Minimum Convex Partitions and Maximum Empty Polytopes

Adrian Dumitrescu* Sariel Har-Peled† Csaba D. Tóth‡

December 5, 2011

Abstract

Let S be a set of n points in d-space. A convex Steiner partition is a tiling of conv(S) with empty convex bodies. For every integer d, we show that S admits a convex Steiner partition with at most $\lceil (n-1)/d \rceil$ tiles. This bound is the best possible for affine independent points in the plane, and it is best possible apart from constant factors in every dimension $d \geq 3$. We also give the first constant-factor approximation algorithm for computing a minimum Steiner convex partition of an affine independent point set in the plane. Determining the maximum possible volume of a single tile in a Steiner partition is equivalent to a famous problem of Danzer and Rogers. We give a $(1-\epsilon)$-approximation for the maximum volume of an empty convex body when S lies in the d-dimensional unit box $[0,1]^d$.

Keywords: Steiner convex partition, Horton set, VC-dimension, epsilon-net, approximation algorithm.

1 Introduction

Let S be a set of $n \geq d+1$ points in $\mathbb{R}^d$, $d \geq 2$. A convex body C is *empty* if its interior is disjoint from S. A *convex partition* of S is a partition of the convex hull conv(S) into empty convex bodies (called *tiles*) such that the vertices of the tiles are in S. In a *convex Steiner partition* of S the vertices of the tiles are arbitrary: they can be points in S or *Steiner points*. For instance, any triangulation of S is a convex partitions of S, where the convex bodies are simplices, and so conv(S) can always be partitioned into less than dn empty convex tiles.

In this paper, we study the minimum number of tiles that a convex Steiner partition of every n points in $\mathbb{R}^d$ admits, and the maximum volume of a single tile for a given point set. The research is motivated by a longstanding open problem by Danzer and Rogers (Problem E14 in [CFG91], see also [ABFK92, BW71, BC87, FP92, PT11]): what is the maximum volume of an empty convex body $C \subset [0,1]^d$ for a set $S \subset [0,1]^d$ of n points in a unit cube? The current best bounds are $\Omega(1/n)$ and $O(\log n/n)$, respectively (for a fixed d). The lower bound comes from decomposing the unit cube by n parallel hyperplanes, each containing at least one point, into at most $n+1$ empty convex bodies. The upper bound is tight for n uniformly distributed random points in the unit cube. It is suspected that the largest volume grows faster then $\Omega(1/n)$, *i.e.*, it is $\omega(1)/n$ in any dimension $d \geq 2$.

*Department of Computer Science, University of Wisconsin–Milwaukee, WI 53201-0784, USA. Email: `dumitres@uwm.edu`. Supported in part by NSF grant DMS-1001667.

†Department of Computer Science, University of Illinois at Urbana–Champaign, Urbana, IL 61801-2302, USA. Email: `sariel@cs.uiuc.edu`. Work on this paper was partially supported by a NSF AF award CCF-0915984.

‡Department of Mathematics and Statistics, University of Calgary, Canada; and Department of Computer Science, Tufts University, Medford, MA. Email: `cdtoth@ucalgary.ca`. Supported in part by NSERC grant RGPIN 35586.

Minimum number of tiles in a convex partition. A *minimum convex partition* of S is a convex partition of S with a minimum number of tiles. Denote this number by $f_d(S)$. Further define (by slightly abusing notation)

$$f_d(n) = \max\{f_d(S) : S \subset \mathbb{R}^d, |S| = n\}.$$

Similarly define a *minimum Steiner convex partition* of S as one with a minimum number of tiles. and let $g_d(S)$ denote this number. We also define

$$g_d(n) = \max\{g_d(S) : S \subset \mathbb{R}^2, |S| = n\}.$$

There has been substantial work on estimating $f_2(n)$, and computing $f_2(S)$ for a given set S in the plane. It has been shown successively that $f_2(n) \leq \frac{10n-18}{7}$ by Neumann-Lara *et al.* [NRU04], $f_2(n) \leq \frac{15n-24}{11}$ by Knauer and Spillner [KS06], and $f_2(n) \leq \frac{4n-6}{3}$ for $n \geq 6$ by Sakai and Urrutia [SU09]. From the other direction, García-López and Nicolás [GN06] proved that $f_2(n) \geq \frac{12n-22}{11}$, for $n \geq 4$, thereby improving an earlier lower bound $f_2(n) \geq n+2$ by Aichholzer and Krasser [AK01]. Knauer and Spillner [KS06] have also obtained a $\frac{30}{11}$-factor approximation algorithm for computing a minimum convex partition for a given set $S \subset \mathbb{R}^2$, no three of which are collinear. There are also a few exact algorithms, including three fixed-parameter algorithms [FMR01, GL04, Spi08].

