that the boundaries of any pair of pseudodisks intersect at most twice.)

4.1 Transversals

A requirement for our sampling could be that it contains at least one representative element from each range: this is the classical hitting-set problem.

Definition 1.10. A **transversal**, or hitting set, for the set system $(X, \mathcal{R})$ is any subset H of X that intersects all ranges of $\mathcal{R}$:

$$\forall R \in \mathcal{R}, \quad H \cap R \neq \emptyset.$$

Example 1.9. In the Art Gallery optimisation problem, we are given a simple polygonal region $\mathcal{P}$ in the plane—a gallery—and we want to find a smallest subset G of $\mathcal{P}$—consisting of guards—that together see all of $\mathcal{P}$, that is, there is a segment fully contained in $\mathcal{P}$ from G to every point of $\mathcal{P}$. The art gallery problem amounts to computing a transversal for the set system on $\mathcal{P}$ whose ranges are the sets $\{g \in \mathcal{P} : g \text{ sees } x\}$ for $x \in \mathcal{P}$.

Example 1.10. In the Minimum Set Cover optimisation problem, we are given a finite set system and want to find a smallest subset of ranges whose union equals the ground set. This is exactly a smallest transversal of the dual system.

Computing transversals for various classes of set systems is a fundamental algorithmic challenge. However, for some applications small ranges may be inconsequential, but have a significant contribution to the size of transversals. For example the primal set system defined by disks admits all singletons as ranges, meaning that only its ground set is a transversal.

The idea of discarding those small ranges leads to the concept of nets.

4.2 Nets

Definition 1.11. Given a set system $(X, \mathcal{R})$, a parameter $\epsilon \in [0, 1]$, and a finite subset $Y \subseteq X$, let

$$\mathcal{R}_\epsilon = \{R \in \mathcal{R} : |R \cap Y| \geq \epsilon |Y|\}.$$

A **weak ϵ-net** for Y in $(X, \mathcal{R})$ is any transversal of $(X, \mathcal{R}_\epsilon)$, i.e. any $N \subseteq X$ such that

$$\forall R \in \mathcal{R}, \quad |R| \geq \epsilon |Y| \implies N \cap R \neq \emptyset.$$

It is **strong** if it is a subset not only of X but also of Y.

We also say that the system $(X, \mathcal{R})$ 'has ϵ-nets of size $S(\epsilon)$' if it has an ϵ-net of cardinality at most $S(\epsilon)$ for each finite $Y \subseteq X$.

Example 1.11. The primal set system of convex sets in $\mathbb{R}^2$ has weak ϵ-nets of size $O(\epsilon^{-(3/2+\delta)})$ for any positive δ. This bound is independent of $|Y|$ and 'close to linear' in $1/\epsilon$. (This is a very recent development by Rubin [92]. The previous best bound had been $O(\epsilon^{-2})$ since 1992.)

In contrast any strong ϵ-net for a subset $Y \subset \mathbb{R}^2$ in convex position must have size at least $|Y| - \lfloor \epsilon |Y| \rfloor$, a quantity that grows linearly with $|Y|$.

Remark 1.3. In what follows we either talk about ϵ-nets or $(1/r)$-nets ($r \geq 1$), depending on what is most convenient at the time. Furthermore 'net' without a qualifier always means strong net as in Definition 1.11.

4.2.1 Existence of small nets

Beginning with the breakthrough results of Clarkson [37, 38] and Haussler and Welzl [60], nets have been one of the most fundamental structures in combinatorial geometry with many applications in areas such as approximation algorithms, discrete and computational geometry, combinatorial discrepancy theory and learning theory [35, 77, 78, 90].

Haussler and Welzl showed that set systems with finite VC dimension d have $(1/r)$-nets of size $O(dr \log(dr))$.

Chan et al. [32], improving on an earlier, weaker result of Varadarajan [99] for dual set systems induced by geometric objects, generalised this result to descriptions based on SCC. See also the simpler proof by Mustafa, Dutta, and Ghosh [85].

Theorem A. *A set system with SCC $\tilde{\varphi}(m) = O(m^d)$ for some constant d has $(1/r)$-nets of size $O(r \log \tilde{\varphi}(r))$ for any $r \geq 1$. Furthermore, such nets can be computed in deterministic polynomial time.*

Theorem A has been shown to be tight by Kupavskii, Mustafa, and Pach [69] (this is Theorem F on page 62). A recent survey on nets is the one by Mustafa and Varadarajan [88].

5 Combining Set Systems

There are numerous ways to define binary operations on set systems. The most common ones extend those that are already well-studied on graphs, either unions or products [61]. Here we present one that does *not* extend a graph-theoretic construct since it does not preserve the cardinalities of ranges.

Definition 1.12. The **union** $(X_1, \mathcal{R}_1) \cup (X_2, \mathcal{R}_2)$ is $(X_1 \cup X_2, \mathcal{R}_1 \cup \mathcal{R}_2)$.

Definition 1.13. Given two set systems $(X_1, \mathcal{R}_1)$ and $(X_2, \mathcal{R}_2)$ with X_1 and X_2 disjoint, their **$\star$-sum** $(X_1, \mathcal{R}_1) \star (X_2, \mathcal{R}_2)$ has ground set $X_1 \cup X_2$ and ranges $R_1 \cup R_2$ for all $(R_1, R_2) \in \mathcal{R}_1 \times \mathcal{R}_2$.

Discard the disjointness condition by replacing the union of ground sets with a disjoint union. In this way the operation $\star$ is defined on all set systems.

Example 1.12. Consider the set system $\mathcal{S}$ with ground set $\{0, 1\}$ and only range $\{0, 1\}$. The union $\mathcal{S} \cup \mathcal{S}$ is still $\mathcal{S}$ whereas the $\star$-square $\mathcal{S} \star \mathcal{S}$ is (up to isomorphism) the system $(\{0, 1, 2, 3\}, \{\{0, 1, 2, 3\}\})$.

The interest of $\star$ is that it behaves like a sum on ground sets but also like a product on range sets. For example, the n-th $\star$-power of $(X, \mathcal{R})$ has a ground set of cardinality $n\,|X|$, but $|\mathcal{R}|^n$ ranges.

Remark 1.4. On simplicial complexes (set systems whose ranges are all finite and whose range sets are downward closed), $\star$ coincides with the join operation [79].

Remark 1.5. The reader is free to choose her favourite construction for disjoint unions as they are universal in a category-theoretic sense. A possible choice is

$$X \sqcup Y = (X \times \{0\}) \cup (Y \times \{1\})\,.$$

Lemma 1.4. *For any two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ and for all $(m, l) \in \mathbb{N}^2$, one has*

$$s_{\mathcal{S}_1 \cup \mathcal{S}_2}(m, l) \le s_{\mathcal{S}_1}(m, l) + s_{\mathcal{S}_2}(m, l), \tag{1.1}$$

$$s_{\mathcal{S}_1 \star \mathcal{S}_2}(m, l) \le s_{\mathcal{S}_1}(m, l) \cdot s_{\mathcal{S}_2}(m, l), \tag{1.2}$$

and as a consequence

$$\textit{VC-}\textsc{dim}(\mathcal{S}_1 \cup \mathcal{S}_2) \le \textit{VC-}\textsc{dim}(\mathcal{S}_1) + \textit{VC-}\textsc{dim}(\mathcal{S}_2) + 1, \tag{1.3}$$

$$\textit{VC-}\textsc{dim}(\mathcal{S}_1 \star \mathcal{S}_2) = \textit{VC-}\textsc{dim}(\mathcal{S}_1) + \textit{VC-}\textsc{dim}(\mathcal{S}_2). \tag{1.4}$$

Proof. Write $\mathcal{S}_i = (X_i, \mathcal{R}_i)$ for $i \in \{1, 2\}$. To prove (1.1) we may assume $X_1 = X_2$. Then for any $Y \subseteq X_1$, the trace $(\mathcal{R}_1 \cup \mathcal{R}_2)_{|Y}$ is the union of $\mathcal{R}_{1|Y}$ and $\mathcal{R}_{2|Y}$. On the other hand to prove (1.2) we may take X_1 and X_2 disjoint. Let $\mathcal{R}$ be the range set of $\mathcal{S}_1 \star \mathcal{S}_2$. With some slight notational abuse, for any $Y \subseteq X_1 \cup X_2$ the set $\mathcal{R}_{|Y, \le l}$ is in bijection with a subset of $\mathcal{R}_{1|Y \cap X_1, \le l} \times \mathcal{R}_{2|Y \cap X_2, \le l}$. □

Chapter Two

Packing Lemmas

HAUSSLER's packing lemma gives an upper bound on the number of ranges in a set system with finite VC dimension whose ranges have pairwise large symmetric difference [59]. In this chapter we discuss it and present some recent extensions.

The packing lemma has turned out to be the technical foundation for many results in combinatorial discrepancy of geometric set systems using the entropy method (see Matoušek's *Geometric Discrepancy: An Illustrated Guide* [77]) as well as recent work on Zarankiewicz's problem for semi-algebraic graphs by Fox et al. [49]. Recently it was generalised [44, 84] to the shallow packing lemma, applying to set systems as a function of their SCC.

The two new results on packings in this chapter are:

- An optimal lower bound for shallow packings (Theorem 1), which settles the open question in Ezra [47] and Dutta, Ezra, and Ghosh [44].
- The l-wise shallow packing lemma (Theorem 2), which combines the ideas of the shallow packing lemma and the l-wise packing lemma of Fox et al. [49].

Although the shallow packing lemma is worthy of study in its own right, it will be a stepping stone towards the Mnet results in Chapter 3.

These results were obtained with Kunal Dutta, Arijit Ghosh & Nabil H. Mustafa and presented at the 33rd International Symposium on Computational Geometry [45], then invited to a special edition of Discrete & Computational Geometry *dedicated to the Symposium's best papers.*

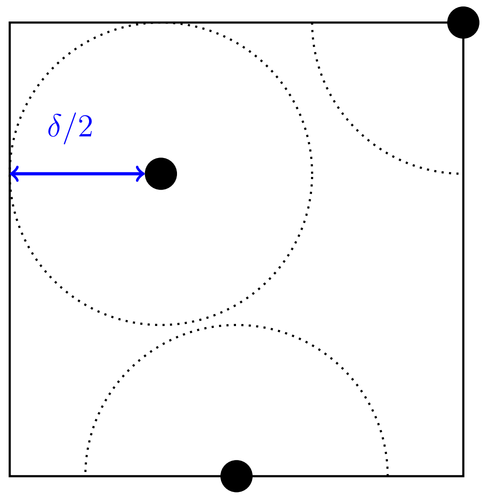


Figure 2.1: Euclidean intuition for Haussler's packing lemma: in $\mathbb{R}^d$, balls of diameter δ have volume $\Theta(\delta^d)$, so only $O((n/\delta)^d)$ of them may be packed within a d-dimensional cube of volume n^d.

1 Classical and Shallow Packing Lemmas

1.1 Haussler's Packing Lemma

Definition 2.1. A set system with range set $\mathcal{R}$ is a **δ-packing** for some $\delta \in \mathbb{N}^*$ if $|R \,\Delta\, S| \geq \delta$ for all distinct $(R, S) \in \mathcal{R}^2$, where Δ is the symmetric difference.

By extension, we also say that $\mathcal{P}$ is a δ-packing of $(X, \mathcal{R})$ if $(X, \mathcal{P})$ is a δ-packing and $\mathcal{P} \subseteq \mathcal{R}$. It is **maximal** if no subset of $\mathcal{R}$ is both a proper superset of $\mathcal{P}$ and a δ-packing.

Remark 2.1. In a (semi-)metric space, a δ-packing is defined as any subset of points with pairwise distance at least δ. This includes Definition 2.1 with the Hamming distance

$$(R, S) \in \mathcal{R}^2 \mapsto |R \,\Delta\, S| \,.$$

Haussler [59] gave the following upper bound on the size of packings.

Theorem B (Packing Lemma). *If $(X, \mathcal{R})$ is a finite δ-packing with VC-DIM$(\mathcal{R}) \leq d$, then*

$$|\mathcal{R}| \leq e(d+1)\left(\frac{2e(|X|+1)}{\delta+2d+2}\right)^d \leq e(d+1)\left(\frac{2e\,|X|}{\delta}\right)^d .$$

Remark 2.2. The same asymptotic bound also holds as long as $s(m) = O(m^d)$, see [77, Chapter 5.3].

Haussler's proof of Theorem B, later simplified by Chazelle [34], is an elegant application of the probabilistic method, and has since been applied to diverse areas ranging from computational geometry and machine learning to Bayesian inference—see e.g. [59, 71, 77]. Haussler also showed that this bound is tight:

Theorem C. *Given any positive integers d, n and $\delta \in [1 \mathinner{..} n]$, there exists a δ-packing $(X, \mathcal{R})$ such that $|X| = n$, VC-DIM$(\mathcal{R}) \leq d$, and*

$$|\mathcal{R}| \geq \left(\frac{n}{2e(\delta + d)}\right)^d .$$

1.2 Shallow Packing Lemma

Recent efforts have been devoted to extending the packing lemma to finer classifications of set systems. For $k \in \mathbb{N}$, call $(X, \mathcal{R})$ **k-shallow** if $|R| \leq k$ for every $R \in \mathcal{R}$. Equivalently $\mathcal{R}$ is included in the Hamming ball of radius k around $\emptyset$ in $\mathcal{P}(X)$.

Example 2.1. Graphs are 2-shallow set systems, d-dimensional simplicial complexes are d-shallow set systems.

After some earlier bounds [47, 86], the following lemma has been recently established for Clarkson–Shor set systems by Dutta, Ezra, and Ghosh [44]. Recall from Example 1.6 that a system is (d, d_1)-Clarkson–Shor when it has $s(m, l) = O(m^{d_1} l^{d-d_1})$. The result was then extended in terms of SCC by Mustafa [84].

Theorem D (Shallow Packing Lemma)**.** *Let $(X, \mathcal{R})$ be a finite and k-shallow δ-packing. If VC-DIM$(\mathcal{R}) \leq d$, then*

$$|\mathcal{R}| \leq 6 \cdot s_{\mathcal{R}} \left(4d \cdot \frac{|X|}{\delta}, 12d \cdot \frac{k}{\delta}\right) .$$

Remark 2.3. When the system is (d, d_1)-Clarkson–Shor, the bound becomes

$$O\left(\frac{|X|^{d_1} k^{d-d_1}}{\delta^d}\right) .$$

Additionally, taking $k = |X|$ gives the bound of Theorem B up to a constant factor.

❧

❧ ❧

2 Contributions

In this chapter we present two new results: a tight lower bound for shallow packings and a generalisation of the shallow packing lemma to l-wise packings.

2.1 Optimality of Shallow Packings (Proof in Section 3)

While Haussler gave a matching lower bound to his packing lemma, the optimality of the *shallow* packing lemma was an open question in previous work [44, 47, 84, 86]. In particular a matching lower bound was presented for one particular case, when $\tilde{\varphi}(m) = m$ [44]. We show that the shallow packing lemma is tight up to a constant factor for the most common case of upper bound on the SCC: the Clarkson–Shor property.

Theorem 1 (Optimality of Shallow Packings). *For any positive integers $d \geq d_1$ and for any large positive integer n, there exists a set system on n elements with*

$$s(m, l) \leq 2^{d_1} m^{d_1} l^{d-d_1}$$

that, for any δ and $k \geq 4d\delta$, has a k-shallow δ-packing of size

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

The proof is an explicit construction of such a set system.

2.2 l-Wise Shallow Packings (Proof in Section 4)

Definition 2.2. A set system $(X, \mathcal{R})$ is an **l-wise δ-packing** if for all tuples of l distinct ranges $R_1, \ldots, R_l \in \mathcal{R}$ we have

$$\left|(R_1 \cup \cdots \cup R_l) \setminus (R_1 \cap \cdots \cap R_l)\right| \geq \delta.$$

In particular δ-packings are 2-wise δ-packings. Building on the proof of the packing lemma by Chazelle together with Turán's theorem on independent sets in graphs, Fox et al. [49, Lemma 2.5] proved the following:

Theorem E (l-Wise Packing Lemma). *Let $(X, \mathcal{R})$ be a finite set system with $s_{\mathcal{R}}(m) = O(m^d)$. If $\mathcal{R}$ is an l-wise δ-packing, for an integer $l \geq 2$ and $\delta \in [1 \,..\, |X|]$, then*

$$|\mathcal{R}| = O\left(\left(\frac{|X|}{\delta}\right)^d\right),$$

where the constant in the asymptotic notation depends on l and d.

Building on the proofs by Matoušek [77] and Mustafa [84], we prove the following, which simultaneously generalises the theorems of Haussler [59] (Theorem B), Fox et al. [49] (Theorem E), and Dutta, Ezra, and Ghosh [44] (Theorem D).

Theorem 2 (l-Wise Shallow Packing Lemma)**.** *Let $(X, \mathcal{R})$ be a k-shallow l-wise δ-packing with $|X| = n$. If VC-DIM$(\mathcal{R}) \leq d$, then*

$$|\mathcal{R}| = O\left(\frac{l^3 dn}{\delta} \cdot \varphi_{\mathcal{R}}\left(s, 4l \cdot \frac{ks}{n}\right)\right), \text{ where } s = \frac{8l(l-1)dn}{\delta} - 1.$$

Corollary 2.1. *Theorems B, D (up to a constant factor) and E.*

Proof. Theorem D is Theorem 2 with l set to 2. To obtain Theorem E, set $k = |X|$ in Theorem 2. Theorem B is the special case of Theorem E when $l = 2$. □

3 Building Large Shallow Packings

In this section we prove Theorem 1 by building a set system with the desired SCC that contains a large shallow packing.

Proof of Theorem 1. Without loss of generality we assume that n is an integer multiple of d. Let $\hat{n} = n/d$.

Define the following range set $\mathcal{P}$ on $[1 \mathinner{..} \hat{n}]$:

$$\mathcal{P} = \left\{[2^\alpha\beta + 1 \mathinner{..} 2^\alpha(\beta + 1)] : \ 0 \leq \alpha \leq \log_2 \hat{n}, \ 0 \leq \beta < 2^{-\alpha}\hat{n}\right\}.$$

Intuitively, consider a balanced binary tree $\mathcal{T}$ with its leaves labelled from 1 to $\hat{n}$ (see Figure 2.2). Then for each node of $\mathcal{T}$, $\mathcal{P}$ contains a set consisting of the leaves of the sub-tree rooted at that node. Here α is the height of the node (so 2^α is the number of elements in the corresponding subset), while β identifies one of the nodes of that height (among the $2^{\log(\hat{n})-\alpha} = 2^{-\alpha} \cdot \hat{n}$ choices), see Figure 2.2.

Claim 2.1. $s_{\mathcal{P}}(m, l) \leq 2m$.

Proof. For any $Y \subseteq [1 \mathinner{..} \hat{n}]$, the sets in $\mathcal{P}_{|Y}$ are in a one-to-one correspondence with the nodes of $\mathcal{T}$ whose left and right sub-trees, if they exist, both contain leaves labelled by Y. If the nodes of $\mathcal{T}$ corresponding to Y form a connected sub-tree, then these nodes define a new binary tree whose leaves are still labelled by Y, and thus their number is at most $2|Y| - 1$. Otherwise, the statement holds by induction on the number of connected components of Y in $\mathcal{T}$. □

Next define the range set

$$\mathcal{Q} = \{[1 \mathinner{..} \gamma] : \gamma \in [1 \mathinner{..} \hat{n}]\},$$

whose elements can be seen as prefix sets of the list $\langle 1, \ldots, \hat{n} \rangle$.

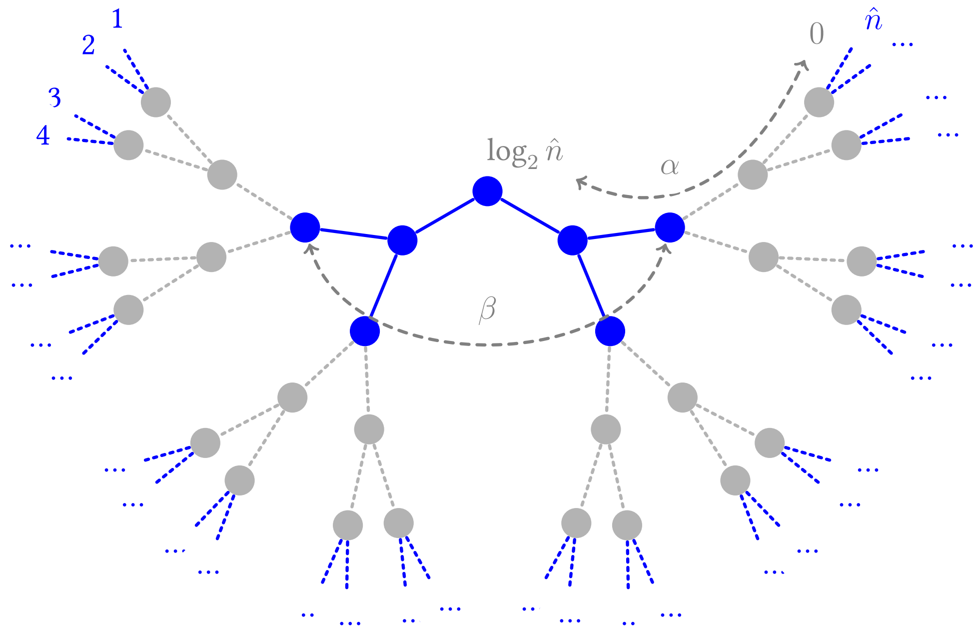


Figure 2.2: A set system with linear growth. Each inner node corresponds to the set of leaves of its sub-tree, a sub-interval of $[1 \,..\, \hat{n}]$, and is identified by a pair (α, β) corresponding to its depth and its position relative to other nodes at the same depth. Claim 2.1 establishes that it has $s(m, l) \le 2m$.