The state of affairs is much different in regard to convex Steiner partitions. As pointed out in [DT11], no corresponding results are known for the variant with Steiner points. Here we take the first steps in this direction, and obtain the following results.

Theorem 1. *For $n \geq d+1$, we have $g_d(n) \leq \left\lceil \frac{n-1}{d} \right\rceil$. For $d = 2$, this bound is the best possible, that is, $g_2(n) = \lceil (n-1)/2 \rceil$; and for every fixed $d \geq 2$, we have $g_d(n) \geq \Omega(n)$.*

A set of points in $\mathbb{R}^d$ is *affine independent* if every k-dimensional affine subspace contains at most $k+1$ points for $0 \leq k \leq d$. We show that in the plane every convex Steiner partition for n affine independent points, i of which lie in the interior of the convex hull, has at least $\Omega(i)$ tiles. This leads to a simple constant-factor approximation algorithm.

Theorem 2. *Given a set S of n affine independent points in $\mathbb{R}^2$, a ratio 3 approximation of a minimum convex Steiner partition of S can be computed in $O(n \log n)$ time.*

The *average* volume of a tile in a Steiner partition of n points in the unit cube $[0,1]^d$ is an obvious lower bound for the maximum possible volume of a tile, and for the maximum volume of any empty convex body $C \subset [0,1]^d$. The lower bound $g_d(n) \geq \Omega(n)$ in Theorem 1 shows that the average volume of a tile is $O(1/n)$ in some instances, where the constant of proportionality depends only on the dimension. This implies that a simple "averaging" argument is not a viable avenue for finding a solution to the problem of Danzer and Rogers.

Maximum empty polytope among n points in a unit cube. In the second part of the paper, we consider the following problem: Given a set of n points in rectangular box B in $\mathbb{R}^d$, find a maximum-volume empty convex body $C \subset B$. Since the ratio between volumes is invariant under affine transformations, we may assume without loss of generality that $B = [0,1]^d$. We therefore have the problem of computing a maximum volume empty convex body $C \subset [0,1]^d$ for a set of n points in $[0,1]^d$. An exact algorithm is known in the plane. Eppstein *et al.* [EORW92] find the maximum area empty convex k-gon among n points in the unit square $[0,1]^2$ in $O(kn^3)$ time. Since a maximum area convex body is a convex polygon with at most n vertices, it can be computed exactly in $O(n^4)$ time with their algorithm. Note also that by John's ellipsoid theorem, the maximum volume empty ellipsoid in $[0,1]^d$ gives a $\frac{1}{d^d}$-approximation.

We present a $(1-\varepsilon)$-approximation for the maximum volume empty convex body C_{opt} by first guessing an approximate inscribed ellipsoid of C_{opt}, which is an empty ellipsoid whose volume is an $\frac{1}{d^d}$-approximation of $\text{vol}(C_{\text{opt}})$, and then refining it to a sufficiently close approximation of C_{opt}.

Theorem 3. *Given a set S of n points in $[0,1]^d$, one can $(1-\varepsilon)$-approximate the maximum-volume empty convex body in $[0,1]^d$. The running time of the approximation algorithm is*

$$O\left(n^{1+d(d-1)/2} 2^{O\left(1/\varepsilon^d\right)} \log^d n\right).$$

Related work. Decomposing polygonal domains into convex sub-polygons has been also studied extensively. We refer to the article by Keil [Kei00] for a survey of results up to the year 2000. For instance, when the polygon may contain holes, obtaining a minimum convex partition is NP-hard, regardless of whether Steiner points are allowed. For polygons without holes, Chazelle and Dobkin [CD79] obtained an $O(n + r^3)$ time algorithm for the problem of decomposing a polygon with n vertices, r of which are reflex, into convex parts, with Steiner points permitted. As remarked by Keil [Kei00], although there are an infinite number of possible locations for the Steiner points, surprisingly a dynamic programming approach is amenable to obtain an exact (optimal) solution; see also [KS02, She92].

Fevens *et al.* [FMR01] designed a polynomial time algorithm for computing a minimum convex partition for a given set of n points in the plane if the points are arranged on a constant number of convex layers. The problem of minimizing the total Euclidean length of the edges of a convex partition has been also considered. Grantson and Levcopoulos [GL04], and Spillner [Spi08] proved that the shortest convex partition and convex Steiner partition problems are fixed parameter tractable, where the parameter is the number of points of P lying in the interior of $\text{conv}(P)$. Dumitrescu and Tóth [DT11] proved that every set of n points in $\mathbb{R}^2$ admits a convex Steiner partition which is at most $O(\log n/\log\log n)$ times longer than the minimum spanning tree, and this bound cannot be improved. Without Steiner points, the best upper bound for the ratio of the minimum length of a convex partition and the length of a minimum spanning tree (MST) is $O(n)$ [Kir80].

For finding a maximum volume empty axis-parallel box amidst n points in $[0,1]^d$, Backer and Keil [BK10] reported an algorithm with worst-case running time of $O(n^d \log^{d-2} n)$. An empty axis-aligned box whose volume is at least $(1-\varepsilon)$ of the maximum can be computed in $O\left(\left(\frac{8ed}{\varepsilon^2}\right)^d n \log^d n\right)$ time by the algorithm of Dumitrescu and Jiang [DJ09].

Lawrence and Morris [LM09] studied the minimum integer $k_d(n)$ such that the complement $\mathbb{R}^d \setminus S$ of any n-element set $S \subset \mathbb{R}^d$, not in a hyperplane, can be *covered* by $k_d(n)$ convex sets. They prove $k_d(n) \geq \Omega(\log n/d \log\log n)$. Bounds on $k_d(n)$ are also related to the *invisibility graph* of a point set [CKM$^+$10]. Note that covering the complement of n uniformly distributed points in $[0,1]^d$ requires at least $\Omega(n/d\log n)$ convex sets, which follows from the upper bound in the problem of Danzer and Rogers.

2 Combinatorial bounds

In this section we prove Theorem 1. We start with the upper bound. The following very simple algorithm returns a convex Steiner partition with at most $\lceil (n-1)/d \rceil$ tiles for any n points in $\mathbb{R}^d$.

Algorithm **A1**:
Step 1. Compute the convex hull $R \leftarrow \text{conv}(S)$ of S. Let $A \subseteq S$ be the set of hull vertices, and let $B = S \setminus A$ denote the remaining points.

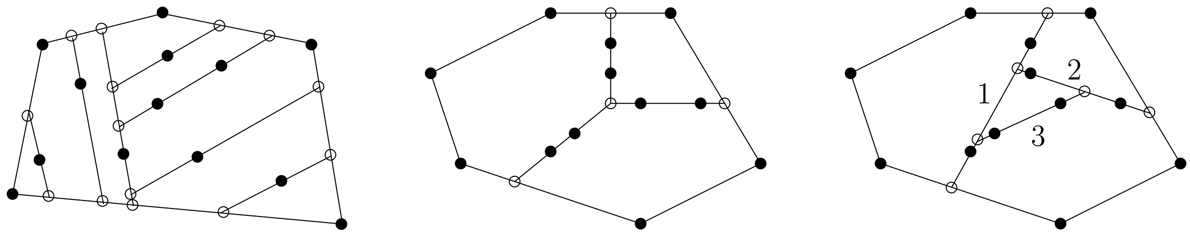

Figure 1: Steiner convex partitions with Steiner points drawn as hollow circles. Left: a convex Steiner partition of a set of 13 points. Middle: A Steiner partition of a set of 10 points into three tiles. Right: A Steiner partition of the same set of 10 points into 4 tiles, generated by Algorithm **A1** (the labels reflect the order of execution).

STEP 2. Compute $\text{conv}(B)$, and let H be the supporting hyperplane of an arbitrary $(d-1)$-dimensional face of $\text{conv}(B)$. Denote by H^+ the halfspace that contains B, and $H^- = \mathbb{R}^d \setminus H^+$. The hyperplane H contains d points of B, and it decomposes R into two convex bodies: $R \cap H^-$ is empty and $R \leftarrow R \cap H^+$ contains all points in $B \setminus H$. Update $B \leftarrow B \setminus H$ and $R \leftarrow R \cap H^+$.