Claim 2.2. $s_{\mathcal{Q}}(m, l) \le l$.

Proof. The number of sets of size at most l is $|\mathcal{Q}|_{Y, \le l}| = \min\{l, |Y|\} \le l$. □

Write $\mathcal{S}_1$ for the system $([1 \,..\, \hat{n}], \mathcal{P})$ and $\mathcal{S}_2$ for $([1 \,..\, \hat{n}], \mathcal{Q})$. Recall the $\star$ operation (Definition 1.13) and consider the system

$$(X, \mathcal{R}) = \underbrace{\mathcal{S}_1 \star \cdots \star \mathcal{S}_1}_{d_1 \text{ times}} \star \underbrace{\mathcal{S}_2 \star \cdots \star \mathcal{S}_2}_{d - d_1 \text{ times}}.$$

By Lemma 1.4 one has the desired bound on $s_{\mathcal{R}}$:

$$s_{\mathcal{R}}(m, l) \le (s_{\mathcal{P}}(m, l))^{d_1} \cdot (s_{\mathcal{Q}}(m, l))^{d - d_1} = (2m)^{d_1} l^{d - d_1}. \tag{2.1}$$

Now that we have built $\mathcal{R}$, it remains to construct a subset of $\mathcal{R}$ which is a large shallow

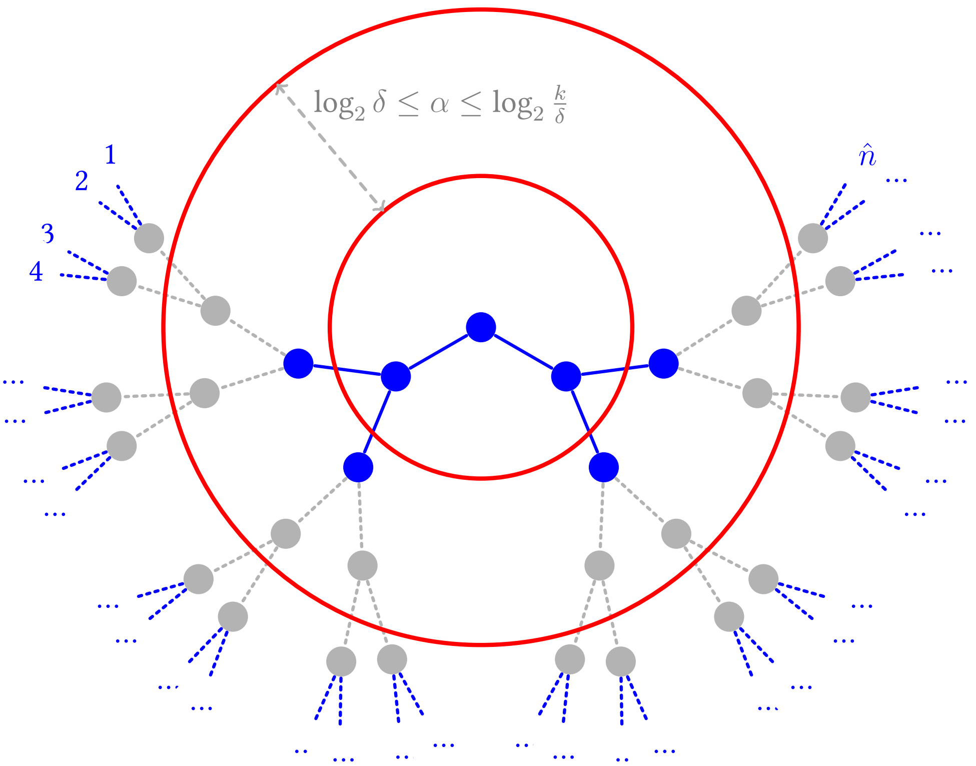


Figure 2.3: Selecting ranges corresponding to low-depth nodes.

packing. For every choice of the parameters k, δ, define:

$$\mathcal{P}^{(k,\delta)} = \left\{ [2^\alpha\beta + 1 \mathinner{..} 2^\alpha(\beta+1)] : \begin{array}{l} \alpha, \beta \in \mathbb{N} \\ \log_2 \delta \le \alpha \le \log_2(\frac{k}{d}) \\ 0 \le \beta < 2^{-\alpha}\hat{n} \end{array} \right\} \subseteq \mathcal{P},$$

$$\mathcal{Q}^{(k,\delta)} = \left\{ [1 \mathinner{..} \gamma\delta] : \gamma \in \mathbb{N},\ 1 \le \gamma \le \frac{k}{d\delta} \right\} \subseteq \mathcal{Q}.$$

The intuition here is that we pick only the nodes in our binary tree $\mathcal{T}$ with height at least $\log_2 \delta$ (and thus a symmetric difference of at least δ elements), see Figure 2.3. Similarly in $\mathcal{Q}$ we pick every δ-th set only. All these sets have size at most k/d. This is straightforward for $\mathcal{Q}^{(k,\delta)}$; on the other hand, a set in $\mathcal{P}^{(k,\delta)}$ defined by the pair (α, β) has size $2^\alpha \le k/d$.

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

Write $\mathcal{S}_1^{(k,\delta)}$ for the system $([1\,..\,\hat{n}], \mathcal{P}^{(k,\delta)})$ and $\mathcal{S}_2^{(k,\delta)}$ for $([1\,..\,\hat{n}], \mathcal{Q}^{(k,\delta)})$. The system

$$\underbrace{\mathcal{S}_1^{(k,\delta)} \star \cdots \star \mathcal{S}_1^{(k,\delta)}}_{d_1 \text{ times}} \star \underbrace{\mathcal{S}_2^{(k,\delta)} \star \cdots \star \mathcal{S}_2^{(k,\delta)}}_{d - d_1 \text{ times}}$$

is a δ-packing of $(X, \mathcal{R})$ which is k-shallow.

Its number of ranges is:

$$\begin{aligned}
\left|\mathcal{P}^{(k,\delta)}\right|^{d_1} \cdot \left|\mathcal{Q}^{(k,\delta)}\right|^{d-d_1} &= \left(\hat{n} \sum_{\alpha=\lceil \log_2 \delta \rceil}^{\lfloor \log_2(\frac{k}{d}) \rfloor} 2^{-\alpha}\right)^{d_1} \cdot \left(\frac{k}{d\delta}\right)^{d-d_1} \\
&\geq d^{-d} \left(2^{1-\lceil \log_2 \delta \rceil} - 2^{\lfloor \log_2(\frac{k}{d}) \rfloor}\right)^{d_1} n^{d_1} \left(\frac{k}{\delta}\right)^{d-d_1} \\
&\geq d^{-d} \left(\frac{1}{\delta} - \frac{2d}{k}\right)^{d_1} n^{d_1} \left(\frac{k}{\delta}\right)^{d-d_1} \\
&\geq d^{-d} (2\delta)^{-d_1} n^{d_1} \left(\frac{k}{\delta}\right)^{d-d_1} \\
&= \Omega\left(\frac{n^{d_1} k^{d-d_1}}{\delta^d}\right). \qquad \square
\end{aligned}$$

The gist of Haussler's probabilistic lower bound construction for Theorem C was to consider this same $\mathcal{R}$ with $d_1 = 0$—that is, no trees—and randomly build a maximal packing (select a range at random, remove its Hamming ball of radius δ, recurse).

4 Proving the l-Wise Shallow Packing Lemma

The proof of Theorem 2 uses a technical lemma that combines the ideas in [49, 77, 84].

Lemma 2.1. *Let $(X, \mathcal{R})$ be an l-wise δ-packing on n elements with VC-DIM$(\mathcal{R}) \leq d$. If $A \subseteq X$ is a uniformly selected random sample of size $8l(l-1)d\frac{n}{\delta} - 1$, then*

$$\mathbf{E}\left|\mathcal{R}_{|A}\right| \geq \frac{|\mathcal{R}|}{2l}.$$

Proof. Let $t = 8l(l-1)d\frac{n}{\delta}$ and pick a random sample $S = \{x_1, \ldots, x_t\}$ one element at a time without replacement, from x_1 to x_t, from X. Let $(\mathcal{R}_{|S}, E_{\mathcal{R}})$ be the unit distance graph on the trace $\mathcal{R}_{|S}$, with an edge between any two sets whose symmetric difference is a singleton. Define the weight of a set $Q \in \mathcal{R}_{|S}$ as the number of ranges of $\mathcal{R}$ whose projection is Q, i.e.

$$w(Q) = \left|\{R \in \mathcal{R} : R \cap S = Q\}\right|,$$

and the weight of an edge $\{Q_1, Q_2\} \in E_{\mathcal{R}}$ as $w(Q_1, Q_2) = \min\{w(Q_1), w(Q_2)\}$. Finally let

$$W = \sum_{E_{\mathcal{R}}} w.$$

In [77, Chapter 5, Proof 5.14] it was shown that

$$W \leq 2d \cdot |\mathcal{R}| \,. \tag{2.2}$$

On the other hand, we lower-bound the expected value of W. Let W_i be the weight of the edges in $E_{\mathcal{R}}$ for which the symmetric difference is the singleton $\{x_i\}$. By symmetry, we have $\mathbf{E}\, W = \sum_{i=1}^{t} \mathbf{E}\, W_i = t \cdot \mathbf{E}\, W_t$.

Let $A = \{x_i : i \in [1\,..\,t-1]\}$, i.e. $A = S \setminus \{x_t\}$. A proof of the following conditional lower bound on the expected value of W_t is deferred until the end of this section:

$$t \cdot \mathbf{E}(W_t|A) \geq 4d\left(|\mathcal{R}| - l \cdot \left|\mathcal{R}_{|A}\right|\right). \tag{2.3}$$

It follows from it that

$$\mathbf{E}\, W = t \cdot \mathbf{E}\, W_t = t \cdot \mathbf{E}(\mathbf{E}\,(W_t|A)) \geq 4d\left(|\mathcal{R}| - l\,\mathbf{E}\left|\mathcal{R}_{|A}\right|\right). \tag{2.4}$$

Finally combine the upper bound (2.2) with the lower bound (2.4):

$$2d\,|\mathcal{R}| \geq \mathbf{E}\, W \geq 4d\,|\mathcal{R}| - 4dl \cdot \mathbf{E}\left|\mathcal{R}_{|A}\right|,$$

and then $|\mathcal{R}| \leq 2l \cdot \mathbf{E}\left|\mathcal{R}_{|A}\right|$. □

We can now prove Theorem 2.

Proof of Theorem 2. Let $A \subseteq X$ be a random sample of size $8l(l-1)d\frac{n}{\delta} - 1$. Also let

$$\hat{\mathcal{R}} = \left\{ R \in \mathcal{R} : |R \cap A| \leq 4l \cdot \frac{k}{n}\,|A| \right\}.$$

Each element of X belongs to A with probability $\frac{|A|}{n}$, and thus $\mathbf{E}\,|R \cap A| \leq \frac{k}{n}\,|A|$ as $|R| \leq k$ for all $R \in \mathcal{R}$. Markov's inequality then bounds the probability of a range of $\mathcal{R}$ not belonging to $\hat{\mathcal{R}}$:

$$\mathbf{Pr}\,(R \notin \hat{\mathcal{R}}) = \mathbf{Pr}\left(|R \cap A| > 4l \cdot \frac{k}{n}\,|A|\right) \leq \frac{1}{4l},$$

from which:

$$\mathbf{E}\left|\mathcal{R} \setminus \hat{\mathcal{R}}\right| \leq \frac{|\mathcal{R}|}{4l}.$$

Let $t = 4l\frac{k}{n}|A|$ and $\hat{\mathcal{R}}_{|A}$ is exactly $\mathcal{R}_{|A,\le t}$, so

$$\left|\hat{\mathcal{R}}_{|A}\right| = \left|\mathcal{R}_{|A,\le t}\right| \le |A| \cdot \varphi(|A|, t).$$

From this it follows that

$$\mathbf{E}\left|\mathcal{R}_{|A}\right| \le \mathbf{E}\left|\hat{\mathcal{R}}_{|A}\right| + \mathbf{E}\left|\mathcal{R} \setminus \hat{\mathcal{R}}\right| \le |A| \cdot \varphi\Big(|A|, 4l \cdot \frac{k}{n}|A|\Big) + \frac{|\mathcal{R}|}{4l}.$$

Now the bound follows from Lemma 2.1. $\square$

We end by proving the inequality (2.3) which was used to obtain Lemma 2.1.

Proof of (2.3). Consider a set $Q \in \mathcal{R}_{|A}$, and let $\mathcal{R}_Q$ consist of the ranges of $\mathcal{R}$ sent to Q by the projection on A, i.e. those R such that $R \cap A = Q$. Depending on the choice of x_t, $\mathcal{R}_Q$ is split into those ranges containing x_t—forming a set $\mathcal{A}$—and those not containing x_t—in a set $\mathcal{B}$. From the definition of weights, the expected contribution of the ranges of $\mathcal{R}_Q$ to edge weight is

$$\mathbf{E}\min\{|\mathcal{A}|, |\mathcal{B}|\} \ge \mathbf{E}\left(\frac{|\mathcal{A}||\mathcal{B}|}{|\mathcal{A}| + |\mathcal{B}|}\right) = \frac{1}{|\mathcal{R}_Q|} \cdot \mathbf{E}(|\mathcal{A}||\mathcal{B}|). \tag{2.5}$$

The above inequality follows from $\min\{|\mathcal{A}|, |\mathcal{B}|\} \ge |\mathcal{A}||\mathcal{B}|/(|\mathcal{A}| + |\mathcal{B}|)$. Now:

$$\begin{aligned}
\mathbf{E}(|\mathcal{A}||\mathcal{B}|) &= \frac{1}{|X \setminus A|} \sum_{x \in X \setminus A} |\{R \in \mathcal{R}_Q : x \in R\}| \, |\{R \in \mathcal{R}_Q : x \notin R\}| \\
&= \frac{1}{|X \setminus A|} \sum_{x \in X \setminus A} \sum_{R_1 \in \mathcal{R}_Q} \sum_{R_2 \in \mathcal{R}_Q} \mathbb{1}_{R_1}(x)(1 - \mathbb{1}_{R_2}(x)) \\
&= \frac{1}{|X \setminus A|} \sum_{R_1 \in \mathcal{R}_Q} \sum_{R_2 \in \mathcal{R}_Q} |(X \setminus A) \cap (R_1 \setminus R_2)| \\
&= \frac{1}{|X \setminus A|} \sum_{\{R_1, R_2\} \in \binom{\mathcal{R}_Q}{2}} |(X \setminus A) \cap (R_1 \,\Delta\, R_2)|.
\end{aligned}$$

Observe that R_1 and R_2 agree on A, so this simplifies to

$$\mathbf{E}(|\mathcal{A}||\mathcal{B}|) = \frac{1}{|X \setminus A|} \sum_{\{R_1, R_2\} \in \binom{\mathcal{R}_Q}{2}} |R_1 \,\Delta\, R_2|. \tag{2.6}$$

For all l sets $R_1, \ldots, R_l \in \mathcal{R}_Q$,

$$\bigcup_{2 \le j \le l} R_1 \,\Delta\, R_j = (R_1 \cup \cdots \cup R_l) \setminus (R_1 \cap \cdots \cap R_l),$$

and since $\mathcal{R}$ is an l-wise δ-packing

$$\sum_{2\leq j\leq l} |R_1 \,\Delta\, R_j| \geq |(R_1 \cup \cdots \cup R_l) \setminus (R_1 \cap \cdots \cap R_l)| \geq \delta.$$

So for every l-tuple there exists one pair (R_1, R_j) with $|R_1 \,\Delta\, R_j| \geq \frac{\delta}{l-1}$. Consider the graph $(\mathcal{R}_Q, E_Q)$, where $\{R_1, R_2\} \in E_Q$ if $|R_1 \,\Delta\, R_2| \geq \frac{\delta}{l-1}$. As $\mathcal{R}_Q$ is an l-wise δ-packing this graph does not have independent sets of size l. Turán's theorem (see Appendix A) gives a lower bound for its number of edges:

$$|E_Q| \geq \frac{|\mathcal{R}_Q|\,(|\mathcal{R}_Q| - l)}{2l}. \tag{2.7}$$

Recall (2.6):

$$\begin{aligned}
\mathbf{E}(|\mathcal{A}|\,|\mathcal{B}|) &= \frac{1}{|X \setminus A|} \sum_{\{R_1,R_2\}\in\binom{\mathcal{R}_Q}{2}} |R_1 \,\Delta\, R_2| \\
&\geq \frac{1}{|X \setminus A|} \sum_{\{R_1,R_2\}\in E_Q} |R_1 \,\Delta\, R_2| \\
&\geq \frac{|E_Q|}{|X \setminus A|} \cdot \frac{\delta}{l-1} \\
&\geq \frac{\delta\,|R_Q|\,(|R_Q| - l)}{|X \setminus A|\,(l-1)} \quad \text{by (2.7)}.
\end{aligned}$$

Now come back to (2.5):

$$\mathbf{E}(\min\{|\mathcal{A}|\,,|\mathcal{B}|\}) \geq \frac{\mathbf{E}(|\mathcal{A}|\,|\mathcal{B}|)}{|\mathcal{R}_Q|} \geq \frac{\delta(|R_Q| - l)}{|X \setminus A|\,(l-1)},$$

where $|X \setminus A| = n - (s-1) \geq \left(1 - \frac{8l(l-1)}{\delta}\right) n$. Summing up over all sets of $\mathcal{R}_{|A}$,

$$\begin{aligned}
\mathbf{E}(W_s|A) &\geq \frac{1}{2l(l-1)} \sum_{Q\in\mathcal{R}_{|A}} \frac{\delta}{n}\,(|\mathcal{R}_Q| - l) \\
&= \frac{\delta/n}{2l(l-1)}\left(|\mathcal{R}| - l\,|\mathcal{R}_{|A}|\right).
\end{aligned}$$

□

5 Remarks

Lower bound for the Shallow Packing Lemma. The lower bound construction given in the proof of Theorem 1, showing the optimality of the Shallow Packing Lemma (Theorem D), is constructive. Also observe that it can be realised in a number of simple ways, for example with points on a square grid and sets induced by some specific $(2d)$-gons, i.e. a semi-algebraic set system with constant description complexity.

Chapter Three

Combinatorial Macbeath Regions

COMBINATORIAL analogues of decomposition theorems that use Macbeath regions in convex geometry are obtained using the packing lemmas from the previous chapter. Building on the results of Chapter 2, this chapter contains several applications to the combinatorics of geometric set systems:

1. Improved bounds on combinatorial Macbeath regions, or *Mnets*, a combinatorial analogue to Macbeath regions in convex geometry [23, 74] (Theorem 3). This resolves one of the main open problems in Mustafa and Ray [86], where some bounds were obtained by disparate techniques.

2. Mnets provide a general, more powerful framework from which the state-of-the-art unweighted ϵ-net results of Varadarajan [99] and Chan et al. [32] follow immediately (Corollary 3.1).

3. Our upper bounds on Mnets for general semi-algebraic set systems of fixed VC dimension are asymptotically tight (Theorem 4). We also give a more precise lower bound in terms of SCC, based on a similar bound for nets by Kupavskii, Mustafa, and Pach [69].

Besides using the packing lemma and a combinatorial construction, our proofs combine tools from the polynomial partitioning technique of Guth and Katz [54] and the probabilistic method.

Together with those on packings (Chapter 2) these results were presented at the 33rd International Symposium on Computational Geometry [45], then invited to a special edition of Discrete & Computational Geometry *dedicated to the Symposium's best papers.*

1 Macbeath Regions and Mnets

Given a convex body K in $\mathbb{R}^d$ and a small parameter $\epsilon > 0$, Brönnimann, Chazelle, and Pach [23] use a classical construction of Macbeath [74] to build a collection of smaller convex regions $K_1, \dots, K_l \subset K$ collectively known as *Macbeath regions* of K such that

- $l = O\left(\left(\frac{1}{\epsilon}\right)^{1-\frac{2}{d+1}}\right)$,
- $\operatorname{vol} K_i = \Theta(\epsilon \operatorname{vol} K)$ for each i, and
- for any half-space H with $\operatorname{vol}(H \cap K) \geq \epsilon \operatorname{vol}(K)$, there exists a j such that $K_j \subseteq H$.

Details on Macbeath regions can be found in the surveys by Bárány [16] and Bárány and Larman [17]. *Mnets*[1] (or *combinatorial Macbeath regions*), introduced by Mustafa and Ray [86], are a combinatorial analogue of Macbeath regions for set systems, replacing the Lebesgue measure with the counting measure.

Definition 3.1. Given a set system $(X, \mathcal{R})$, two positive parameters ϵ and C, and a finite subset $Y \subseteq X$, a **C-heavy ϵ-Mnet** for Y in $(X, \mathcal{R})$ is a collection $\{M_1, \dots, M_l\}$ of subsets of Y such that

- $|M_i| \geq C\epsilon\,|Y|$ for each i, and
- for any $R \in \mathcal{R}_{|Y}$ with $|R| \geq \epsilon\,|Y|$, there exists an index j such that $M_j \subseteq R$.

Mnets can be seen as a generalisation of strong nets, where the ranges are hit not just with a singleton, but with a set of at least $\lceil C\epsilon\,|Y|\rceil$ elements. Just as we do for strong nets, we say that the set system 'has ϵ-Mnets of size $S(\epsilon)$' if there is an ϵ-Mnet of this size for each finite $Y \subseteq X$.

2 Contributions

The main contribution of this chapter is a new construction of Mnets using the shallow packing lemma. The key idea of this construction is a combination of the polynomial partitioning technique and the shallow packing lemma from Chapter 2.

1. The recommended pronunciation is 'em-net'. The M stands for Macbeath.

2.1 Mnets for Semi-algebraic Set Systems (Proof in Section 3)

Recall that $\Gamma_{d,\Delta,s}$ is the family of all semi-algebraic sets in $\mathbb{R}^d$ obtained by taking Boolean operations on at most s polynomial inequalities, each of degree at most Δ. For our purposes d, Δ, and s are all regarded as constants and therefore the sets in $\Gamma_{d,\Delta,s}$ have constant description complexity. For a detailed introduction to this topic, see the book by Basu, Pollack, and Roy [18].

Given a set X of points in $\mathbb{R}^d$ and a range set $\mathcal{R}$ on X, we say that $(X, \mathcal{R})$ is a **semi-algebraic set system** generated by $\Gamma_{d,\Delta,s}$ if $\mathcal{R}$ is a subset of $\{X \cap \gamma : \gamma \in \Gamma_{d,\Delta,s}\}$.

Theorem 3 (Mnets). *Let d, d_0, Δ and s be integers. Semi-algebraic set systems generated by $\Gamma_{d,\Delta,s}$ with VC dimension d_0 and SCC φ non-decreasing in the first argument have $(1/r)$-Mnets of size*

$$O\left(d_0 r \cdot \varphi(8d_0 r, 48d_0)\right)$$

for all $r \geq 1$. In particular this simplifies to $O\left(r \cdot \tilde{\varphi}\left(8d_0 r\right)\right)$ for SCC in one variable. Constants depend on d, Δ, and s; the second one also depends on d_0.

Most of the time this bound simplifies further to $O\left(r \cdot \tilde{\varphi}(r)\right)$. The proof of Theorem 3 uses the shallow packing lemma (Theorem D), as well as the *polynomial partitioning* method of Guth and Katz [54], specifically a multilevel refinement due to Matoušek and Patáková [81].

First we point out that Theorem 3 immediately implies the best known bounds on unweighted nets—though with the additional restriction that the set system be semi-algebraic—because any transversal for an Mnet is a net.

Corollary 3.1. *Set systems with $\frac{1}{k}$-heavy $\frac{1}{r}$-Mnets of size $M(r,k)$ have $\frac{1}{r}$-nets of size*

$$rk \log(M(r,k)/r) + r.$$

In particular a set system $(X, \mathcal{R})$ with VC-DIM$(\mathcal{R}) \leq d$ has $\frac{1}{r}$-nets of size

$$O\left(r \log \varphi_{\mathcal{R}}(8dr, 48d)\right).$$

Proof. Let $\mathcal{M}$ be an $\frac{1}{k}$-heavy $\frac{1}{r}$-Mnet for $(X, \mathcal{R})$ and $p = \frac{kr}{n} \log(|\mathcal{M}|/r)$. We can assume that $p \in (0,1)$ since if $p \leq 0$ then $\mathcal{M} \leq r$, and picking one element from each set of $\mathcal{M}$ is enough, whereas if $p \geq 1$ then X is a small enough net. Pick each point of X into a random sample R independently with probability p. This sample R is disjoint from any fixed $M_i \in \mathcal{M}$ with probability

$$\mathbf{Pr}(R \cap M_i = \emptyset) \leq (1-p)^{n/(kr)} \leq e^{-np/(kr)} = \frac{r}{|\mathcal{M}|}.$$

Therefore the expected number of sets of $\mathcal{M}$ not hit by R is at most r; let the set S consist of an arbitrary point from each such set. As $\mathbf{E}\,|S| \leq r$, we have that $R \cup S$ is a $\frac{1}{r}$-net of expected size at most $np + r$. $\square$

Set system	Primal/Dual	Size of $(1/r)$-Mnets
Objects w. union complexity κ	D	$O(\kappa(r))$
α-fat triangles	D	$O(r \log^* r)$
Locally γ-fat objects	D	$r \cdot 2^{O(\log^* r)}$
Triangles of approx. same size	D	$O(r)$
Axis-parallel rectangles in $\mathbb{R}^2$	P	$O(r \log r)$
Lines in $\mathbb{R}^2$	**P**	$O(r^2)$
Strips in $\mathbb{R}^2$	**P**	$O(r^2)$
Cones in $\mathbb{R}^2$	**P**	$O(r^2)$
Semi-algebraic pseudodisks in $\mathbb{R}^2$	P/D	$O(r)$
Half-spaces in $\mathbb{R}^d$	P/D	$O(r^{\lfloor d/2 \rfloor})$
Hyperplanes in $\mathbb{R}^d$	P/D	$O(r^d)$

Table 3.1: Bounds on the size of Mnets. Many known results follow from Theorem 3 via their SCC. Poly-logarithmic improvements are in bold. The bounds are for $r \in [1, +\infty)$.

Second, Theorem 3 unifies and generalises several previous statements. Mustafa and Ray [86] gave some results on Mnets using different techniques: for the dual set system induced by regions of union complexity κ using cuttings, for rectangles using divide-and-conquer constructions, and for triangles using nets. All these and more results follow as immediate corollaries of Theorem 3.

Corollary 3.2 (See Table 3.1). *There exist $(1/r)$-Mnets of size*

(i) $O(\kappa(r))$ *for the dual set system induced by semi-algebraic objects in $\mathbb{R}^2$ with union complexity κ. In particular, $O(r \log^* r)$ for the dual set system induced by α-fat triangles*[2], $O\left(r \cdot 2^{O(\log^* r)}\right)$ *for the dual set system induced by locally γ-fat semi-algebraic objects*[3] *in the plane, and $O(r)$ for the dual set system induced by triangles of approximately same size [80].*

(ii) $O(r \log r)$ *for the primal set system induced by axis-parallel rectangles in the plane.*

(iii) $O(r^2)$ *for the primal system induced by lines, strips and cones in the plane, improving the previous-best results by poly-logarithmic factors. They were $O(r^2 \log^2 r)$, $O(r^2 \log^3 r^2)$ and $O(r^2 \log^4 r)$ respectively.*

(iv) $O(r)$ *for the primal set system of semi-algebraic pseudodisks.*

(v) $O(r^{\lfloor d/2 \rfloor})$ *for the primal set system of half-spaces in $\mathbb{R}^d$.*

2. For a fixed parameter α with $0 < \alpha \leq \pi/3$, a triangle is α-fat if all three of its angles are at least α.
3. For a fixed parameter γ with $0 < \gamma \leq 1/4$, a planar semi-algebraic object o is called locally γ-fat if, for any disk D centred in o and that does not fully contain o in its interior, we have area$(D \sqcap o) \geq \gamma \cdot$ area(D), where $D \sqcap o$ is the connected component of $D \cap o$ that contains the centre of D.

(vi) $O(r^d)$ *for the primal and dual set systems of hyperplanes in* $\mathbb{R}^d$.

The main open question in [86] was the following interesting pattern that was observed: for all cases studied, a set system that had $(1/r)$-nets of size $O(r \log \tilde{\varphi}(r))$ had Mnets of size $O(r\tilde{\varphi}(r))$. Theorem 3 now shows that this was not a coincidence. By Theorem A, a set system with SCC $\tilde{\varphi}(\cdot)$ has $(1/r)$-nets of size $O(r \log \tilde{\varphi}(r))$. And now, from Theorem 3, it follows that it has Mnets of size $O(r\tilde{\varphi}(r))$.

2.2 Lower Bounds on the Size of Mnets (Proof in Section 4)

We obtain some lower bounds on the size of $(1/r)$-Mnets in semi-algebraic settings. Mustafa and Ray [86] already gave a number of non-linear lower bounds (in terms of r), e.g. $\Omega\left(r^2\right)$ for primal systems induced by points and lines in the plane and $\Omega\left(r^{\lfloor (d+1)/3 \rfloor}\right)$ for primal set systems induced by points and half-spaces in $\mathbb{R}^d$.

We establish that Theorem 3 is tight for general semi-algebraic set systems with finite VC dimension.

Theorem 4. *Given $r > 0$ and integers n and d, there exists a set P of n points in $\mathbb{R}^d$ such that any c-heavy $(1/r)$-Mnet for the primal set system induced on P by hyperplanes has size $\Omega\left(r^d\right)$, where the constant depends on d and c.*

Note that said primal set system has VC dimension d. Using Lemma 1.1 and Theorem 3, we get that it has $(1/r)$-Mnets of size $O(r^d)$, and Theorem 4 shows that this bound is asymptotically tight.

Kupavskii, Mustafa, and Pach [69] recently gave the following lower bound on the size of nets in terms of SCC.

Theorem F. *Let $\tilde{\varphi}\colon \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonically increasing sub-multiplicative[4] function which tends to infinity and is bounded from above by a polynomial of constant degree. Given any $\delta \in (0, 1/10)$ and $r > r_0(\delta) > 0$, there exists a set system with SCC $\tilde{\varphi}$ that does not have $(1/r)$-nets smaller than*

$$\left(\frac{1}{2} - \delta\right) r \log \tilde{\varphi}(r).$$

Note that the family of sub-multiplicative functions includes commonly appearing functions in discrete and combinatorial geometry such as polynomials, log, $\log \circ \log$, $\log^*$, inverse Ackermann.

Through Corollary 3.1 this lower bound on the size of nets translates directly to a lower bound on the size of Mnets.

4. A function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is **sub-multiplicative** if there exists $x_0 \in \mathbb{R}^+$ such that for every $\alpha \in (0,1)$ and $x > x_0$ we have $f^\alpha(x^{1/\alpha}) \le f(x)$.

Theorem 5. *Let $\tilde{\varphi}\colon \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonically increasing sub-multiplicative function which tends to infinity and is bounded from above by a polynomial of constant degree. Given any $\delta \in (0, 1/10)$ and $r > r_0(\delta) > 0$, there exists a set system with SCC $\tilde{\varphi}$ that does not have $\frac{1}{k}$-heavy $(1/r)$-Mnets smaller than*

$$\Omega\left(r\tilde{\varphi}^{\frac{1-2\delta}{2k}}(r)\right).$$

3 Construction of Mnets

In this section we prove Theorem 3 on the existence of small Mnets in semi-algebraic set systems, then we derive from it the bounds listed in Corollary 3.2. We begin with a brief overview of a technical tool that is used in the proof of Theorem 3.

3.1 Preliminaries on Polynomial Partitioning

We use the following theorem of Matoušek and Patáková [81]. For two subsets γ and ω of $\mathbb{R}^d$, say that γ **crosses** ω if $\omega \cap \gamma \notin \{\emptyset, \omega\}$.

Theorem G (Multilevel polynomial partitioning). *For every integer $d > 1$, there are positive constants K and C such that the following holds. Given a finite set $P \subset \mathbb{R}^d$ and a parameter $\eta > 1$, there exists a partition of P*

$$P = \Sigma^* \cup \bigcup_{k=1}^{d} \bigcup_{l=1}^{t_k} \Sigma_{kl},$$

where $|\Sigma^*| \leq \eta^K$ and the following properties hold for each $k \in [1 \,..\, d]$:

1. *$t_k \leq C\eta^C$, and for $l \in [1 \,..\, t_k]$, $|\Sigma_{kl}| \leq \frac{|P|}{\eta_k}$ with $\eta_k \in [\eta, \eta^K]$.*
2. *there exists a family of semi-algebraic regions $\mathcal{S}_k = \{S_{k1}, \ldots, S_{kt_k}\}$ such that for each $l \in [1 \,..\, t_k]$,*
   a) *S_{kl} is connected, defined by $O(\eta^C)$ polynomial inequalities of degree $O(\eta^C)$,*
   b) *$\Sigma_{kl} \subseteq S_{kl}$, and*
   c) *every set $\gamma \in \Gamma_{d,\Delta,s}$ crosses at most $C_{d,\Delta,s} \cdot \eta_k^{1-1/d}$ of the sets in $\mathcal{S}_k$, where the constant $C_{d,\Delta,s}$ depends only on d, Δ and s.*

In other words, the point set P can be partitioned into a constant number of parts defined by a set $\mathcal{S}_k$ of semi-algebraic regions (for all $k \in [1 \,..\, d]$), such that every set in $\Gamma_{d,\Delta,s}$ either contains, or is disjoint from, most of these regions. Theorem G extends the Guth–Katz *polynomial partitioning technique* [54], a partition of $\mathbb{R}^d$ by an algebraic variety that is balanced with respect to the set P. Here partitioning is applied not once but recursively on varieties of decreasing dimension. This allows us to dispense with assumptions of genericity on P.

3.2 Proofs

We now give proofs of Theorem 3 and Corollary 3.2.

Proof of Theorem 3. Here is a first overview of the proof. Suppose that all ranges in $\mathcal{R}$ are of size exactly ϵn. Consider a maximal $(\epsilon n/2)$-packing $\mathcal{P} \subseteq \mathcal{R}$. It is necessarily (ϵn)-shallow so its cardinality $|\mathcal{P}|$ is bounded by Theorem D. Further, each range of $\mathcal{R}$ has size ϵn and must have an intersection of size at least $\epsilon n/2$ with some range of $\mathcal{P}$, since the latter is maximal. So, to get an ϵ-Mnet, one just needs to get a $(1/2)$-Mnet for each of the set systems $(R, \mathcal{R}_{|R})$ for each $R \in \mathcal{P}$, and take their union. This is where the polynomial partitioning theorem comes in. Fix such a set system $(R, \mathcal{R}_{|R})$. Suppose we could partition the points in R into a constant number c of roughly equal-sized parts, such that each set of size at least $\epsilon n/2$ in $\mathcal{R}|_R$ was a union of some parts in the partition. Then the average size of a part inside a set of size $\epsilon n/2$ in $\mathcal{R}|_R$ is at least $\epsilon n/2c$, and so each such set contains at least one part of size $\geq \epsilon n/2c$. Thus by choosing those parts in R which have at least $\epsilon n/2c$ elements, we are guaranteed to cover all the sets in $\mathcal{R}|_R$ which have at least $\epsilon n/2$ elements. Since we have chosen at most c parts from each $R \in \mathcal{P}$, the total number of sets in $\mathcal{M}$ is at most $c\,|P|$, which is what we wanted. The actual proof involves a finer analysis to show that these ideas go through even when we require that all sets of size *at least* ϵn have to be covered, and we are guaranteed only a partition with a small but possibly nonzero number of crossings.

Now we come to the formal proof. Note that if $\epsilon n = O(1)$, then the trivial collection of singleton sets $\{\{p\} : p \in X\}$ is an ϵ-Mnet for $(X, \mathcal{R})$, of size $n = O(\epsilon^{-1})$. Therefore we may restrict ourselves to the case when

$$\epsilon > \frac{4\,(16 \cdot d \cdot C_{d,\Delta,s})^{Kd}}{n}. \tag{3.1}$$

For $i \in [0..\lceil \log(1/\epsilon) \rceil]$, let $\mathcal{R}_i \subseteq \mathcal{R}$ be an inclusion-maximal $(2^{i-1}\epsilon\, n)$-packing amongst the ranges of $\mathcal{R}$ whose cardinality is in $[2^i \epsilon n,\ 2^{i+1}\epsilon n)$. Denote the elements of $\mathcal{R}_i$ by $\mathcal{R}_{ij}$ for $j \in [1 \,..\, |R_i|]$. From Theorem D, we have

$$|\mathcal{R}_i| \leq \frac{48 d_0}{2^i \epsilon} \cdot \varphi\left(\frac{8 d_0}{2^i \epsilon}, 48 d_0\right). \tag{3.2}$$

For a parameter η to be fixed later, consider the multilevel polynomial partitioning of $\mathcal{R}_{ij}$ as in Theorem G. We write

$$\mathcal{R}_{ij} = \Sigma^*_{ij} \cup \bigcup_{k=1}^{d} \bigcup_{l=1}^{t_{ijk}} \Sigma_{ijkl},$$

where

1. The point set Σ_{ijkl} is included in a connected semi-algebraic region S_{ijkl} of $\mathbb{R}^d$; see Theorem G.
2. $\eta_{ij1},\, \eta_{ij2},\, \ldots,\, \eta_{ijd} \in [\eta,\, \eta^K]$ with $K = K(d)$ as in Theorem G.
3. For all $k = 1,\, 2,\, \ldots,\, d$, $t_{ijk} \leq C\eta^C$, with $C = C(d)$ as in Theorem G. This implies $\sum_{i=1}^{d} t_{ijk} \leq Cd\eta^C$.
4. $|\Sigma_{ijkl}| \leq \frac{|\mathcal{R}_{ij}|}{\eta_{ijk}}$ for all k and l.
5. $\left|\Sigma^*_{ij}\right| \leq \eta^K$.
6. For all $\gamma \in \Gamma_{d,\Delta,s}$ and every $k = 1,\, 2,\, \ldots,\, d$, the number of S_{ijkl} crossed by γ is at most $C_{d,\Delta,s}\eta_{ijk}^{1-1/d}$, where the constant $C_{d,\Delta,s}$ is defined in Theorem G.

The indices $ijkl$ will appear frequently in the sequel. For ease of reference, we remind the reader that i stands for the packing $\mathcal{R}_i$, j stands for the j-th range $\mathcal{R}_{ij}$ in $\mathcal{R}_i$, k indicates the level in the multilevel polynomial partitioning of the set $\mathcal{R}_{ij}$, and l stands for the l-th part in the k-th level.

The Mnet $\mathcal{M}$ is the union of a family $(\mathcal{M}_i)$ of set collections. Let

$$\mathcal{M}_i = \left\{ \Sigma_{ijkl} : |\Sigma_{ijkl}| \geq \frac{2^i \epsilon n}{8Cd\eta^C} \right\} \text{ for each index } i \text{, and}$$
$$\mathcal{M} = \bigcup_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \mathcal{M}_i.$$

It remains to show that $\mathcal{M}$ is the required Mnet for an appropriate value of η. Namely,

(i) the promised bound on $|\mathcal{M}|$ holds,

(ii) each set in $\mathcal{M}$ has size $\Omega(\epsilon n)$, and

(iii) for every $R \in \mathcal{R}$ with $|R| \geq \epsilon n$, there is a choice of (i, j, k, l) such that $R \supseteq \Sigma_{ijkl} \in \mathcal{M}$.

Let $\eta = (16dC_{d,\Delta,s})^d$, ensuring that $\eta^K < \frac{1}{4}\epsilon n$.

To see $i)$, observe that

$$|\mathcal{M}_i| = O\big(d\eta^C \cdot |\mathcal{R}_i|\big) = O\left(\frac{d_0}{2^i\epsilon} \cdot \varphi\Big(\frac{8d_0}{2^i\epsilon}, 48d_0\Big)\right),$$

so as φ is non-decreasing in the first variable,

$$|\mathcal{M}_i| = O\left(\frac{d_0}{2^i\epsilon} \cdot \varphi\Big(\frac{8d_0}{\epsilon}, 48d_0\Big)\right),$$

$$|\mathcal{M}| = \sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} |\mathcal{M}_i| = O\left(\frac{d_0}{\epsilon} \cdot \varphi\Big(\frac{8d_0}{\epsilon}, 48d_0\Big)\right).$$

To see $ii)$, observe that by definition each set added to $\mathcal{M}$ satisfies

$$|\Sigma_{ijkl}| \geq \frac{2^i\epsilon n}{8Cd\eta^C} = \Omega\big(\epsilon n\big).$$

To see $iii)$, let $R \in \mathcal{R}$ be any set such that $|R| \geq \epsilon n$, and let i be the index such that $|R| \in [2^i\epsilon n,\, 2^{i+1}\epsilon n)$. Since $\mathcal{R}_i$ is a maximal packing, there must exist an index j such that $\mathcal{R}_{ij} \in \mathcal{R}_i$ and $|R \,\Delta\, \mathcal{R}_{ij}| \leq 2^{i-1}\epsilon n$. Note that then

$$|R| - |R \,\Delta\, \mathcal{R}_{ij}| \leq |R \cap \mathcal{R}_{ij}| \leq |R|,$$

so that $|R \cap \mathcal{R}_{ij}| \in [2^{i-1}\epsilon n, 2^{i+1}\epsilon n)$.

We shall show that in fact, $R \cap \mathcal{R}_{ij}$ must contain a set $\Sigma_{ijkl} \in \mathcal{M}_i$, by contradiction. Suppose it does not, then consider the contribution to the points in the disjoint union

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

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

$$Cd\eta^C \cdot \frac{2^i\epsilon n}{8Cd\eta^C} = \frac{2^i\epsilon n}{8},$$

because there are at most $Cd\eta^C$ such sets.

b. All points in the parts Σ_{ijkl}, such that the semialgebraic set γ defining $R \cap \mathcal{R}_{ij}$ crosses the connected component S_{ijkl} corresponding to Σ_{ijkl}. It may be that $R \cap \mathcal{R}_{ij} = \mathcal{R}ij$, in which case the number of such points is zero, because for all k, l, $R \cap \mathcal{R}ij = \mathcal{R}ij \supset \Sigma_{ijkl}$. By Theorem G, there are at most $C_{d,\Delta,s}\,\eta_{ijk}^{1-1/d}$ such sets γ, and by the property $(1.)$ of multilevel partitioning, each such region contains at most $\frac{2^{i+1}\epsilon n}{\eta_{ijk}}$ points of X.

c. The points of $R \cap \mathcal{R}_{ij}$ contained in Σ^*_{ij}: at most η^K.