STEP 3. Repeat STEP 2 with the new values of R and B until B is the empty set. (If $|B| < d$, then any supporting hyperplane of B completes the partition.)

It is obvious that the algorithm generates a Steiner convex partition of S. An illustration of Algorithm **A1** on a small planar example appears in Figure 1 (right). Let h and i denote the number of hull and interior points of S, respectively, so that $n = h + i$. Each hyperplane used by the algorithm removes d interior points of S (with the possible exception of the last round if i is not a multiple of d). Hence the number of convex tiles is $1 + \lceil i/d \rceil$, and we have $1 + \lceil i/d \rceil = \lceil (i+d)/d \rceil \leq \lceil (n-1)/d \rceil$, as required.

Lower bound in the plane. A matching lower bound in the plane is given by the following construction. For $n \geq 3$, let $S = A \cup B$, where A is a set of 3 non-collinear points in the plane, and B is a set of $n-3$ points that form a regular $(n-3)$-gon in the interior of $\text{conv}(A)$, so that $\text{conv}(S) = \text{conv}(A)$ is a triangle. If $n = 3$, then $\text{conv}(S)$ is an empty triangle, and $g_2(S) = 1$. If $4 \leq n \leq 5$, S is not in convex position, and so $g_2(S) \geq 2$. Suppose now that $n \geq 6$.

Consider an arbitrary convex partition of S. Let o be a point in the interior of $\text{conv}(B)$ such that the lines os, $s \in S$, do not contain any edges of the tiles. Refer to Figure 2 (left). For each point $s \in B$, choose a *reference point* $r(s) \in \mathbb{R}^2$ on the ray $\overrightarrow{os}$ in $\text{conv}(A) \setminus \text{conv}(B)$ sufficiently close to point s, and lying in the interior of a tile. Note that the convex tile containing o cannot contain any reference points. We claim that any tile contains at most 2 reference points. This immediately implies $g_2(S) \geq 1 + \lceil (n-3)/2 \rceil = \lceil (n-1)/2 \rceil$.

Suppose, to the contrary, that a tile τ contains 3 reference points r_1, r_2, r_3, corresponding to the points s_1, s_2, s_3. Refer to Figure 2. Note that o cannot be in the interior of τ, otherwise τ would contain all points s_1, s_2, s_3 in its interior. Hence $\text{conv}\{o, s_1, s_2, s_3\}$ is a quadrilateral, and $\text{conv}\{o, r_1, r_2, r_3\}$ is also a quadrilateral, since the reference points are sufficiently close to the corresponding points in B. We may assume w.l.o.g. that vertices of $\text{conv}\{o, s_1, s_2, s_3\}$ are o, s_1, s_2, s_3 in counterclockwise order. Then s_2 lies in the interior of $\text{conv}\{o, r_1, r_2, r_3\}$. We conclude that every tile τ contains at most 2 reference points, as required.

Lower bounds for $d \geq 3$. The above argument does not extend to higher dimensions, since an empty convex tile contains at most k reference points only if the cone with apex o spanned by *any*

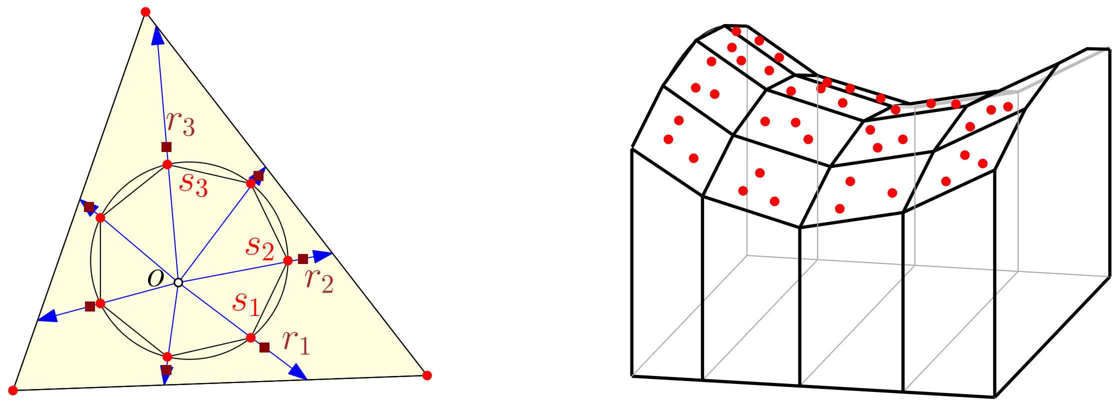

Figure 2: Left: lower bound construction in $\mathbb{R}^2$. Right: affine independent points on a saddle surface in $\mathbb{R}^3$.

k reference points is empty. A similar construction works in for any $d \geq 2$, but the lower bound no longer matches the upper bound $g_d(n) \leq \lceil (n-1)/d \rceil$ for $d \geq 3$.

Recall that a *Horton set* [Hor83] is a set S of n points in the plane such that the convex hull of any 7 points is nonempty. Horton's construction, as presented in [BF87] is defined inductively when n is a power of 2. However, we can construct Horton sets with any number of points n, by successively deleting points from the boundary of the convex hull of a Horton set of size $2^{\lceil \log n \rceil}$.

Valtr [Val92] generalized Horton sets to d-space. For every $d \in \mathbb{N}$, there exists a minimal integer $h(d)$ with the property that for every $n \in \mathbb{N}$ there is a set S of n affine independent points in $\mathbb{R}^d$ such that the convex hull of any $h(d)+1$ points in S is nonempty. It is known that $h(2) = 6$, Valtr proved that $h(3) \leq 22$, and in general that $h(d) \leq 2^{d-1}(N(d-1)+1)$, where $N(d-1)$ is the product of the first $d-1$ primes.

We construct a set S of $n \geq d+1$ points in $\mathbb{R}^d$ as follows. Let $S = A \cup B$, where A is a set of $d+1$ affine independent points in $\mathbb{R}^d$, and B is a generalized Horton set of $n-(d+1)$ points in the interior of $\text{conv}(A)$, such that the interior of any $h(d)+1$ points from B contains some point in B.

Consider an arbitrary convex Steiner partition of S. Every point $b \in B$ is in the interior of $\text{conv}(S)$, and so it lies on the boundary of at least 2 convex tiles. For each $b \in B$, place two *reference points* in the interiors of 2 distinct tiles incident to b. Every tile contains at most $h(d)$ reference points. Indeed, if a tile contains $h(d)+1$ reference points, then it is incident to $h(d)+1$ points in B, and some point of B lies in the interior of the convex hull of these points, a contradiction.

We have $2(n-d-1)$ reference points, and every tile contains at most $h(d)$ of them. So the number of tiles is at least $\lceil 2(n-d-1)/h(d) \rceil$.

3 Approximating the minimum convex Steiner partition in $\mathbb{R}^2$

In this section we prove Theorem 2 by showing that our simple-minded algorithm **A1** from Section 2 achieves a constant-factor approximation in the plane if the points in S are affine independent.

Approximation ratio. Recall that algorithm **A1** computes a Steiner partition of $\text{conv}(S)$ into at most $1 + \lceil i/2 \rceil$ parts, where i stands for the number of interior points of S.

If $i = 0$, the algorithm computes an optimal partition, *i.e.*, ALG = OPT = 1. Assume now that $i \geq 1$. Consider an optimal convex Steiner partition Π of S with OPT tiles. We construct a planar multigraph $G = (V, E)$ as follows. The *faces* of G are the convex tiles and the exterior of conv(S) (the outer face). The *vertices* V are the points in the plane incident to at least 3 faces (counting the outer face as well). Since $i \geq 1$, G is non-empty and we have $|V| \geq 2$. Each *edge* in E is a Jordan arc on the common boundary of two faces. An edge between two bounded faces is a straight line segment, and so it contains at most two interior points of S. An edge between the outer face and a bounded face is a convex arc, containing hull points from S. Double edges are possible if two vertices of the outer face are connected by a straight line edge and a curve edge along the boundary—in this case these two parallel edges bound a convex face. No loops are possible in G. Since Π is a convex partition, G is connected.

Let v, e, and f, respectively, denote the number of vertices, edges, and bounded (convex) faces of G; in particular, $f = \text{OPT}$. By Euler's formula for planar multigraphs, we have $v - e + f = 1$, that is, $f = e - v + 1$. By construction, each vertex of G is incident to at least 3 edges, and every edge is incident to two vertices. Therefore, $3v \leq 2e$, or $v \leq 2e/3$. Consequently, $f = e - v + 1 \geq e - 2e/3 + 1 = e/3 + 1$. Since S is affine independent, each straight-line edge of G contains at most 2 interior points from S. Curve edges along the boundary do not contain interior points. Hence each edge in E is incident to at most two interior points in S, thus $i \leq 2e$. Substituting this into the previous inequality on f yields $\text{OPT} = f \geq e/3 + 1 \geq i/6 + 1$. Comparing this lower bound with the upper bound $\text{ALG} \leq \lceil i/2 \rceil + 1$, we conclude that

$$\frac{\text{ALG}}{\text{OPT}} \leq \frac{\lceil i/2 \rceil + 1}{i/6 + 1} \leq 3\,\frac{i+3}{i+6} < 3,$$

and the approximation ratio of 3 follows.