Using the fact that η is sufficiently large in terms of d, Δ and s, we sum these three contributions and obtain:

$$\begin{aligned} |R \cap \mathcal{R}_{ij}| &\leq 2^{i-3}\epsilon n + \sum_{k=1}^{d} \left(\frac{2^{i+1}\epsilon n}{\eta_{ijk}} \cdot C_{d,\Delta,s}\eta_{ijk}^{1-1/d} \right) + \eta^K \\ &< 2^{i-3}\epsilon n + \frac{dC_{d,\Delta,s}2^{i+1}\epsilon n}{\eta^{1/d}} + \eta^K \\ &< 2^{i-1}\epsilon n. \end{aligned}$$

This last inequality follows from the fact that $\eta_{ijk} \geq \eta$, $\eta = (16dC_{d,\Delta,s})^d$ and $\eta^K < 2^{i-2}\epsilon n$. We get a contradiction to the fact that $|R \cap \mathcal{R}_{ij}| \geq 2^{i-1}\epsilon n$, which completes the proof. □

Proof of Corollary 3.2.

1. Remember that the dual of a semi-algebraic set system is itself semi-algebraic (Lemma 1.3). The SCC of the dual set system induced by objects with union complexity κ is $\tilde{\varphi}(m) = O(\kappa(m)/m)$, which together with Theorem 3 implies the stated bound. The value of $\kappa(m)$ is $O(m)$ for triangles with approximately same size [80], $O(m \log^* m)$ for α-fat triangles [48](where the constant of proportionality depends only on α), and $O(m2^{\log^* m})$ for locally γ-fat objects [9], where the constant of proportionality in the linear term depends only on γ.

2. Let $(X, \mathcal{R})$ be the set system induced on a set X of n points in $\mathbb{R}^2$ by the family of axis-parallel rectangles. Aronov, Ezra, and Sharir [10] show that there is another set system $\mathcal{R}'$ on X with SCC $\varphi_{\mathcal{R}'}(m, l) = O(l \cdot \log m)$, such that for any $R \in \mathcal{R}$, there exists a $R' \in \mathcal{R}'$ such that $R' \subseteq R$ and $|R'| \geq |R|/2$. Thus a $(1/2r)$-Mnet for $\mathcal{R}'$ is a $(1/r)$-Mnet for $\mathcal{R}$, of size $O(r \log r)$.

3. The other results follow from the facts that the SCC $\varphi(m, l)$
   - of the primal set system induced on $\mathbb{R}^2$ is $O(m)$ for lines, $O(ml)$ for strips, and $O(ml^2)$ for cones [86].
   - of the primal set system induced on $\mathbb{R}^2$ by pseudodisks is $O(l^2)$ [27].
   - of the primal set system induced by half-spaces on $\mathbb{R}^d$ is $O(m^{\lfloor d/2 \rfloor - 1} l^{\lceil d/2 \rceil})$.
   - of the primal and dual set systems induced by hyperplanes on $\mathbb{R}^d$ is $O(m^d)$.

□

4 Main Lower Bound on Mnets

To prove Theorem 4, we use the following result from [86].

Lemma 3.1. *For $\epsilon > 0$ and positive integers d and k, there exists a set P of n points in $\mathbb{R}^2$ and a set $\mathcal{C}$ of curves of polynomial functions of degree at most d satisfying the following conditions:*

(i) $|\mathcal{C}| = \Omega\left(\epsilon^{-d}\right)$ *where the constant of proportionality depends on d and k.*

(ii) *For all $C_i \in \mathcal{C}$, $|C_i \cap P| \geq \epsilon n$.*

(iii) *For any two distinct C_1 and C_2 in $\mathcal{C}$, we have $|C_1 \cap C_2 \cap P| \leq \frac{\epsilon n}{k}$.*

Let P and $\mathcal{C}$ be as in Lemma 3.1. By the standard Veronese lifting [78], they are mapped to a set P' of n points in $\mathbb{R}^d$ and a family of hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ satisfying the following conditions:

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

Now any $\frac{1}{k}$-heavy ϵ-Mnet for the primal set system defined by $\mathcal{H}$ on P' must have at least $|\mathcal{H}| = \Omega\left(\epsilon^{-d}\right)$ sets, which proves Theorem 4.

5 Remarks

New approach to understand ϵ-nets. We showed that Mnets gives a new approach to understand ϵ-nets for geometric set systems. Corollary 3.1 gives a proof of optimal ϵ-nets that is independent of Haussler and Welzl's work.

Applications of Mnets. Corollary 3.1 shows that the existence of small nets follows immediately from the more general structure of Mnets. Macbeath regions for convex bodies have found algorithmic applications such as volume estimation of convex bodies [11, 13]. We believe that Mnets will also find important applications and connections to various aspects of set systems with bounded VC dimension.

5.1 Computing Mnets

In the real RAM model of computation one can compute exactly with arbitrary real numbers and each arithmetic operation takes unit time. Matoušek and Patáková [81] gave the following algorithmic counterpart of Theorem G.

Theorem H (Algorithmic Multilevel Polynomial Partitioning). *The sets Σ^*, Σ_{ij}, S_{ij} from Theorem G can be computed in time $O(nr^C)$, where the constant C is the same as in Theorem G.*

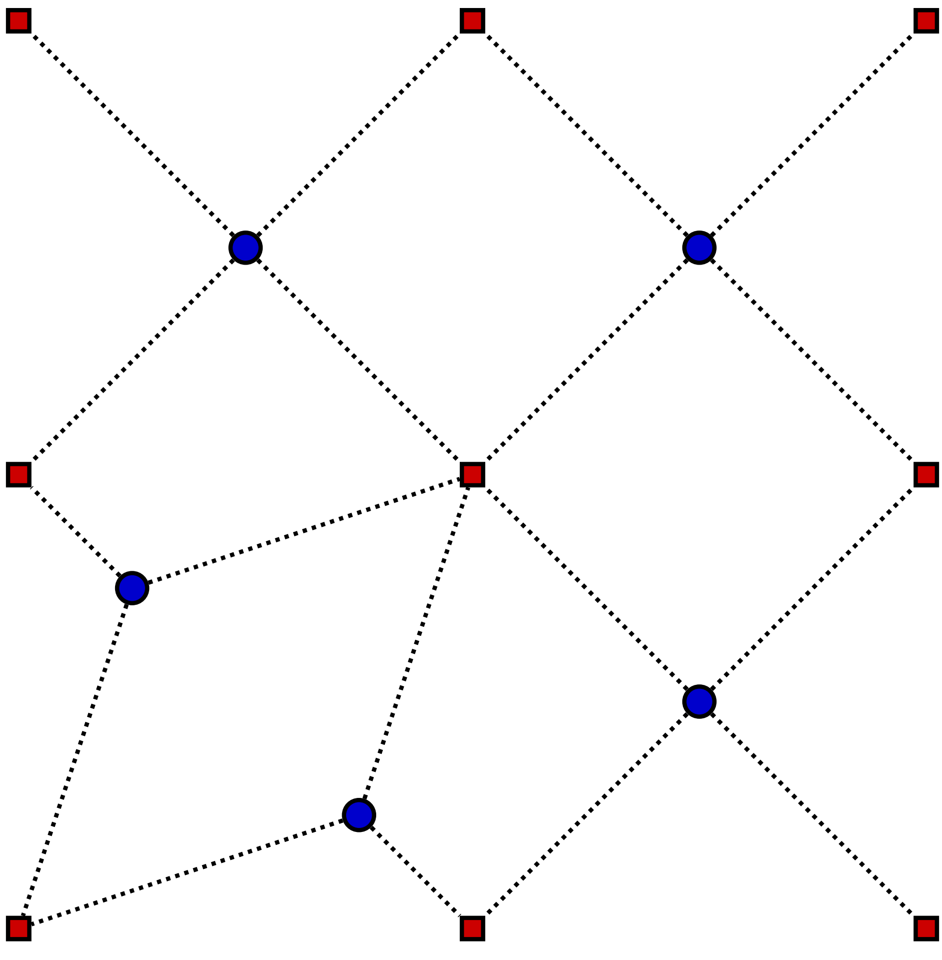
Using this result and the construction in the proof of Theorem 3, we can get a randomised algorithm with time complexity poly $(n, 1/\epsilon)$ that computes Mnets for semi-algebraic set systems matching the upper bound on the size of Mnets from Theorem 3.

We finish with our main open problem.

Question. Can the optimal bounds on *weighted* nets (see e.g. [99], [32]) be obtained via Mnets?

Part II

Geometric Local Search



Chapter Four

Hardness of Approximation

IN THIS CHAPTER we introduce a few complexity classes related to optimisation problems, especially classes associated with approximation algorithms and parametrised tractability. This is not intended as a description of complexity theory from the ground up; several surveys such as [43, 82] contain all notions needed and more. A convenient online resource is the Complexity Zoo [40]. Our goal is to prepare for Chapter 5 by explaining why some geometric optimisation problems are hard to approximate.

1 Combinatorial Optimisation Problems

We mostly follow the introduction to approximation algorithms by Jansen [63]. Where decision problems are yes–no questions, optimisation problems may have multiple possible answers, out of which one has to pick the best.

Definition 4.1. An **optimisation problem** is a tuple $(\mathcal{I}, \mathcal{S}, \mathrm{val}, \prec)$ where $\mathcal{I}$ is a set of instances, $\mathcal{S}$ is a function with domain $\mathcal{I}$ that maps each instance to its nonempty set of feasible solutions, the operator val associates a positive integer to every solution, and $\prec$ is either $\leq$ (*minimisation*) or $\geq$ (*maximisation*).

Example 4.1. MINIMUM VERTEX COVER is a minimisation problem whose instances are all finite graphs with at least one edge. Its function $\mathcal{S}$ maps a graph to the set of its vertex covers (its sets of vertices that cover all edges) and the value of a solution is its cardinality.

Example 4.2. TRAVELLING SALESPERSON or TSP is a minimisation problem whose instances are all finite complete graphs with edge weights. The feasible solutions of an instance are its Hamiltonian circuits and the value of a solution is the sum of its edge weights.

We write $\mathrm{OPT}(\Pi, I)$ for the $\prec$-minimal value of all solutions of the problem Π for the instance I, or simply OPT when both the problem and the instance are clear from the context. A solution that attains this value is an **optimal solution**.

The optimisation problem belongs to the class **NPO** or **NP**-optimisation if $\mathcal{I}$ is in **P** (can be recognised in polynomial time), for each instance I the set $\mathcal{S}(I)$ is in **P**, and val is computable in polynomial time.

To any such optimisation problem corresponds a natural decision problem: given an instance and a positive integer k, is $\mathrm{OPT} \prec k$? If the optimisation problem is in **NPO**, then its decision version is in **NP**. Furthermore **PO** (for **P**-optimisation) is the class of optimisation problems whose corresponding decision versions are in **P**.

2 Approximation Algorithms

Definition 4.2. An **approximation algorithm** with factor $r \in [1, +\infty)$ for some optimisation problem (an r-approximation for short) is an algorithm that for each instance returns a solution whose value is within a factor r of OPT.

Example 4.3. There is a well-known 2-approximation for Minimum Vertex Cover. Find a maximal matching and let t be its number of edges. The set of all its endpoints is a vertex cover with $2t$ vertices and it is clear that $\mathrm{OPT} \geq t$.

Optimisation problems can be classified according to their approximability properties. A problem in **NPO** is in **APX** if it admits a polynomial-time approximation, in **PTAS** if it admits a polynomial-time $(1+\epsilon)$-approximation for all $\epsilon > 0$, in **EPTAS** if said algorithms have running time $f(\epsilon) \cdot \mathrm{poly}\, n$ for some arbitrary function f, and in **FPTAS** when said running time is $\mathrm{poly}(\epsilon^{-1}, n)$. Somewhat confusingly perhaps, algorithms of those types are also called PTAS, EPTAS and FPTAS respectively. These abbreviations stand for (*efficient* or *fully*) *polynomial-time approximation scheme.*

3 Parametrised Complexity

3.1 Fixed-Parameter Tractability

The decision version of an optimisation problem is in the class **FPT** if there is an algorithm that for every instance I and positive integer k decides $\mathrm{OPT} \prec k$ in time $f(k) \cdot \mathrm{poly}\, |I|$ with f a computable function. (There are many more ways to *parametrise* a problem, but as we only use this parametrisation by the value we may specialise our discussion of **FPT**.)

Example 4.4. The decision version of Minimum Vertex Cover is the problem of deciding whether an input graph on n vertices admits a vertex cover of cardinality at most k. This can be answered in time $O(2^k n)$, meaning that MINIMUM VERTEX COVER is in **FPT**.

3.2 Lowest Levels of the W Hierarchy

Definition 4.3. An **fpt-reduction** from the decision version of an optimisation problem $\Pi_1 = (\mathcal{I}_1, \mathcal{S}_1, \mathrm{val}_1, \prec_1)$ to another problem of the same kind $\Pi_2 = (\mathcal{I}_2, \mathcal{S}_2, \mathrm{val}_2, \prec_2)$ is an algorithm

$$\begin{aligned} M_{1\to 2}\colon \mathcal{I}_1 \times \mathbb{N}^* &\to \mathcal{I}_2 \times \mathbb{N}^* \\ (I_1, k_1) &\mapsto (I_2, k_2) \end{aligned}$$

such that

- $M_{1\to 2}$ runs in time at most $g(k_1)\,\mathrm{poly}\,|I_1|$,
- $k_2 \leq f(k_1)$,
- and $\mathrm{OPT}(\Pi_1, I_1) \prec_1 k_1 \iff \mathrm{OPT}(\Pi_2, I_2) \prec_2 k_2$

for some computable functions f and g.

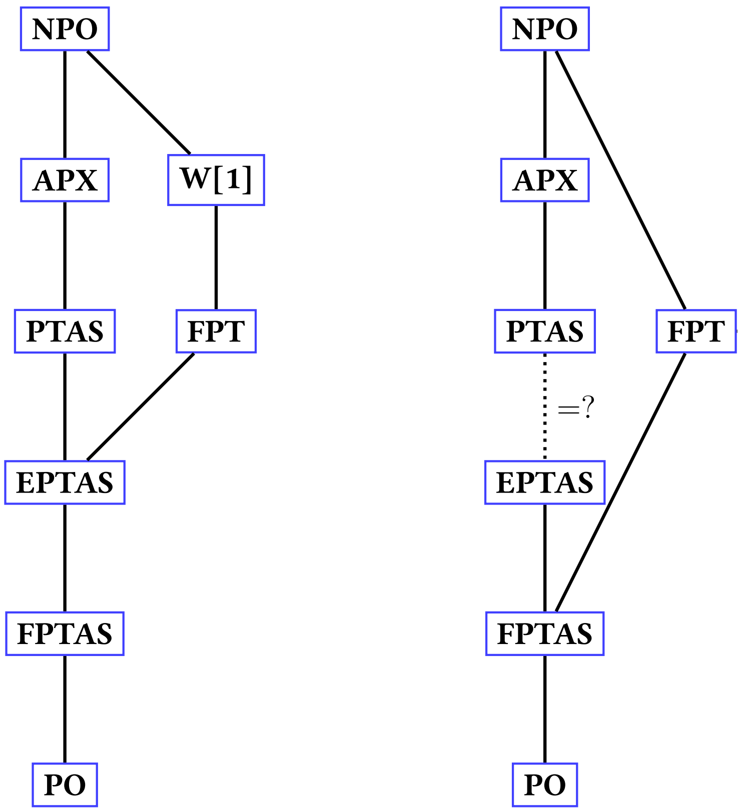
Example 4.5. The reduction from INDEPENDENT SET to CLIQUE is straightforward. To a pair (G, k) with G a graph of order n it associates $(\tilde{G}, k)$, where $\tilde{G}$ is the complement of G (the graph with the same vertex set whose edges are those not present in G). Clearly $\tilde{G}$ can be computed from G in time $\mathrm{poly}\,n$, the identity of $\mathbb{N}^*$ can be taken for f, and G has an independent set of cardinality at least k if and only if $\tilde{G}$ has a clique of cardinality at least k.

The class **W[1]** is defined with Boolean circuits of restricted complexity. For our purposes it is sufficient to know that a problem is **W[1]**-hard if it is fpt-reducible to the decision version of CLIQUE, or to any other **W[1]**-hard problem. As the previous example indicates this is the case for INDEPENDENT SET.

3.3 Complexity Bounds

The inclusion **FPT** $\subseteq$ **W[1]** is commonly thought to be strict as equality of those classes would imply failure of the exponential time hypothesis. Under this assumption, the inclusions of complexity classes presented in Figure 4.1a are all strict (with some abuse of language in that optimisation problems are identified with their decision version).

Recall that an EPTAS is a PTAS that runs in time $f(\epsilon)\,n^c$ for an arbitrary function f and $c \geq 0$. The following observation first appeared in C. Bazgan's master's thesis [19, 29].



(a) State of the world if **FPT** is a proper subclass of **W[1]**.

(b) State of the world if **W[1]** $=$ **FPT** but **P** $\neq$ **NP**.

Figure 4.1: Hasse diagram of some complexity classes depending on the relative status of **FPT** and **W[1]**.

Theorem I. *Combinatorial optimisation problems with integral-valued objective functions and efficient PTASs are in **FPT**.*

Proof. Given a minimisation problem with an efficient PTAS in $f(\epsilon)\,n^c$, for each instance we can find in time $f(1/2k)n^c$ a solution with value V such that

$$\left(1+\frac{1}{2k}\right)\cdot \mathrm{OPT} \geq V \geq \mathrm{OPT}.$$

If $V \leq k$ this implies OPT $\leq k$, whereas if $V \geq k+1$ then

$$\mathrm{OPT} \geq \frac{2kV}{2k+1} \geq \frac{2k^2+2k}{2k+1} > k,$$

so computing V answers the question 'Does the instance have OPT $\leq k$?' in time $f(1/2k)n^c$, i.e. the problem parameterised by the size of its solution is in **FPT**. $\square$

Since the problems that we will consider in the sequel are **W[1]**-hard, we conclude that none of them admits efficient PTASs unless the inclusion of **FPT** in **W[1]** is an equality.

4 A Few Optimisation Problems

Let us define the problems that we will encounter in the next chapters.

4.1 Problems on General Graphs and Set Systems

In the **Minimum Hitting Set** problem, we are given a (finite) set system and the goal is to find a smallest transversal. (Recall Definition 1.10.) The restricted version of this problem when the input set system is a graph—that is, all ranges have cardinality two—is **Minimum Vertex Cover**.

In the **Minimum Set Cover** problem, the input is again a set system but the goal is to find a smallest subset of ranges whose union is the ground set. Minimum Hitting Set and Minimum Set Cover are equivalent problems via set-system duality.

In the **Maximum Independent Set** problem, the input is a graph and the goal is to compute a largest subset of pairwise non-adjacent vertices (such subsets are *independent*).

In the **Minimum Dominating Set** problem, the input is a graph and we search for a smallest subset D of vertices such that every vertex of the graph is in D or has an edge to a vertex of D.

4.2 Geometric Problems

Geometric optimisation problems are typically obtained by specialising the inputs to these general problems. For example, the inputs to Minimum Hitting Set of pseudodisks are points and pseudodisks in $\mathbb{R}^2$. Remember that a family of pseudodisks consists of the interior regions of a family of Jordan curves such that any two curves intersect at most twice. Here are the problems that we will encounter in the next chapter.

Minimum Hitting Set for pseudodisks: given a set $\mathcal{P}$ of points and a family $\mathcal{D}$ of pseudodisks in the plane, compute a smallest subset of $\mathcal{P}$ that intersects all pseudodisks in $\mathcal{D}$.

Maximum Independent Set of pseudodisks: given a family $\mathcal{D}$ of pseudodisks in the plane, compute a maximum size subset of pairwise disjoint pseudodisks in $\mathcal{D}$.

Minimum Dominating Set of pseudodisks: given a family $\mathcal{D}$ of pseudodisks in the plane, compute a smallest subset of pseudodisks of $\mathcal{D}$ that together intersect all other pseudodisks of $\mathcal{D}$.

Minimum Set Cover for disks: given a set $\mathcal{P}$ of points and a family $\mathcal{D}$ of disks in the plane, return a smallest subset of disks in $\mathcal{D}$ that together cover all points of $\mathcal{P}$.

Unit-Capacity Point-Packing: given a set of points $\mathcal{P}$ and a set of disks $\mathcal{D}$, compute a largest subset of $\mathcal{P}$ that hits no disk of $\mathcal{D}$ more than once. This last problem was introduced by Ene, Har-Peled, and Raichel [46].

5 Hardness Results

We prepare for the next chapter by proving the **W[1]**-hardness of some geometric problems. Under the assumption that **FPT** is strictly included in **W[1]** this means that they cannot be fixed-parameter tractable (when parametrised by the cardinality of the optimal solution) and then by Theorem I that they cannot admit efficient PTASs.

Marx already gave hardness results for two of our problems, even when the input is restricted to disks of unit radius.

Theorem J ([75]). *Maximum Independent Set of unit disks is **W[1]**-complete (and in particular **W[1]**-hard).*

Theorem K ([76]). *Minimum Dominating Set of unit disks is **W[1]**-hard.*

This second result extends to Hitting Set by a simple reduction. All reductions that we give are very simple examples of fpt-reductions as in Definition 4.3 with $f : k \mapsto k$ and $g \equiv 1$.

Corollary 4.1. *Minimum Hitting Set of unit disks is **W[1]**-hard.*

Proof. Given an instance of Dominating Set with n unit disks, let $\mathcal{P}$ consist of their centres and $\mathcal{D}$ consist of the n radius-2 disks centred at $\mathcal{P}$. The hitting sets of $\mathcal{D}$ in $\mathcal{P}$ are exactly the dominating sets of the initial unit disks. □

Corollary 4.2. *Minimum Set Cover by unit disks is **W[1]**-hard.*

Proof. Given an instance of Hitting Set with unit disks $\mathcal{D}$ and points $\mathcal{P}$, let $\mathcal{P}'$ consist of the centres of $\mathcal{D}$ and $\mathcal{D}'$ contain one unit disk centred at each point in $\mathcal{P}$. The set covers of $\mathcal{P}'$ in $\mathcal{D}'$ are exactly the hitting sets of $\mathcal{D}$ in $\mathcal{P}$. □

Now recall the Unit-Capacity Point-Packing problem defined on page 76. Unit-Capacity Disk-Packing, also introduced in [46], is the dual problem where we are given a set $\mathcal{D}$ of disks and at set $\mathcal{P}$ of points and must return a largest subset of $\mathcal{D}$ that covers every point of $\mathcal{P}$ at most once.

Corollary 4.3. *Unit-Capacity Disk-Packing is **W[1]**-hard even with unit disks.*

Proof. Given an instance of Maximum Independent Set with n unit disks, let $\mathcal{D}$ consist of these disks and $\mathcal{P}$ contain one point from each nonempty pairwise intersection in $\mathcal{D}$. The set $\mathcal{P}$ has size $O(n^2)$ and can be computed in $O(n^2)$. Now the feasible solutions of Unit-Capacity Disk-Packing for $\mathcal{D}$ and $\mathcal{P}$ are exactly the independent sets of $\mathcal{D}$. $\square$

Corollary 4.4. *Unit-Capacity Point-Packing is also **W[1]**-hard, even with unit disks.*

Proof. Given an instance of Unit-Capacity Point-Packing with unit disks $\mathcal{D}$ and points $\mathcal{P}$, let $\mathcal{D}'$ consist one unit disk centred at each point in $\mathcal{P}$ and $\mathcal{P}'$ contain the centres of the disks in $\mathcal{D}$. The feasible solutions of Unit-Capacity Disk-Packing for $\mathcal{D}$ and $\mathcal{P}$ are exactly the feasible solutions of Unit-Capacity Point-Packing for $\mathcal{D}'$ and $\mathcal{P}'$. $\square$

Chapter Five

Approximation Guarantee of Geometric Local Search

In this chapter we analyse geometric algorithms for combinatorial optimisation based on local search. Such algorithms have been extremely fruitful in the last decade.

More precisely, the algorithmic status of several basic **NP**-complete problems in geometric combinatorial optimisation were unresolved. This included the existence of polynomial-time approximation schemes (PTASs) for Hitting Set, Set Cover, Dominating Set, Independent Set, and other problems for some basic geometric objects. Other the past past nine years all have been solved—interestingly, with the same algorithm: local search. In fact, it was shown that for many of these problems, local search with radius λ gives a $(1 + O(1/\sqrt{\lambda}))$-approximation with running time $n^{O(\lambda)}$. Setting $\lambda = \Theta(\epsilon^{-2})$ yields a PTAS with a running time of $n^{O(\epsilon^{-2})}$.

On the other hand, hardness results suggest that a PTAS for any of these problems cannot run in $\text{poly}(n) \cdot f(\epsilon)$ for any arbitrary f. Thus the main question left open in previous work is whether the exponent of n can be decreased to $o(\epsilon^{-2})$.

We are able to establish a negative result: we show that in fact the approximation guarantee of local search cannot be improved for any of these problems. The key ingredient, of independent interest, is a new lower bound on locally expanding planar graphs, which is then used to show the impossibility results. Our construction extends to other graph families with small separators.

These results were obtained with Nabil H. Mustafa and presented at the 34th International Symposium on Computational Geometry [64].

1 General Principles of Local Search

1.1 Space of Feasible Solutions

Computing an optimal solution for **NP**-hard combinatorial optimisation problems typically involves exploring all (or a constant fraction of) feasible solutions, which is costly. The strict inclusion of **P** in **NP** and the exponential time hypothesis are two common assumptions that extend this observation; they posit respectively that **NP**-complete problems cannot be solved in polynomial or even sub-exponential time, i.e. an algorithm that enumerates and evaluates all solutions is roughly optimal.

1.2 Local Search

This motivates techniques such as **local search** that restrict the domain to be investigated in the solution space. Given a notion of neighbourhood in this space, a local search starts from an initial solution then visits neighbouring solutions as long as such moves improve the value of the solution.

Algorithm 1 Local search algorithm

```
 1: procedure FEASIBLE(S)                    ▷ Problem-specific feasibility test.
 2:     ...
 3: procedure NEIGHBOURS(S)                  ▷ Lists all neighbouring solutions.
 4:     ...
 5: procedure LOCALSEARCH(s_0)
 6:     s ← s_0
 7:     for s' in Neighbours(s) do
 8:         if val s' ≺ val s then
 9:             s ← s'
10:             jump to line 7
11:     return s                             ▷ The solution s is locally optimal.
```

Given a solution space $\mathcal{S}$ and a neighbourhood operator $\mathcal{N}\colon \mathcal{S} \to \mathcal{P}(\mathcal{S})$ with $s \in \mathcal{N}(s)$ for every $s \in \mathcal{S}$, the $\mathcal{N}$-**locally optimal solutions** are the $s \in \mathcal{S}$ such that

$$\forall s' \in \mathcal{N}(s), \quad \mathrm{val}\, s \prec \mathrm{val}\, s'. \tag{5.1}$$

All problems that we are interested in have for their feasible solutions elements of $\mathcal{P}(X)$, where the finite set X is part of the instance. For example, the feasible solutions of MINIMUM VERTEX COVER are subsets of the vertex set of the graph given in the instance. Additionally for all problems considered the value of a solution will be its cardinality (which we either want to maximise or minimise depending on $\prec \in \{\leq, \geq\}$).

In this setting, our notion of neighbourhood is based on the Hamming distance (the solutions can be seen as binary words of length $|X|$). In a λ-local search we try to improve the current solution by removing t elements from it and adding at most $t-1$ new ones for any $t \in [1 \mathrel{..} \lambda]$ (for a minimisation problem). For this notion of neighbourhood, condition (5.1) becomes

$$\forall s' \in \mathcal{S}, \quad |s \,\Delta\, s'| < 2\lambda \implies |s| \prec |s|' .$$

Letting $n = |X|$, every improvement step is executed in time $n^{O(\lambda)}$ and there cannot be more than n such steps since each decreases the cardinality of the solution. Thus λ-local search obtains a λ-locally optimal solution in time $n^{O(\lambda)}$.

Algorithm 2 Local search algorithm with radius λ, cardinalities as objective

1: **procedure** FEASIBLE(s) ▷ Problem-specific feasibility test.
2:     …
3: **procedure** LOCALSEARCH(X)
4:     $s \leftarrow X$
5:     **for** $s_1 \subseteq s, |s_1| \leq \lambda$ **do**
6:         **for** $s_2 \subseteq X \setminus s, |s_2| \leq \lambda$ **do**
7:             $s' \leftarrow s \cup s_2 \setminus s_1$
8:             **if** Feasible(s') and $|s'| \prec |s|$ **then**
9:                 $s \leftarrow s'$
10:                 **jump to** line 5
11:     **return** s ▷ s is locally optimal.

1.3 Locality Gap and the Efficiency of Local Search

Based on prior use by Arya et al. [14] and mirroring the *integrality gap* from linear programming, we introduce the notion of locality gap.

Definition 5.1. For a given optimisation problem, the **locality gap of an instance** is the function

$$\lambda \in \mathbb{N} \mapsto \max \left\{ \max \left\{ \frac{|s|}{\mathrm{OPT}}, \frac{\mathrm{OPT}}{|s|} \right\} : s \ \lambda\text{-locally optimal solution} \right\},$$

which is the maximal multiplicative ratio between the value of an optimal solution and that of a λ-locally optimal one. The **locality gap of the problem** itself is the supremum of this ratio over all instances.

Since $(\lambda+1)$-local optimality also implies λ-local optimality, these are both non-increasing functions of λ.

There is no obvious reason why the locality gap of a problem should be bounded or even finite, and it often is not. However it turns out—and this is the foundation of local search techniques in computational geometry—that for many optimisation problems arising from geometry it is a somewhat quickly decreasing function of λ.

In fact it has been shown that for several problems that are somehow characterised by planarity this ratio is asymptotically $1 + O(\lambda^{-1/2})$. This is described in Section 2. An immediate consequence is that those problems admit $(1+\epsilon)$-approximations in time $O(n^{O(\epsilon^{-2})})$ and belong thus to the **PTAS** class, although they may be otherwise computationally hard. Amongst our contributions to the study of geometric local search techniques, we prove a matching lower bound of $1 + \Omega(\lambda^{-1/2})$, meaning that the existing analysis of local search for those problems is tight.

2 Geometric PTASs Based on Local Search

Within the past decade polynomial-time approximation schemes (PTASs) have been proposed for a number of long-standing open problems in geometric approximation algorithms, including the following **NP**-hard problems (see [36, 62] for hardness results):

1. Minimum Hitting Set for pseudodisks [87],
2. Maximum Independent Set of pseudodisks [3, 33],
3. Minimum Dominating Set of pseudodisks [52, 53],
4. Minimum Set Cover for disks [15, 31],
5. Unit-Capacity Point-Packing [46].

2.1 Approximation Guarantees for Local Search

Surprisingly, the PTAS for all these problems is essentially the same: local search. Let X be the set of base elements of the problem (for example, this would be the input points in the Hitting Set problem), and let the **search radius** $\lambda \geq 3$ be an integer. Then start with any feasible solution $\mathcal{L} \subseteq X$ and increase (in the case of a maximisation problem, e.g. Maximum Independent Set) or decrease (in the case of a minimisation problem, e.g. Minimum Hitting Set) its size by local improvement steps while maintaining feasibility. Here a **local improvement step** is to swap a subset $\mathcal{L}'$ of at most λ elements of the current solution $\mathcal{L}$ with a subset of $X \setminus \mathcal{L}$ of size at least $|\mathcal{L}'| + 1$ (for maximisation problems) or at most $|\mathcal{L}'| - 1$ (for minimisation problems), as long as the new solution is still feasible. The algorithm finishes when no local improvement step is possible. Such a solution is called **λ-locally optimal**.

All these algorithms are analysed in a similar way, as follows. Let $\mathcal{L}$ be a λ-locally optimal solution and $\mathcal{O}$ be an optimal solution. We can assume that these solutions are disjoint by considering $\mathcal{L} \setminus \mathcal{O}$ and $\mathcal{O} \setminus \mathcal{L}$. To relate the cardinalities of $\mathcal{L}$ and $\mathcal{O}$, a bipartite **exchange graph** is built on vertex sets $\mathcal{L}$ and $\mathcal{O}$ with a local vertex expansion property:

$$\textbf{Minimisation:} \text{ for all } \mathcal{L}' \subseteq \mathcal{L} \text{ of size at most } \lambda,\ |N(\mathcal{L}')| \geq |\mathcal{L}'|\,. \tag{5.2}$$

$$\textbf{Maximisation:} \text{ for all } \mathcal{O}' \subseteq \mathcal{O} \text{ of size at most } \lambda,\ |N(\mathcal{O}')| \geq |\mathcal{O}'|\,. \tag{5.3}$$

(When a graph G is clear from the context and V' is a subset of its vertices, $N(V')$ denotes the set of neighbours of the vertices of V' in G.)

The construction of exchange graphs is problem-specific and exploits the geometric properties of optimal and local solutions. For example, in the Minimum Vertex Cover problem on a graph G this would simply be the bipartite subgraph of G induced by $\mathcal{L}$ and $\mathcal{O}$; condition (5.2) follows from the local optimality of $\mathcal{L}$.

The key in the analysis lies in a general theorem on local expansion in sparse graphs. A bipartite graph on vertex sets (B, R) is **λ-expanding** if for all $B' \subseteq B$ of size at most λ we have $|N(B')| \geq |B'|$. Note that the roles of B and R in this definition are *not* symmetric. A **(vertex) separator** of a graph on n vertices is a subset of vertices whose removal leaves connected components of cardinality at most $\frac{2}{3}n$. A class of graphs $\mathcal{G}$ has the **separator property** with parameter $s \in [0, 1]$ if there exists a positive constant c such that any graph in $\mathcal{G}$ has a separator of size at most cn^{1-s}, where n is the number of vertices. For example, trees have this property with $s = 1$ as they have constant-sized separators, whereas planar graphs have the separator property with parameter $s = \frac{1}{2}$. In fact, the separator property with $s = \frac{1}{2}$ actually holds for graphs excluding fixed minors and in particular for minor-closed classes other than the class of all graphs, e.g. graphs of bounded genus [7]. A class of graphs closed under taking subgraphs is called **monotone**.

Theorem L ([15, 33, 87])**.** *If a finite and λ-expanding bipartite graph on (B, R) belongs to a monotone family with the separator property with parameter $s \in (0, 1)$ and $\lambda \geq \lambda_s$, then $|B| \leq (1 + c_s\lambda^{-s}) \cdot |R|$, where c_s and λ_s are positive constants that depend only on s.*

In an independent paper, Cabello and Gajser [28] describe a subcase of this theorem for K_h-minor-free graphs, which have separators of size $O\left(h^{3/2}\sqrt{n}\right)$. Finally, Har-Peled and Quanrud [57, 58] observe that intersection graphs of low-density objects in $\mathbb{R}^d$ have the separator property with $s = 1/d$.

To complete the analysis for minimisation problems, apply Theorem L with $B = \mathcal{L}$ and $R = \mathcal{O}$, and get $|\mathcal{L}| \leq (1 + c_s\lambda^{-s}) \cdot |\mathcal{O}|$. For maximisation problems, take $B = \mathcal{O}$ and $R = \mathcal{L}$, and get $|\mathcal{O}| \leq (1 + c_s\lambda^{-s}) \cdot |\mathcal{L}|$ or equivalently $|\mathcal{L}| \geq (1 - c'_s\lambda^{-s}) \cdot |\mathcal{O}|$.

2.2 Computational Efficiency of Geometric Local Search

Given a positive parameter ϵ, local search with radius $\lambda = \Theta(\epsilon^{-1/s})$ provides a $(1+\epsilon)$-approximate solution to problems whose exchange graphs have the separator property with parameter s. This can be implemented in $n^{O(\lambda)}$ time by considering all possible local improvements, thus yielding a PTAS in time $n^{O(\epsilon^{-1/s})}$, and in particular $n^{O(\epsilon^{-2})}$ for the five problems listed on page 81.

The parameterised versions of these problems are **W[1]**-hard: even for unit disks, Maximum Independent Set is **W[1]**-complete [75] and Minimum Dominating Setis **W[1]**-hard [76], and the latter is easily reduced to our other three problems. Under the common assumption that **FPT** $\subsetneq$ **W[1]**, which follows from the exponential time hypothesis, these problems do not admit PTASs with time complexity poly$(n) \cdot f(\epsilon)$ for any arbitrary function f. In other words, the dependence of the exponent of n on ϵ is inevitable. The details were given in the previous chapter, Section 5.

Still, this running time is prohibitively expensive, and there have been two complementary approaches towards further progress: firstly, careful implementations of local search that find local improvements more efficiently than by brute force [25]. The second, more structural approach is to better analyse the quality of solutions resulting from local search algorithms, mainly by studying the properties of exchange graphs [8].

2.3 Contributions: Limits of Geometric Local Search

The construction that we give in Section 3 shows that Theorem L is asymptotically tight whenever $1/s$ is an integer.

Theorem 6. *Given a positive integer d, there are positive constants c_d and λ_d such that, for every integer $\lambda \geq \lambda_d$, there is a family of bipartite graphs $(B_n, R_n; E_n)$ indexed by $n \in \mathbb{N}$ that*

- *are λ-expanding,*
- *have the separator property for $s = 1/d$, and so do their subgraphs,*
- *satisfy $|B_n|, |R_n| = \Theta(n)$ and $|B_n| \geq (1 + c_d \cdot \lambda^{-\frac{1}{d}}) |R_n| - o(|R_n|)$ as $n \to \infty$.*

Furthermore when $d = 2$ they are Gabriel graphs.

(A graph (V, E) is called **Gabriel** if there exists a mapping $f : V \to \mathbb{R}^2$ such that $\{v_i, v_j\} \in E$ if and only if the circumdisk of the segment $f(v_i)f(v_j)$ contains no other point of $f(V)$. Gabriel graphs are subgraphs of Delaunay triangulations and thus planar.)

Remark 5.1. Since our construction for $d = 2$ is planar, previous analogues of Theorem L restricted to planar graphs are also tight.

2.4 Algorithmic Consequences

The analysis of local search in terms of the radius that achieves a $(1+\epsilon)$-approximation is tight for the five problems listed earlier (which all had $s = \frac{1}{2}$), as well as for a few other problems with small separators (Section 4).

Theorem M ([87]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for* Minimum Hitting Set *for pseudodisks.*

Corollary 5.1. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there is a set $\mathcal{D}$ of at least n disks and two disjoint sets B and R of at least n points in $\mathbb{R}^2$ each such that both B and R are hitting sets for $\mathcal{D}$, $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and B is a λ-locally optimal solution to the* Hitting Set *problem for $\mathcal{D}$ with $\mathcal{P} = B \cup R$.*

Theorem N ([33]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for* Maximum Independent Set *for pseudodisks.*

Corollary 5.2. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two independent sets B and R of at least n disks in $\mathbb{R}^2$ such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the* Independent Set *problem in $B \cup R$.*

Theorem O ([15, 31]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for* Minimum Set Cover *for disks.*

Corollary 5.3. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two independent sets B and R of at least n disks in $\mathbb{R}^2$ and a set $\mathcal{P}$ of $\Theta(|R|)$ points in $\mathbb{R}^2$ such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the* Set Cover *problem for $\mathcal{P}$ in $B \cup R$.*

Theorem P ([52, 53]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for* Minimum Dominating Set *for pseudodisks.*

Corollary 5.4. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there is a set $\mathcal{D}$ of disks in $\mathbb{R}^2$ and two dominating sets B and R of $\mathcal{D}$ of at least n disks each such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and B is a λ-locally optimal solution to the* Dominating Set *problem for $\mathcal{D}$.*

Theorem Q ([46]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for the* Unit-Capacity Point-Packing *problem for disks.*

Corollary 5.5. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two sets B and R of at least n points in $\mathbb{R}^2$ and a set $\mathcal{D}$ of $\Theta(|R|)$ disks in $\mathbb{R}^2$ such that every disk of $\mathcal{D}$ contains one point from B and one point from R, $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the Unit-Capacity Point-Packing problem for $\mathcal{D}$ in $B \cup R$.*

Following Definition 5.1 we say that: 'The locality gap for Maximum Independent Set of disks, Minimum Dominating Set of disks, etc. is $1 + \Theta(\lambda^{-1/2})$'.

3 Lower-Bound Construction

In this section we build a family of graphs that have the properties stated in Theorem 6. Namely, given parameters d, a large enough λ and n, we construct a bipartite graph G with vertex set (B, R) such that:

1. $|R| = n + o(n)$ as $n \to +\infty$,
2. G is λ-expanding,
3. $|B| \geq (1 + c\lambda^{-\frac{1}{d}}) \cdot |R| - o\,(|R|)$ as $n \to +\infty$, where c depends only on d,
4. any subgraph of G on m vertices has a separator of size $O(m^{1-\frac{1}{d}})$, and
5. G is a Gabriel graph when $d = 2$.

The vertices of R are called the **red vertices**, and the vertices of B the **blue vertices**. Our construction is geometric, in that vertices correspond to points in $\mathbb{R}^d$. Thus we use the terminology vertex and point interchangeably. We denote the i-th coordinate of a point $p \in \mathbb{R}^d$ by $x_i\,(p)$.

Let $L \geq 2$ and t be two positive integers whose values will be fixed later as a function of the parameters d, λ and n. Let Ξ be a $L \times \cdots \times L$ regular integer grid in $\mathbb{R}^d$ consisting of the $(L+1)^d$ points in $\{0, \ldots, L\}^d$. It has L^d **cells**, each defined by precisely 2^d vertices of Ξ. In every cell of Ξ, the vertex with the lexicographically minimum coordinates among the 2^d red vertices defining it is called the **anchor vertex** of that cell. Each vertex—apart from those with one of the d coordinate values equal to L—is the anchor vertex of exactly one cell, which is called its **top cell**. The cell with anchor vertex $(0, \ldots, 0)$ is called the **lowest cell** of Ξ.

We define a first bipartite graph $G(d, L)$ as follows. The red vertices of $G(d, L)$ consist of the $(L+1)^d$ points of Ξ. We next place a blue vertex at the centre of each of the L^d cells of Ξ—except for the lowest cell, which contains two blue vertices with coordinates $\left(\frac{1}{4}, \ldots, \frac{1}{4}, \frac{3}{4}\right)$ and $\left(\frac{3}{4}, \ldots, \frac{3}{4}, \frac{1}{4}\right)$. Thus $G(d, L)$ has precisely $L^d + 1$ blue vertices. The edges of $G(d, L)$ consist of 2^d edges from each blue vertex to the 2^d red vertices of its cell. Of the two blue vertices in the lowest cell of Ξ, one is connected to all the red vertices of the cell except for $(0, \ldots, 0, 1)$ (the vertex v that has $x_i(v) = 1$ if and only if $i = d$) and the other to all red vertices except for $(1, \ldots, 1, 0)$.

Our second and final graph $G(d, L, t) = (B, R; E)$ is defined as a $t \times \ldots \times t$ grid composed of t^d **translates** of $G(d, L)$. Each translate of $G(d, L)$ is indexed by a vector $\vec{\tau} \in \{0, \ldots, t-1\}^d$, where by $G^{\vec{\tau}}$ we denote the translate of $G(d, L)$ by $L \cdot \vec{\tau}$. The blue vertices of $G(d, L, t)$ are simply the disjoint union of the blue vertices of each $G^{\vec{\tau}}$; the red vertices are also the union of the red vertices of each $G^{\vec{\tau}}$, except that we identify duplicate red vertices shared by the boundary of two adjacent grids. See Figure 5.1 for an example for the case $d = 2$ and Figure 5.2 for the lowest cell when $d = 3$.

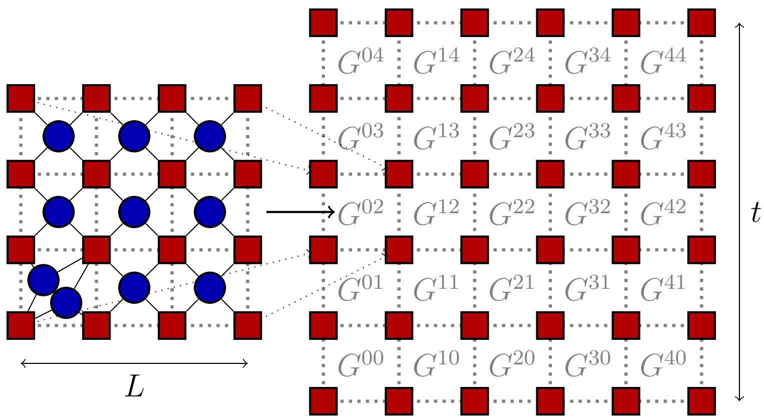


Figure 5.1: Construction of locally-expanding 'unbalanced' bipartite graphs. The graph $G(d, L)$ (shown on the left for $d = 2$ and $L = 3$) has L^d grid cells. It is the basic building block of the graph $G(d, L, t)$ (right, with $t = 5$). Square vertices are red, round vertices are blue.

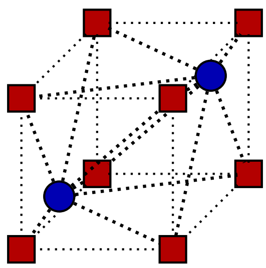


Figure 5.2: Three-dimensional lowest cell of $G(3, L)$.

An explicit, if awkward, description of $G(d, L, t) = (B, R; E)$ is:

$$R = \{0, \ldots, tL\}^d,$$

$$B = \left\{ \left(\frac{1}{2}, \ldots, \frac{1}{2}\right) + \vec{x} : \vec{x} \in \{0, \ldots, tL-1\}^d \setminus \{0, L, \ldots, (t-1)L\}^d \right\}$$
$$\bigcup \left\{ (\mu, \ldots, \mu, 1-\mu) + L \cdot \vec{x} : \vec{x} \in \{0, 1, \ldots, t-1\}^d,\ \mu \in \left\{\frac{1}{4}, \frac{3}{4}\right\} \right\},$$

$$E = \left\{ \{b, r\} : \begin{array}{c} (b, r) \in B \times R, \\ \min\limits_{i \in \{1,\ldots,d\}} |x_i(b) - x_i(r)| \leq \frac{1}{2}, \\ \max\limits_{i \in \{1,\ldots,d\}} |x_i(b) - x_i(r)| \leq 1 \end{array} \right\}.$$

The $L^d + 1$ blue vertices of $G^{\vec{\tau}}$ form the set $B_{\vec{\tau}}$. For the red vertices, note that the outer red vertices of each copy of $G(d, L)$ may be shared between up to 2^d translates. To avoid this overlap, let $R_{\vec{\tau}}$ consist only of the L^d red vertices $v \in G^{\vec{\tau}}$ such that $x_i(v) < L(\vec{\tau}_i + 1)$ for each i. In two dimensions, this amounts to peeling off the $2L + 1$ red vertices located on the top and right boundaries of $G^{\vec{\tau}}$.

Let R_b be the set of red vertices with at least one coordinate value equal to tL. We have

$$B = \bigcup_{\vec{\tau}} B_{\vec{\tau}} \quad \text{and} \quad R = R_b \cup \bigcup_{\vec{\tau}} R_{\vec{\tau}},$$

where all unions are disjoint. Observe that

$$|B| = t^d(L^d + 1) \quad \text{and} \quad |R| = (tL + 1)^d. \tag{5.4}$$

Local expansion. To prove that $G(d, L, t)$ is locally expanding we fix a subset B' of B and let $R' = N(B')$ be the set of its (red) neighbours in $G(d, L, t)$. We show that $|R'| \geq |B'|$ whenever B' is smaller than some function of L and d; later we will set L such that this function turns out to be at least λ.

A grid cell is **non-empty** if it contains a vertex of B' and otherwise **empty**. A vertex of R' that belongs to R_b or whose top cell is empty is called a **boundary vertex**.

We first sketch a proof in two dimensions based on a *charging* argument (a one-to-one mapping from B' to R'): each vertex of B' is charged to a vertex of R' such that each vertex of R' receives at most one charge, implying that $|R'| \geq |B'|$. Charge each blue vertex of B' to the anchor red vertex of its cell. For those $G^{\vec{\tau}}$ containing two blue vertices in the lowest cell, one of them remains uncharged. On the other hand, each red vertex receives one charge, except the boundary vertices which receive zero charge. Now for each $\vec{\tau}$ for which $G^{\vec{\tau}}$ contains at least two boundary red vertices charge the uncharged blue vertex in $G^{\vec{\tau}}$ (if it exists) to any one of these (at least two) boundary vertices.

There still remains an uncharged blue vertex in those $G^{\vec{\tau}}$ with less than two boundary red vertices. However, for each such $\vec{\tau}$, the number of vertices of B' in $G^{\vec{\tau}}$ must be at least $\frac{L^2}{2}$. Thus overall, there can remain at most $\frac{|B'|}{L^2/2} = \frac{2|B'|}{L^2}$ uncharged blue vertices. On the other hand, we argue that the total number of boundary red vertices is at least $c_2 \cdot \sqrt{|B'|}$, for some constant c_2. By our charging scheme, at least half of them—or $\frac{c_2}{2} \cdot \sqrt{|B'|}$—are still uncharged. Thus when $\frac{2|B'|}{L^2} \leq \frac{c_2}{2} \cdot \sqrt{|B'|}$—or equivalently $|B'| \leq c' \cdot L^4$, the number of uncharged blue vertices will be less than the number of uncharged red vertices, and we are done.

Now we present the complete proof for general d. We need two preliminary statements. Let the indicator variable $d_{\vec{\tau}}$ be 1 if both blue vertices in the lowest cell of $G^{\vec{\tau}}$ belong to B' and 0 otherwise. Also let $\delta_{\vec{\tau}}$ be the number of boundary vertices in $R_{\vec{\tau}}$. The total number of boundary vertices in R' is thus

$$\delta = |R_b \cap R'| + \sum_{\vec{\tau}} \delta_{\vec{\tau}}. \tag{5.5}$$

Lemma 5.1. *For each index $\vec{\tau}$, if $d_{\vec{\tau}} = 1$ and $\delta_{\vec{\tau}} < 2$, then $|B' \cap B_{\vec{\tau}}| \geq \frac{L^d}{2}$.*

Proof. For such an index, B' contains both blue vertices from the lowest cell of $G^{\vec{\tau}}$ so R' contains the 2^d red vertices of this cell. If $\delta_{\vec{\tau}} = 0$, that is, $R_{\vec{\tau}}$ contains no boundary vertex, then the blue vertex in each other cell of $G^{\vec{\tau}}$ is present in B', and so B' includes all of $B_{\vec{\tau}}$, which consists of $L^d + 1$ blue vertices. It remains to consider the case when $R_{\vec{\tau}}$ contains one unique boundary vertex $v_r \in R' \cap R_{\vec{\tau}}$.

Without loss of generality, assume that $\vec{\tau} = (0, \ldots, 0)$. As both blue vertices from the lowest cell of $G^{\vec{\tau}}$ belong to B', the boundary vertex v_r cannot be the lowest vertex of $G^{\vec{\tau}}$, which has coordinates $(0, \ldots, 0)$. Thus there must be some $j \in \{1, \ldots, d\}$ for which $x_j(v_r) > 0$. Consider the grid slab Ξ' consisting of all cells whose anchor vertex v has $x_j(v) = 0$. Note that Ξ' contains the lowest cell of $G^{\vec{\tau}}$, which has two vertices of B'. Thus no other cell of Ξ' can be empty, as otherwise that would imply the existence of another boundary red vertex anchoring one of the cells of Ξ'. Now take any cell c of Ξ' whose anchor vertex differs in at least one coordinate other than x_j from v_r; there are $L^{d-1} - 1$ such cells. All the L cells of $G^{\vec{\tau}}$ whose anchor vertex only differs in the j-th coordinate value from the anchor vertex of c must also be non-empty, as otherwise it would imply the existence of a boundary red vertex in one of these L cells.

Thus there are at least $L\left(L^{d-1} - 1\right)$ non-empty cells in $G^{\vec{\tau}}$, i.e. $|B' \cap B_{\vec{\tau}}| \geq L^d - L$ which is at least $\frac{L^d}{2}$ since $L \geq 2$. $\square$

Let T be the set of indices $\vec{\tau}$ with $d_{\vec{\tau}} = 1$ and $\delta_{\vec{\tau}} < 2$. As a consequence of the previous lemma, for every such $\vec{\tau} \in T$, the translate $G^{\vec{\tau}}$ contains at least $\frac{L^d}{2}$ vertices of B', and thus $|T| \leq 2\,|B'|\,L^{-d}$. Now consider the quantity $d_{\vec{\tau}} - \frac{\delta_{\vec{\tau}}}{2}$. If $\vec{\tau} \in T$, we have $d_{\vec{\tau}} = 1$ and $0 \leq \delta_{\vec{\tau}} < 2$ and so $d_{\vec{\tau}} - \frac{\delta_{\vec{\tau}}}{2}$ is at most 1. Otherwise for any $\vec{\tau} \notin T$, it is 0 or

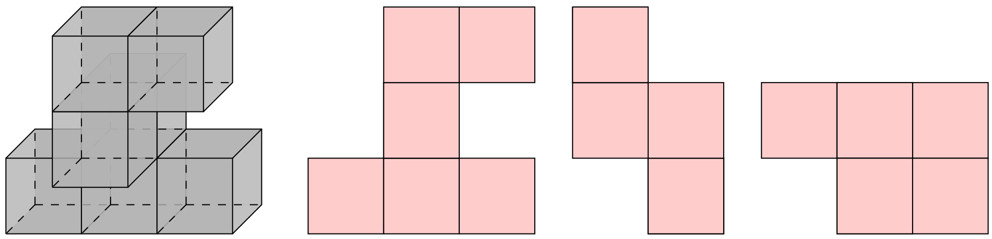


Figure 5.3: Loomis–Whitney inequality for cubes. The number of cubes in a subset of the 'cubical subdivision' of $\mathbb{R}^d$ is controlled by the number of $(d-1)$-dimensional cubes in its projections onto the $\{x_i = 0\}$ hyperplanes. Here $d = 2$ and the inequality is $7^{3-1} \leq 6 \times 4 \times 5$.

negative. Therefore

$$\sum_{\vec{\tau}} \left(d_{\vec{\tau}} - \frac{\delta_{\vec{\tau}}}{2} \right) \leq |T| \leq \frac{2\,|B'|}{L^d}. \tag{5.6}$$

An isoperimetric inequality. Consider the set S of all grid cells containing vertices of B'. As each cell contains at most two blue vertices, $|B'| \leq 2\,|S|$. In the orthogonal projection along the i-th coordinate, S is sent to a set S_i of $(d-1)$-dimensional cells. The preimage of each cell of S_i is a column of d-dimensional cells and must contain at least one boundary vertex, so the total number δ of boundary vertices is at least $|S_i|$. The combinatorial Loomis–Whitney inequality (see Figure 5.3 and Appendix A) relates d- and $(d-1)$-dimensional volumes:

$$\prod_{i=1}^{d} |S_i| \geq |S|^{d-1} \geq \left(\frac{|B'|}{2} \right)^{d-1},$$

from which it follows that

$$\delta^d \geq \left(\frac{|B'|}{2} \right)^{d-1}. \tag{5.7}$$

Now we come to the key claim, which means that the graph $G(d, L, t)$ is $(2^{1-3d}L^{d^2})$-expanding.

Lemma 5.2. *If* $2^{3d-1}\,|B'| \leq L^{d^2}$*, then* $|R'| \geq |B'|$.

Proof. For every index $\vec{\tau}$, by definition, each vertex in the set $R_{\vec{\tau}} \cap R'$ either has its top cell non-empty or is a boundary vertex of $G^{\vec{\tau}}$. The number of non-empty top cells in

$G^{\vec{\tau}}$ is $|B_{\vec{\tau}} \cap B'| - d_{\vec{\tau}}$, while the number of boundary vertices is $\delta_{\vec{\tau}}$. Thus

$$\begin{aligned}
|R'| &= |R_b \cap R'| + \sum_{\vec{\tau}} |R_{\vec{\tau}} \cap R'| = |R_b \cap R'| + \sum_{\vec{\tau}} (|B_{\vec{\tau}} \cap B'| - d_{\vec{\tau}} + \delta_{\vec{\tau}}) \\
&= |R_b \cap R'| + |B'| - \sum_{\vec{\tau}} (d_{\vec{\tau}} - \delta_{\vec{\tau}}) \\
&\geq |B'| - \sum_{\vec{\tau}} \left(d_{\vec{\tau}} - \frac{\delta_{\vec{\tau}}}{2} \right) + \frac{1}{2} \left(|R_b \cap R'| + \sum_{\vec{\tau}} \delta_{\vec{\tau}} \right) \\
&= |B'| - \sum_{\vec{\tau}} \left(d_{\vec{\tau}} - \frac{\delta_{\vec{\tau}}}{2} \right) + \frac{\delta}{2}.
\end{aligned}$$

Use the lower bounds (5.6) and (5.7) for the second and third summands:

$$|R'| \geq |B'| - \frac{2\,|B'|}{L^d} + \frac{1}{2} \left(\frac{|B'|}{2} \right)^{(d-1)/d},$$

and the result $|R'| \geq |B'|$ follows when

$$\begin{aligned}
\frac{2\,|B'|}{L^d} &\leq \frac{1}{2} \left(\frac{|B'|}{2} \right)^{(d-1)/d} \\
2^{3-1/d}\,|B'|^{1/d} &\leq L^d
\end{aligned}$$

or equivalently $2^{3d-1}\,|B'| \leq L^{d^2}$. □

Ball graph structure. A **ball graph** is the intersection graph of a family of n balls in $\mathbb{R}^d$ and is p-ply if it has no clique of size $p+1$. Such graphs have separators of size $O(p^{\frac{1}{d}} n^{1-\frac{1}{d}})$ [83].

A bounded-ply ball graph is obtained from $G(d, L, t)$ by only adding some edges: put a d-dimensional ball of radius $\frac{\sqrt{d}}{4}$ at each vertex of $G(d, L, t)$. The resulting edge set includes that of $G(d, L, t)$—they coincide when $d \leq 3$—so that $G(d, L, t)$ inherits separator properties of ball graphs. In other words, any subgraph of $G(d, L, t)$ on m vertices has a separator of size $O(m^{1-\frac{1}{d}})$ (this is property (4)).

Note that said bound on the graph's ply is a function of d only: the largest number of vertices of $B \cup R$ included in a same ball of radius $\frac{\sqrt{d}}{2}$. See e.g. [41] for estimates on such bounds.

Gabriel graph structure. For $d = 2$, the circumdisk of each blue–red edge in $G(d, L, t)$ contains no vertex but its endpoints, so $G(d, L, t)$ is a Gabriel graph and property (5) is proved.

Remark 5.2. With the understanding that a one-dimensional cell is an interval, the construction covers the case $d = 1$. The graph $G(1, L, t)$ is a path of length $2tL + 1$, seen as blue–red bipartite, with every L-th blue vertex duplicated. It has $|R| = tL + 1$ and $|B| = t(L + 1)$ and is $(L + 2)$-expanding.

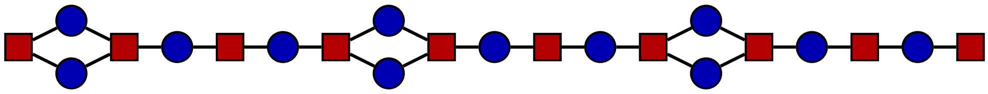


Figure 5.4: The graph $G(1, 3, 3)$.

Setting parameters and concluding the proof. Given d, λ and n, choose

$$L = \max\left\{2, \left\lceil \left(2^{3d-1}\lambda\right)^{1/d^2} \right\rceil\right\} \text{ and } t = \lceil n^{1/d} L^{-1} \rceil.$$

Note that L does not depend on n and L^d is $\Theta(\lambda^{\frac{1}{d}})$ when $\lambda \to +\infty$. Using (5.4), we obtain (1) and (3):

$$\begin{aligned} |R| &= (tL + 1)^d = n + o(n), \\ \frac{|B|}{|R|} &= \frac{t^d \left(L^d + 1\right)}{(tL + 1)^d} = 1 + \frac{1}{L^d} + o(1) \text{ as } n \to +\infty \\ &\geq 1 + c_d \lambda^{-\frac{1}{d}} + o_n(1) \text{ for } \lambda \geq \lambda_d \end{aligned}$$

where the positive constants c_d and λ_d depend only on d. Since $2^{1-3d} L^{d^2} \geq \lambda$, it follows from Lemma 5.2 that (2) holds: G is at least λ-expanding. This concludes the proof of Theorem 6.

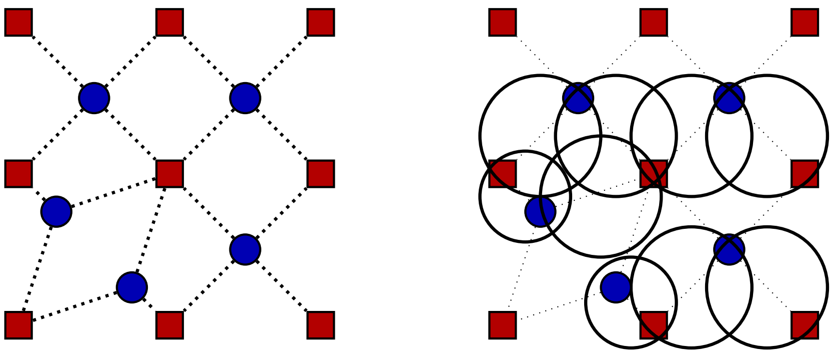
4 Consequences

4.1 Geometric Problems in the Plane

We construct arbitrarily large instances of our five optimisation problems for which some λ-locally optimal solution is $1 + \Omega(\lambda^{-1/2})$ times worse than the optimal solution. Since our instances consist of proper disks rather than just families of pseudodisks, the bound applies also to the restrictions of these problems to disk families.

For $d = 2$ and any given $\lambda \geq \lambda_d$ and n, let $G = (B, R; E)$ be the planar and λ-expanding graph $(B_n, R_n; E_n)$ described in Theorem 6 and built in Section 3. Our instances are based on G: its vertex sets are associated with feasible solutions of the problems. It then suffices to check that the solution associated with B_n (for minimisation problems) or R_n (for maximisation problems) is locally optimal.

4.1.1 Hitting Set for Pseudodisks



(a) Detail of the graph G used in 'bad' instances.

(b) Hitting set (drawing only a few disks for readability).

Figure 5.5: Building a 'tight' instance for the Hitting Set problem.

Theorem M ([87]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for Minimum Hitting Set for pseudodisks.*

Corollary 5.1. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there is a set $\mathcal{D}$ of at least n disks and two disjoint sets B and R of at least n points in $\mathbb{R}^2$ each such that both B and R are hitting sets for $\mathcal{D}$, $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and B is a λ-locally optimal solution to the Hitting Set problem for $\mathcal{D}$ with $\mathcal{P} = B \cup R$.*

Proof. Recall that the circumdisk of each edge of G contains only its two endpoints. The input consists of all such disks, with $\mathcal{P} = B \cup R$, so that the hitting sets are exactly the vertex covers of G. By construction both B and R are feasible solutions.

On this instance, a λ-local improvement for B would remove a set B' of blue vertices with $|B'| \leq \lambda$. To preserve the hitting set property, it would then need to add to the solution the red endpoints of all edges with their blue endpoint in B', i.e. the set $N(B')$. Because the graph is λ-expanding, there are at least $|B'|$ such red neighbours: B is λ-locally optimal. □

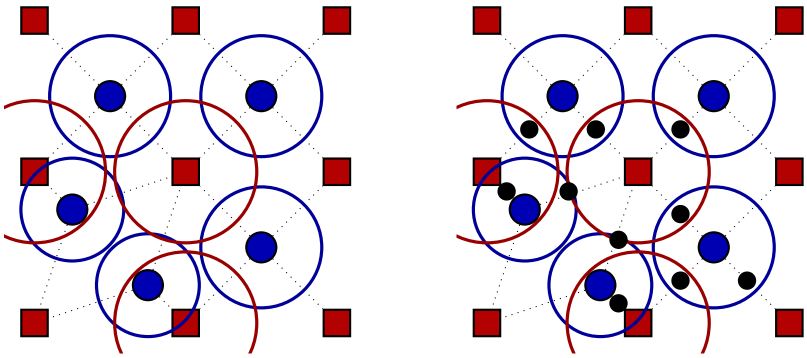
4.1.2 Independent Set of Pseudodisks

Theorem N ([33]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for Maximum Independent Set for pseudodisks.*

Corollary 5.2. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are* *two independent sets B and R of at least n disks in $\mathbb{R}^2$ such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the* Independent Set *problem in $B \cup R$.*

Proof. Realise the graph G as an intersection graph of red and blue disks. As it is planar, the disks could even be taken interior-disjoint by the Koebe–Andreev–Thurston theorem (see Appendix A). The independent sets of disks correspond to the independent sets of G. Since G is bipartite both the blue and red families of disks form independent sets, and the red solution is $(\lambda - 1)$-locally optimal—in maximisation terms: a $(\lambda - 1)$-local improvement for the red solution would remove a set R' of up to $\lambda - 1$ red disks and replace them with a set B' of blue disks such that $N(B') \subseteq R'$ (to preserve independence) and $|B'| > |R'|$. If there exists a subset $B'' \subseteq B'$ of size $|R'| + 1$, which is at most λ, then since G is λ-expanding such a set has $|B''| \leq |N(B'')| \leq |R'|$, a contradiction. Thus R is a $(\lambda - 1)$-locally optimal solution. □

4.1.3 Set Cover for Disks



(a) Independent set. (b) Set cover.

Figure 5.6: 'Tight' instances for Independent set and Set Cover with disks.

Theorem O ([15, 31])**.** *Local search with radius $O(\epsilon^{-2})$ is a $(1 + \epsilon)$-approximation algorithm for* Minimum Set Cover *for disks.*

Corollary 5.3. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two independent sets B and R of at least n disks in $\mathbb{R}^2$ and a set $\mathcal{P}$ of $\Theta(|R|)$ points in $\mathbb{R}^2$ such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the* Set Cover *problem for $\mathcal{P}$ in $B \cup R$.*

Proof. As in the proof of Corollary 5.2, realise G as an intersection graph of blue and red disks. Take for $\mathcal{P}$ one point from each blue–red intersection. The set covers for this instance are exactly the vertex covers of G. □

4.1.4 Dominating Set of Pseudodisks

Theorem P ([52, 53]). *Local search with radius $O(\epsilon^{-2})$ is a $(1 + \epsilon)$-approximation algorithm for Minimum Dominating Set for pseudodisks.*

Corollary 5.4. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there is a set $\mathcal{D}$ of disks in $\mathbb{R}^2$ and two dominating sets B and R of $\mathcal{D}$ of at least n disks each such that $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and B is a λ-locally optimal solution to the Dominating Set problem for $\mathcal{D}$.*

Proof. The instance that was proposed for Set Cover (Figure 5.6b) becomes an instance of Dominating Set when the points of $\mathcal{P}$ are seen as zero-radius disks, i.e. take $\mathcal{D} = \mathcal{P} \cup B \cup R$. A feasible solution that involves some of the zero-radius disks of $\mathcal{P}$ can be transformed into a solution of at most the same cardinality whose support is entirely blue and red since the disks of $\mathcal{P}$ are fully included in the other disks. Thus it suffices to examine the efficiency of local search on blue–red solutions. The blue–red dominating sets of this instance are exactly the covers of points by blue and red disks. □

4.1.5 Unit-capacity Packing Problems

Theorem Q ([46]). *Local search with radius $O(\epsilon^{-2})$ is a $(1+\epsilon)$-approximation algorithm for the Unit-Capacity Point-Packing problem for disks.*

Corollary 5.5. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two sets B and R of at least n points in $\mathbb{R}^2$ and a set $\mathcal{D}$ of $\Theta(|R|)$ disks in $\mathbb{R}^2$ such that every disk of $\mathcal{D}$ contains one point from B and one point from R, $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally optimal solution to the Unit-Capacity Point-Packing problem for $\mathcal{D}$ in $B \cup R$.*

Proof. Take for $\mathcal{D}$ the set of all disks associated with the edges, as in the Hitting Set instance (see Figure 5.5b). Since every such disk contains only two points of $\mathcal{P}$, the 'unit-capacity point-packings' of this instance are exactly the independent sets of G. The result then follows from the analysis in Corollary 5.2. □

Recall that Unit-Capacity Disk-Packing is the dual problem [46].

Corollary 5.6. *There is a positive constant C and a positive integer λ_0 such that for every integer $\lambda \geq \lambda_0$ there is a positive integer n_λ such that for every integer $n \geq n_\lambda$ there are two sets B and R of at least n disks each and a set $\mathcal{P}$ of $\Theta(|R|)$ points such that every point of $\mathcal{P}$ is contained in one disk from B and one disk from R, $|B| \geq (1 + C\lambda^{-\frac{1}{2}})\,|R|$ and R is a λ-locally maximal solution to the Unit-Capacity Disk-Packing problem for $B \cup R$ in $\mathcal{P}$.*

4.2 Other Problems with Hereditary Separators

The paper by Har-Peled and Quanrud [57] is to the best of our knowledge the most extensive study of geometric local search in non-planar settings. The authors study graphs with polynomial expansion, which have strongly sub-linear separators, and in particular intersection graphs of low-density families of objects, where a family of objects in $\mathbb{R}^d$ has **density** ρ if for any $r \geq 0$ any ball of diameter r intersects at most ρ objects of diameter larger than r and has **depth** D if no point of $\mathbb{R}^d$ is contained in $D+1$ objects. A survey on expansion and sparsity is the book by Nešetřil and Ossona de Mendez [89].

We are still able to give some lower bounds on the local search radii that achieve PTASs. Fix the positive integers $d, \lambda \geq \lambda_d$ and n and let G be the λ-expanding graph on vertex sets B_n and R_n that has $|B_n|, |R_n| = \Theta(n)$ and achieves $|B_n| \geq (1 + c\lambda^{-1/d} - o(1))\,|R_n|$, as built in Section 3. Recall that G and its subgraphs have the separator property with $s = 1/d$.

By combining Theorem 3.2.1 and Lemma 2.2.9 from [57], we obtain the following.

Theorem R. *On graphs with hereditary separators of size $O(n^{1-s})$, local search with radius $O(\epsilon^{-s})$ is a $(1+\epsilon)$-approximation algorithm for Maximum Independent Set.*

Corollary 5.7. *For every positive integers d and λ, there are arbitrarily large bipartite graphs on vertex sets (B, R) with hereditary separators of size $O(n^{1-1/d})$ such that $|B| \geq 1 + \Omega(\lambda^{-1/d})\,|R|$ and R is a λ-locally maximal independent set.*

Proof. Since the graph G is bipartite, both B_n and R_n are independent sets, and by the same analysis as in the proof of Corollary 5.2 the feasible solution R_n is $(\lambda-1)$-locally optimal. □

Theorem S ([57])**.** *On graphs with hereditary separators of size $O(n^{1-s})$, local search with radius $O(\epsilon^{-O(1)})$ is a $(1+\epsilon)$-approximation algorithm for Minimum Vertex Cover.*

Corollary 5.8. *For every positive integers d and λ, there are arbitrarily large bipartite graphs on vertex sets (B, R) with hereditary separators of size $O(n^{1-1/d})$ such that $|B| \geq 1 + \Omega(\lambda^{-1/d})\,|R|$ and B is a λ-locally minimal vertex cover.*

Proof. In G both B_n and R_n are vertex covers. Since G is λ-expanding, B_n is λ-locally optimal. □

4.3 Matchings and Local Versions of Hall's Theorem

With our terminology, Hall's theorem is as follows (see also Appendix A).

Theorem T (Hall's marriage theorem)**.** *Any bipartite graph on vertex sets (B, R) that is $|B|$-expanding has a matching with $|B|$ edges.*

Restricting the condition to λ-expansion for some fixed λ breaks Hall's theorem: for example, the matchings of $K_{|B|,\lambda}$ have at most λ edges. However it was observed by Antunes, Mathieu and Mustafa [8] that a strengthening of Hall's theorem holds for planar graphs.

Theorem U. *There is an absolute constant $c > 0$ such that, for every given integer $\lambda \geq 3$, any bipartite* planar *graph on vertex sets (B, R) that is λ-expanding has a matching with at least $(1 - c\lambda^{-\frac{1}{2}})\,|B|$ edges.*

Now it follows from our constructions that this is tight.

Corollary 5.9. *There are absolute constants $c_0, \lambda_0 > 0$ such that, for every given integer $\lambda \geq \lambda_0$, some bipartite, λ-expanding planar graph on vertex sets (B, R) does not have matchings with more than $(1 - c_0\lambda^{-\frac{1}{2}})\,|B|$ edges.*

5 Remarks

We emphasise that our results apply to standard, non-specialised local-search techniques. Although the approximation quality of a previously successful one-size-fits-all approach cannot be improved, custom algorithms tailored for specific problems can bypass this bound, especially when the exchange graphs are extremely sparse.

5.1 Local Search and Terrain Guarding

For example we do not yet know whether our constructions can be transformed into a local-search-defeating instance for Terrain Guarding, a question that can be formulated as follows.

Question. Are the exchange graphs of Gibson et al. [51] for Terrain Guarding sparser than other planar graphs? What is the minimum size of their separators?

5.2 Local Search with Small Radius

When used with small radii rather than $\lambda = \Theta(1/\epsilon^2)$, geometric local search yields constant-factor approximations instead of PTASs. For planar exchange graphs and such small radii, the combinatorial analysis based on Theorem L can be specialised: for $k \in \mathbb{N}^*$, let c_k be the smallest c such that all planar k-expanding graphs on vertex sets B and R have $|B| \leq c\,|R| + O(1)$.

By combining Theorem L with our lower bound (Theorem 6), the asymptotic behaviour of c_k is

$$c_k = 1 + \Theta\left(\frac{1}{\sqrt{k}}\right).$$

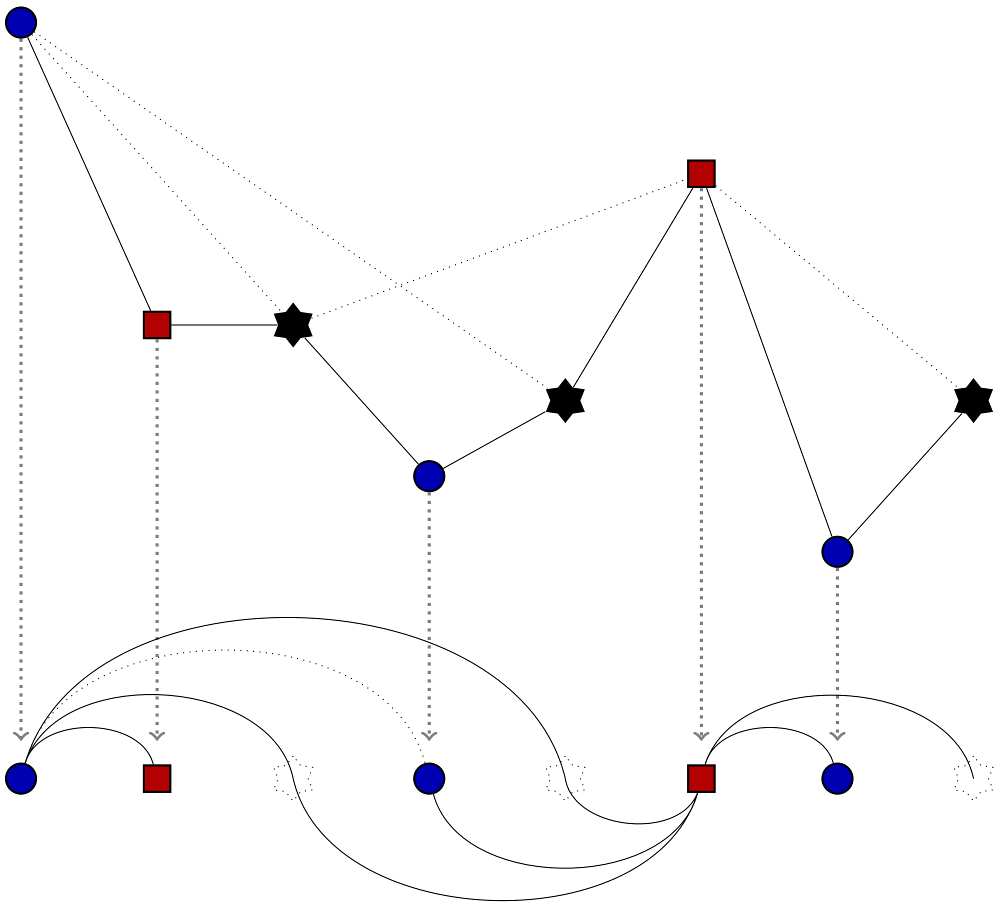


Figure 5.7: Exchange graph construction for TERRAIN GUARDING, as it appears in Gibson et al. [51]. Edges are drawn between *guard* vertices based on the targets (stars in this picture) that they jointly see.

It is known that $c_1 = c_2 = +\infty$ (consider the sequence of 2-expanding planar bicliques $K_{n,2}$ as $n \to \infty$), that $c_3 = 8$ [25] and that $c_4 = 4$ [8]. Thus local search with radius 2 and 3 achieves approximation ratios of respectively 8 and 4 on the problems in this chapter.

Conjecture. We believe that the value of c_5 is 3.

At any rate we are able to construct graphs showing that $c_5 \geq 3$: see Figure 5.8.

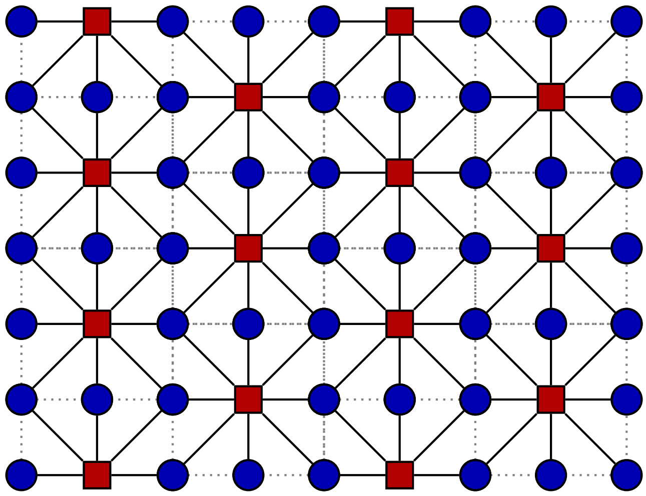


Figure 5.8: Lower bound construction for $c_5 \geq 3$: as the plane is tiled in this fashion, the infinite blue–red bipartite graph is 5-expanding and planar, and the asymptotic blue-to-red ratio is 3.

Appendices

Appendix A

Some Classical Theorems

Here we gather some well-known theorems from discrete mathematics that are referred to in the body of this dissertation.

1 Circle Packing Theorem of Koebe, Andreev and Thurston

A circle packing is a family of disks of $\mathbb{R}^2$ whose interiors are pairwise disjoint. To every such packing one associates a *coin graph* that has a vertex for each disk and whose edges correspond to intersecting boundaries. The circle packing theorem asserts that the planar graphs are exactly the coin graphs [67].

2 Cubical Loomis–Whitney Inequality

This lemma is the critical step in the proof of the more general Loomis–Whitney inequality [73]. Consider $\mathbb{R}^d$ tiled with all unit cubes of the form $x + [0, 1]^d$ where x is a tuple of d integers. Let S be a finite subset of such cubes and for each index i let S_i be the set of $(d-1)$-dimensional cubes in the orthogonal projection along the i-th coordinate. Then

$$\prod_{i=1}^{d} |S_i| \geq |S|^{d-1} .$$

3 Hall's Marriage Theorem

This theorem was first formulated by Philip Hall [55] in terms of the combinatorics of set systems, but can also be seen as a result in graph theory. In a bipartite graph on finite vertex sets L and R, the following properties are equivalent.

1. Every subset L' of L has at least as many neighbours in R as its own cardinality, i.e. at least $|L'|$.

2. There is a subset of edges, no two of them sharing a vertex (a *matching*), that covers all vertices of L.

In this dissertation we discuss some results (Theorem U) that extend this theorem.

4 Graph Separator Theorems

A separator in a graph of order n is a subset of vertices whose complement does not induce a subgraph with connected components containing more than $2/3n$ vertices.

Planar graphs admit separators of cardinality only $O(\sqrt{n})$. This was first shown by Lipton and Tarjan [72]. Many other proofs have been found. For example Har-Peled [56] observed that it follows from the circle packing theorem.

A minor of a graph is a smaller graph obtained by a sequence of the following operations: contracting an edge so that its two vertices become identified, deleting an edge, and deleting a vertex that is not part of any edge. A minor can be embedded in the same space as the original graph. For this reason planar graphs form one of the graph families closed under taking minors. From the point of view of separators, such minor-closed families all behave like planar graphs in the sense that each of them except for the family of all graphs has a separator theorem in $O(\sqrt{n})$. Precisely, all graphs of order up to $h \in \mathbb{N}$ are minors of the complete graph K_h, so a minor-closed family that does not include all graphs must exclude K_h for some h. Then a theorem of Alon, Seymour, and Thomas [7] states that the graphs in this family have separators in $O(h^{3/2}\sqrt{n})$.

5 Turán's Theorem

A t-clique in a graph ($t \in \mathbb{N}$) is a subset of t vertices that span all possible edges between them, i.e. $\binom{t}{2}$ edges. A theorem of Pál Turán [96] states that a graph of order n that has no $(t+1)$-clique must possess at most

$$\frac{t-1}{t} \cdot \frac{n^2}{2}$$

edges. By considering the complement graph one obtains the equivalent result that a graph of order n that does not have $t+1$ independent vertices must have at least

$$\binom{n}{2} - \frac{t-1}{t} \cdot \frac{n^2}{2} = \frac{n(n-t)}{2t}$$

edges, which is the statement that we use in Chapter 2.

Appendix B

Open Questions and Remaining Problems

Here are a few problems encountered in this dissertation that remain open.

1 On Mnets

- The proof of the Mnet theorem (Theorem 3) uses algebraic tools, which is why the theorem applies to semi-algebraic set systems only. Can this be avoided? In other words, come up with an alternative to polynomial partitioning.
- The usual bounds on unweighted nets are easy consequences of said theorem (see Corollary 3.1 in Chapter 3). Can the same be said of weighted nets?
- Give a sharp lower bound on the size of Mnets in terms of shallow-cell complexity, i.e. improve the exponent in Theorem 5.
- The recent employment of Macbeath regions in efficient and elegant data structures for convex body approximation and approximate polytope membership [2, 12] has renewed interest in those structures. Can Mnets—a discrete analogue—find similar uses?
- The result of Brönnimann, Chazelle, and Pach [23] on the small family of Macbeath regions that approximate a given convex body holds in a Lebesgue setting. Propose extensions to other measures, absolutely continuous or singular.

2 On Local Search

- The lower bound on the efficiency of local search (Theorem 6) is proved for d a positive integer, but there is no reason why it should not hold for any real d in

$[1, +\infty)$. Perhaps some kind of interpolation could be defined on grid graphs of distinct dimensions.

- Describe the exchange graphs for TERRAIN GUARDING (see page 96). Can our lower bound construction be adapted to this problem?

- Prove the conjecture on page 97: the correct value of c_5 (the left-to-right ratio for *planar* 5-expanding graphs) is 3. More generally, determine c_k for all k.

- In practice, local search algorithms may be initialised from a feasible solution chosen at random. The choice of local improvements could also be random whenever several candidates are available. What can still be said about the efficiency of local search? Are there instances were not only some but *most* possible search paths end at a solution whose value is far from the optimum?

Appendix C

Tentative English–French Lexicon

> 5.6—Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.
>
> WITTGENSTEIN

Discussing my work in French has proved difficult. With the amount of French-speaking literature dwindling over time, many concepts do not have a well-established name. Here are a few proposals.

ground set	univers
hitting set	transversal(e), ensemble intersectant
locality gap	*écart de localité*
Mnet	*M-réseau, réseau de Macbeath*
net (set system)	réseau
(ϵ-)packing (semimetric space)	partie (ϵ-)discernable [1]
packing (set system)	*famille discernable*
polynomial partitioning	*partition par polynômes*
range of a set system	(hyper)arête d'un hypergraphe [2]
shallow-cell complexity	*complexité en cellules peu profondes*
shallow set system	hypergraphe de rang borné [2]
shatter dimension / exponent	densité réelle [1]
shatter function	coefficient de pulvérisation
shattered set	ensemble pulvérisé
VC dimension	dimension VC, densité entière [1]

[1] Patrick ASSOUAD. 'Densité et dimension'. French. In: *Annales de l'institut Fourier* 33.3 (1983), pp. 233–282.

[2] Claude BERGE. *Hypergraphes: combinatoire des ensembles finis*. French. Gauthier-Villars, 1987.

.

References

[1] Scoot Aaronson. *The Scientific Case for $\boldsymbol{P} \neq \boldsymbol{NP}$*. Mar. 2014. URL: https://www.scottaaronson.com/blog/?p=1720 (visited on 05/09/2018) (cit. on p. 19).

[2] Ahmed Abdelkader and David M. Mount. 'Economical Delone Sets for Approximating Convex Bodies'. In: *16th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2018)*. Vol. 101. 2018, 4:1–4:12 (cit. on p. 102).

[3] Pankaj K. Agarwal and Nabil H. Mustafa. 'Independent set of intersection graphs of convex objects in 2D'. In: *Computational Geometry* 34.2 (2006), pp. 83–95 (cit. on pp. 9, 30, 81).

[4] Pankaj K. Agarwal, János Pach, and Micha Sharir. 'State of the Union (of Geometric Objects): A Review'. In: *Computational Geometry: Twenty Years Later.* American Mathematical Society, 2008, pp. 9–48 (cit. on pp. 23, 42).

[5] Pankaj K. Agarwal and Jiangwei Pan. 'Near-Linear Algorithms for Geometric Hitting Sets and Set Covers'. In: *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*. 2014, 271:271–271:279 (cit. on p. 28).

[6] Noga Alon. 'A Non-linear Lower Bound for Planar Epsilon-nets'. In: *Discrete & Computational Geometry* 47.2 (Mar. 2012), pp. 235–244 (cit. on p. 22).

[7] Noga Alon, Paul Seymour, and Robin Thomas. 'A separator theorem for non-planar graphs'. In: *Journal of the American Mathematical Society* 3.4 (Oct. 1990), pp. 801–808 (cit. on pp. 82, 101).

[8] Daniel Antunes, Claire Mathieu, and Nabil H. Mustafa. 'Combinatorics of Local Search: An Optimal 4-Local Hall's Theorem for Planar Graphs'. In: *25th Annual European Symposium on Algorithms (ESA 2017)*. Vol. 87. 2017, 8:1–8:13 (cit. on pp. 33, 83, 96, 97).

[9] Boris Aronov, Mark de Berg, Esther Ezra, and Micha Sharir. 'Improved Bounds for the Union of Locally Fat Objects in the Plane'. In: *SIAM Journal on Computing* 43.2 (2014), pp. 543–572 (cit. on p. 67).

[10] Boris Aronov, Esther Ezra, and Micha Sharir. 'Small-Size ϵ-Nets for Axis-Parallel Rectangles and Boxes'. In: *SIAM Journal on Computing* 39.7 (2010), pp. 3248–3282 (cit. on p. 67).

[11] Sunil Arya, Guilherme D. da Fonseca, and David M. Mount. 'On the Combinatorial Complexity of Approximating Polytopes'. In: *Discrete & Computational Geometry* 58.4 (Dec. 2017), pp. 849–870 (cit. on pp. 24, 68).

[12] Sunil Arya, Guilherme D. da Fonseca, and David M. Mount. 'Optimal Approximate Polytope Membership'. In: *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 270–288 (cit. on p. 102).

[13] Sunil Arya, Guilherme D. da Fonseca, and David M. Mount. 'Optimal Area-Sensitive Bounds for Polytope Approximation'. In: *Proceedings of the 28th Annual ACM Symposium on Computational Geometry, SoCG*. 2012, pp. 363–372 (cit. on p. 68).

[14] Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. 'Local Search Heuristics for k-Median and Facility Location Problems'. In: *SIAM Journal on Computing* 33.3 (2004), pp. 544–562 (cit. on p. 80).

[15] Rom Aschner, Matthew J. Katz, Gila Morgenstern, and Yelena Yuditsky. 'Approximation Schemes for Covering and Packing'. In: *WALCOM: Algorithms and Computation: 7th International Workshop. Proceedings*. 2013, pp. 89–100 (cit. on pp. 30, 34, 81, 82, 84, 93).

[16] Imre Bárány. 'Random polytopes, convex bodies, and approximation'. In: *Stochastic Geometry*. Springer, 2007, pp. 77–118 (cit. on p. 59).

[17] Imre Bárány and David G. Larman. 'Convex bodies, economic cap coverings, random polytopes'. In: *Mathematika* 35 (1988), pp. 274–291 (cit. on p. 59).

[18] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. Springer-Verlag, 2003 (cit. on p. 60).

[19] Cristina Bazgan. 'Schémas d'approximation et complexité paramétrée'. French. Unobtainable online. Mémoire de DEA. Université Paris Sud, 1995 (cit. on pp. 29, 73).

[20] Shai Ben-David and Michael Lindenbaum. 'Localization vs. Identification of Semi-Algebraic Sets'. In: *Machine Learning* 32.3 (Sept. 1998), pp. 207–224 (cit. on p. 43).

[21] Claude Berge. *Hypergraphes: combinatoire des ensembles finis*. French. Gauthier-Villars, 1987.

[22] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. 'Classifying Learnable Geometric Concepts with the Vapnik-Chervonenkis Dimension'. In: *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*. 1986, pp. 273–282 (cit. on p. 21).

[23] Hervé Brönnimann, Bernard Chazelle, and János Pach. 'How hard is half-space range searching?' In: *Discrete & Computational Geometry* 10.2 (Aug. 1993), pp. 143–155 (cit. on pp. 8, 24, 58, 59, 102).

[24] Hervé Brönnimann and Michael T. Goodrich. 'Almost optimal set covers in finite VC-dimension'. In: *Discrete & Computational Geometry* 14.4 (Dec. 1995), pp. 463–479 (cit. on pp. 9, 28).

[25] Norbert Bus, Shashwat Garg, Nabil H. Mustafa, and Saurabh Ray. 'Limits of Local Search: Quality and Efficiency'. In: *Discrete & Computational Geometry* 57.3 (Apr. 2017), pp. 607–624 (cit. on pp. 33, 83, 97).

[26] Norbert Bus, Shashwat Garg, Nabil H. Mustafa, and Saurabh Ray. 'Tighter estimates for ϵ-nets for disks'. In: *Computational Geometry* 53 (2016), pp. 27–35 (cit. on p. 28).

[27] Sarit Buzaglo, Rom Pinchasi, and Günter Rote. 'Topological Hypergraphs'. In: *Thirty Essays on Geometric Graph Theory*. Springer, 2013 (cit. on p. 67).

[28] Sergio Cabello and David Gajser. 'Simple PTAS's for families of graphs excluding a minor'. In: *Discrete Applied Mathematics* 189 (2015), pp. 41–48 (cit. on pp. 30, 34, 82).

[29] Marco Cesati and Luca Trevisan. 'On the efficiency of polynomial time approximation schemes'. In: *Information Processing Letters* 64.4 (1997), pp. 165–171 (cit. on p. 73).

[30] Timothy M. Chan. 'Improved Deterministic Algorithms for Linear Programming in Low Dimensions'. In: *Proceedings of the Twenty-seventh Annual ACM-SIAM Symposium on Discrete Algorithms*. 2016, pp. 1213–1219 (cit. on p. 20).

[31] Timothy M. Chan and Elyot Grant. 'Exact algorithms and APX-hardness results for geometric packing and covering problems'. In: *Computational Geometry* 47.2, Part A (2014). Special Issue: 23rd Canadian Conference on Computational Geometry (CCCG11), pp. 112–124 (cit. on pp. 10, 30, 81, 84, 93).

[32] Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. 'Weighted Capacitated, Priority, and Geometric Set Cover via Improved Quasi-Uniform Sampling'. In: *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms, SODA*. 2012, pp. 1576–1585 (cit. on pp. 5, 7, 23, 40, 45, 58, 69).

[33] Timothy M. Chan and Sariel Har-Peled. 'Approximation Algorithms for Maximum Independent Set of Pseudo-Disks'. In: *Discrete & Computational Geometry* 48.2 (Sept. 2012), pp. 373–392 (cit. on pp. 9, 10, 30, 81, 82, 84, 92).

[34] Bernard Chazelle. 'A Note on Haussler's Packing Lemma'. See Section 5.3 from *Geometric Discrepancy: An Illustrated Guide* by J. Matoušek. 1992 (cit. on pp. 49, 50).

[35] Bernard Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, 2000 (cit. on p. 45).

[36] Brent N. Clark, Charles J. Colbourn, and David S. Johnson. 'Unit disk graphs'. In: *Discrete Mathematics* 86.1 (1990), pp. 165–177 (cit. on p. 81).

[37] Kenneth L. Clarkson. 'A Randomized Algorithm for Closest-Point Queries'. In: *SIAM Journal on Computing* 17.4 (1988), pp. 830–847 (cit. on p. 45).

[38] Kenneth L. Clarkson. 'New Applications of Random Sampling in Computational Geometry'. In: *Discrete & Computational Geometry* 2 (1987), pp. 195–222 (cit. on p. 45).

[39] Kenneth L. Clarkson and Kasturi Varadarajan. 'Improved Approximation Algorithms for Geometric Set Cover'. In: *Discrete & Computational Geometry* 37.1 (Jan. 2007), pp. 43–58 (cit. on p. 23).

[40] *Complexity Zoo*. University of Waterloo. URL: `https://complexityzoo.uwaterloo.ca/Complexity_Zoo` (visited on 05/09/2018) (cit. on p. 71).

[41] John Conway and Neil J. A. Sloane. *Sphere Packings, Lattices and Groups*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag New York, 1999 (cit. on p. 90).

[42] Irit Dinur and David Steurer. 'Analytical Approach to Parallel Repetition'. In: *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*. 2014, pp. 624–633 (cit. on p. 28).

[43] Rod Downey. 'A Parameterized Complexity Tutorial'. In: *Language and Automata Theory and Applications, LATA*. 2012, pp. 38–56 (cit. on p. 71).

[44] Kunal Dutta, Esther Ezra, and Arijit Ghosh. 'Two Proofs for Shallow Packings'. In: *Discrete & Computational Geometry* 56.4 (2016), pp. 910–939 (cit. on pp. 26, 40, 41, 47, 49–51).

[45] Kunal Dutta, Arijit Ghosh, Bruno Jartoux, and Nabil H. Mustafa. 'Shallow Packings, Semialgebraic Set Systems, Macbeath Regions, and Polynomial Partitioning'. In: *33rd International Symposium on Computational Geometry (SoCG 2017)*. Vol. 77. 2017, 38:1–38:15.

[46] Alina Ene, Sariel Har-Peled, and Benjamin Raichel. 'Geometric packing under non-uniform constraints'. In: *Proceedings of the Twenty-eighth Annual Symposium on Computational Geometry*. 2012, pp. 11–20 (cit. on pp. 30, 76, 77, 81, 84, 94).

[47] Esther Ezra. 'A Size-Sensitive Discrepancy Bound for Set Systems of Bounded Primal Shatter Dimension'. In: *SIAM Journal on Computing* 45.1 (2016), pp. 84–101 (cit. on pp. 40, 41, 47, 49, 50).

[48] Esther Ezra, Boris Aronov, and Micha Sharir. 'Improved Bound for the Union of Fat Triangles'. In: *Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms, SODA*. 2011, pp. 1778–1785 (cit. on p. 67).

[49] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. 'A Semi-Algebraic Version of Zarankiewicz's Problem'. In: *Journal of the European Mathematical Society* 19.6 (2017), pp. 1785–1810 (cit. on pp. 47, 50, 51, 54).

[50] Greg N. Frederickson. 'Fast Algorithms for Shortest Paths in Planar Graphs, with Applications'. In: *SIAM J. Comput.* 16.6 (Dec. 1987), pp. 1004–1022 (cit. on p. 32).

[51] Matt Gibson, Gaurav Kanade, Erik Krohn, and Kasturi R. Varadarajan. 'An Approximation Scheme for Terrain Guarding'. In: *12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009*. 2009, pp. 140–148 (cit. on pp. 96, 97).

[52] Matt Gibson and Imran A. Pirwani. 'Algorithms for Dominating Set in Disk Graphs: Breaking the $\log n$ Barrier'. In: *Proceedings of the 18th Annual European Symposium on Algorithms (ESA)*. 2010, pp. 243–254 (cit. on pp. 9, 10, 30, 32, 81, 84, 94).

[53] Sathish Govindarajan, Rajiv Raman, Saurabh Ray, and Aniket Basu Roy. 'Packing and Covering with Non-piercing regions'. In: *Proceedings of the 22nd Annual European Symposium on Algorithms (ESA)*. Vol. 57. 2016, 47:1–47:17 (cit. on pp. 30, 81, 84, 94).

[54] Larry Guth and Nets H. Katz. 'On the Erdős distinct distances problem in the plane'. In: *Annals of Mathematics* 181.1 (2015), pp. 155–190 (cit. on pp. 8, 25, 58, 60, 63).

[55] Philip Hall. 'On Representatives of Subsets'. In: *Journal of the London Mathematical Society* s1-10.1 (1935), pp. 26–30 (cit. on p. 100).

[56] Sariel Har-Peled. 'A Simple Proof of the Existence of a Planar Separator'. In: *CoRR* abs/1105.0103 (2011) (cit. on p. 101).

[57] Sariel Har-Peled and Kent Quanrud. 'Approximation Algorithms for Polynomial-Expansion and Low-Density Graphs'. In: *Proceedings of the 23rd Annual European Symposium on Algorithms (ESA)*. 2015, pp. 717–728 (cit. on pp. 30, 34, 82, 95).

[58] Sariel Har-Peled and Kent Quanrud. 'Notes on Approximation Algorithms for Polynomial-Expansion and Low-Density Graphs'. In: *CoRR* abs/1603.03098 (2016) (cit. on p. 82).

[59] David Haussler. 'Sphere Packing Numbers for Subsets of the Boolean n-Cube with Bounded Vapnik-Chervonenkis Dimension'. In: *Journal of Combinatorial Theory, Series A* 69.2 (1995), pp. 217–232 (cit. on pp. 8, 26, 47–51).

[60] David HAUSSLER and Emo WELZL. 'Epsilon-nets and simplex range queries'. In: *Discrete & Computational Geometry* 2 (1987), pp. 127–151 (cit. on pp. 7, 21, 22, 45, 68).

[61] Marc HELLMUTH, Lydia OSTERMEIER, and Peter F. STADLER. 'A Survey on Hypergraph Products'. In: *Mathematics in Computer Science* 6.1 (Mar. 2012), pp. 1–32 (cit. on p. 46).

[62] Dorit S. HOCHBAUM and Wolfgang MAASS. 'Fast approximation algorithms for a nonconvex covering problem'. In: *Journal of Algorithms* 8.3 (1987), pp. 305–323 (cit. on pp. 28, 81).

[63] Thomas JANSEN. 'Introduction to the theory of complexity and approximation algorithms'. In: *Lectures on Proof Verification and Approximation Algorithms*. Springer Berlin Heidelberg, 1998, pp. 5–28 (cit. on p. 71).

[64] Bruno JARTOUX and Nabil H. MUSTAFA. 'Optimality of Geometric Local Search'. In: *34th International Symposium on Computational Geometry (SoCG 2018)*. 2018. Forthcoming.

[65] Richard M. KARP. 'Reducibility among Combinatorial Problems'. In: *Complexity of Computer Computations*. Springer US, 1972, pp. 85–103 (cit. on p. 28).

[66] Klara KEDEM, Ron LIVNE, János PACH, and Micha SHARIR. 'On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles'. In: *Discrete & Computational Geometry* 1.1 (Mar. 1986), pp. 59–71 (cit. on p. 42).

[67] Paul KOEBE. 'Kontaktprobleme der konformen Abbildung'. German. In: *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse* 88 (1936), pp. 141–164 (cit. on p. 100).

[68] János KOMLÓS, János PACH, and Gerhard WOEGINGER. 'Almost Tight Bounds for ϵ-nets'. In: *Discrete & Computational Geometry* 7.2 (Mar. 1992), pp. 163–173 (cit. on p. 22).

[69] Andrey KUPAVSKII, Nabil H. MUSTAFA, and János PACH. 'Near-Optimal Lower Bounds for ϵ-nets for Half-spaces and Low Complexity Set Systems'. In: *A Journey Through Discrete Mathematics: A Tribute to Jiří Matoušek*. Springer, 2017 (cit. on pp. 45, 58, 62).

[70] Andrey KUPAVSKII and Nikita ZHIVOTOVSKIY. 'When are epsilon-nets small?' In: *CoRR* abs/1711.10414 (2017) (cit. on p. 20).

[71] Yi LI, Philip M. LONG, and Aravind SRINIVASAN. 'Improved Bounds on the Sample Complexity of Learning'. In: *Journal of Computer and System Sciences* 62.3 (2001), pp. 516–527 (cit. on p. 49).

[72] Richard J. LIPTON and Robert E. TARJAN. 'A Separator Theorem for Planar Graphs'. In: *SIAM Journal on Applied Mathematics* 36.2 (1979), pp. 177–189 (cit. on pp. 31, 101).

[73] Lynn H. LOOMIS and Hassler WHITNEY. 'An inequality related to the isoperimetric inequality'. In: *Bulletin of the American Mathematical Society* 55.10 (1949), pp. 961–962 (cit. on p. 100).

[74] Alexander M. MACBEATH. 'A Theorem on Non-Homogeneous Lattices'. In: *Annals of Mathematics* 56 (1952), pp. 269–293 (cit. on pp. 24, 58, 59).

[75] Dániel MARX. 'Efficient Approximation Schemes for Geometric Problems?' In: *Algorithms – ESA 2005: 13th Annual European Symposium, Palma de Mallorca, Spain, October 3-6, 2005. Proceedings.* 2005, pp. 448–459 (cit. on pp. 76, 83).

[76] Dániel MARX. 'Parameterized Complexity of Independence and Domination on Geometric Graphs'. In: *Parameterized and Exact Computation: Second International Workshop, IWPEC 2006, Zürich, Switzerland, September 13-15, 2006. Proceedings.* 2006, pp. 154–165 (cit. on pp. 76, 83).

[77] Jiří MATOUŠEK. *Geometric Discrepancy: An Illustrated Guide*. Algorithms and Combinatorics 18. Springer, 1999 (cit. on pp. 45, 47–49, 51, 54, 55).

[78] Jiří MATOUŠEK. *Lectures in Discrete Geometry*. Springer-Verlag, 2002 (cit. on pp. 45, 68).

[79] Jiří MATOUŠEK. *Using the Borsuk-Ulam Theorem. Lectures on Topological Methods in Combinatorics and Geometry*. Springer-Verlag, 2003 (cit. on p. 46).

[80] Jiří MATOUŠEK, János PACH, Micha SHARIR, Shmuel SIFRONY, and Emo WELZL. 'Fat Triangles Determine Linearly Many Holes'. In: *SIAM Journal on Computing* 23.1 (1994), pp. 154–169 (cit. on pp. 61, 67).

[81] Jiří MATOUŠEK and Zuzana PATÁKOVÁ. 'Multilevel Polynomial Partitions and Simplified Range Searching'. In: *Discrete & Computational Geometry* 54.1 (2015), pp. 22–41 (cit. on pp. 60, 63, 69).

[82] Ernst W. MAYR, Hans Jürgen PRÖMEL, and Angelika STEGER, eds. *Lectures on Proof Verification and Approximation Algorithms*. Lecture Notes in Computer Science 1367. Springer-Verlag Berlin Heidelberg, 1998. 348 pp. (cit. on p. 71).

[83] Gary L. MILLER, Shang-Hua TENG, William THURSTON, and Stephen A. VAVASIS. 'Separators for Sphere-packings and Nearest Neighbor Graphs'. In: *Journal of the ACM* 44.1 (Jan. 1997), pp. 1–29 (cit. on p. 90).

[84] Nabil H. MUSTAFA. 'A Simple Proof of the Shallow Packing Lemma'. In: *Discrete & Computational Geometry* 55.3 (Apr. 2016), pp. 739–743 (cit. on pp. 26, 47, 49–51, 54).

[85] Nabil H. Mustafa, Kunal Dutta, and Arijit Ghosh. 'A Simple Proof of Optimal Epsilon-nets'. In: *Combinatorica* (2017) (cit. on p. 45).

[86] Nabil H. Mustafa and Saurabh Ray. 'ϵ -Mnets: Hitting Geometric Set Systems with Subsets'. In: *Discrete & Computational Geometry* 57.3 (2017), pp. 625–640 (cit. on pp. 8, 24, 49, 50, 58, 59, 61, 62, 67, 68).

[87] Nabil H. Mustafa and Saurabh Ray. 'Improved Results on Geometric Hitting Set Problems'. In: *Discrete & Computational Geometry* 44.4 (Dec. 2010), pp. 883–895 (cit. on pp. 9, 10, 30–32, 81, 82, 84, 92).

[88] Nabil H. Mustafa and Kasturi R. Varadarajan. 'Epsilon-approximations and Epsilon-nets'. In: *Handbook of Discrete and Computational Geometry*. CRC Press LLC, 2017 (cit. on p. 45).

[89] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity. Graphs, Structures, and Algorithms*. Algorithms and Combinatorics 28. Springer-Verlag Berlin Heidelberg, 2012 (cit. on p. 95).

[90] János Pach and Pankaj K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, 1995 (cit. on p. 45).

[91] János Pach and Gábor Tardos. 'Tight lower bounds for the size of epsilon-nets'. In: *Journal of the American Mathematical Society* 26 (2013), pp. 645–658 (cit. on p. 23).

[92] Natan Rubin. 'An Improved Bound for Weak Epsilon-Nets in the Plane'. In: *CoRR* abs/1808.02686 (2018) (cit. on p. 45).

[93] Norbert Sauer. 'On the Density of Families of Sets'. In: *Journal of Combinatorial Theory, Series A* 13.1 (1972), pp. 145–147 (cit. on p. 39).

[94] Micha Sharir. 'On k-sets in arrangements of curves and surfaces'. In: *Discrete & Computational Geometry* 6.4 (Dec. 1991), pp. 593–613 (cit. on p. 42).

[95] Saharon Shelah. 'A Combinatorial Problem, Stability and Order for Models and Theories in Infinitary Languages'. In: *Pacific Journal of Mathematics* 41 (1972), pp. 247–261 (cit. on p. 39).

[96] Pál Turán. 'Egy gráfelméleti szélsőérték feladatról'. Hungarian and German. In: *Matematikai és Fizikai Lapok* 48 (1941), pp. 436–452 (cit. on p. 101).

[97] Vladimir N. Vapnik and Alexey Ya. Chervonenkis. 'Theory of uniform convergence of frequencies of events to their probabilities and problems of search for an optimal solution from empirical data.' In: *Autom. Remote Control* 1971 (1971), pp. 207–217 (cit. on pp. 21, 38, 39). Trans. of 'Теория равномерной сходимости частот появления событий к их вероятностям и задачи поиска оптималыного решения по эмпирическим данным'. Russian. In: *Автоматика и телемеханика* 2 (1971), pp. 42–53.

[98] Kasturi R. Varadarajan. 'Epsilon Nets and Union Complexity'. In: *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*. 2009, pp. 11–16 (cit. on p. 23).

[99] Kasturi R. Varadarajan. 'Weighted Geometric Set Cover via Quasi–Uniform Sampling'. In: *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC*. 2010, pp. 641–648 (cit. on pp. 5, 7, 23, 40, 45, 58, 69).

Index