Tightness of the approximation ratio. We first show that the above ratio 3 is tight for Algorithm **A1**. We construct a planar point set S as follows. Consider a large (say, hexagonal) section of a hexagonal lattice. Place Steiner vertices at the lattice points, and place two points in S on each lattice edge. Slightly perturb the lattice, and add a few more points in S near the boundary, and a few more Steiner points, so as to obtain a convex Steiner partition of S with no three points collinear. Denote by v, e, and f, the elements of the planar multigraph G as before. Since we consider a large lattice section, we have $v, e, f \to \infty$. We write $a \sim b$, whenever $a/b \to 1$. As before, we have $f + v = e + 1$, and since each non-boundary edge is shared by two convex faces, we have $e \sim 6f/2 = 3f$. By construction, $i \sim 2e \sim 6f$, hence $f \sim i/6$. Therefore the convex partition constructed above has $f \sim i/6$, while Algorithm **A1** constructs one with about $i/2$ faces. Letting $e \to \infty$, then $i \to \infty$, and the ratio ALG/OPT approaches 3 in the limit: $\text{ALG}/\text{OPT} \sim (i/2)/(i/6) = 3$.

Time analysis. It is easy to show that Algorithm **A1** runs in $O(n \log n)$ time for a set S of n points in the plane. We employ the semi-dynamic (delete only) convex hull data structure of Hershberger and Suri [HS92]. This data structure supports point deletion in $O(\log n)$ time, and uses $O(n)$ space and $O(n \log n)$ preprocessing time. We maintain the boundary of a convex polygon R in a binary search tree, a set $B \subset S$ of points lying in the interior of R, and the convex hull conv(B) with the above semi-dynamic data structure [HS92]. Initially, $R = \text{conv}(S)$, which can be computed in $O(n \log n)$ time; and $B \subset S$ is the set of interior points. In each round of the algorithm, consider the supporting line H of an arbitrary edge e of conv(B) such that B lies in the halfplane H^+. The two intersection points of H with the boundary of R can be computed in $O(\log n)$ time. At the end of the round, we can update $B \leftarrow B \setminus H$ and conv(B) in $O(k \log n)$ time, where k is the number of points removed from B; and we can update $R \leftarrow R \cap H^+$ in $O(\log n)$ time. Every point is removed from B exactly once, and the number of rounds is at most $\lceil (n-3)/2 \rceil$, so the total update time is $O(n \log n)$ throughout the algorithm.

Remark. Interestingly enough, in dimensions 3 and higher, Algorithm **A1** does not give a constant-factor approximation. For every integer n, one can construct a set S of n affine independent points in $\mathbb{R}^3$ such that $i = n - 4$ of them lie in the interior of $\text{conv}(S)$, but the minimum convex Steiner partition has only $O(\sqrt{n})$ tiles. In contrast, Algorithm **A1** computes a Steiner partition with $i/3 = (n-4)/3$ convex tiles.

We first construct the convex tiles, and then describe the point set S. Specifically, S consists of 4 points of a large tetrahedron, and 3 affine independent points on the common boundary of certain pairs of adjacent tiles.

Let $k = \lceil \sqrt{(n-4)/3} \rceil$. Place $(k+1)^2$ Steiner points $(a, b, a^2 - b^2)$ on the saddle surface $z = x^2 - y^2$ for pairs of integers $(a, b) \in \mathbb{Z}^2$, $-\lfloor k/2 \rfloor \le a, b \le \lceil k/2 \rceil$. The four points $\{(x, y, x^2 - y^2) : x \in \{a, a+1\}, y \in \{b, b+1\}\}$ form a parallelogram for every $(a, b) \in \mathbb{Z}^2$, $-\lfloor k/2 \rfloor \le a, b \le \lceil k/2 \rceil - 1$. Refer to Figure 2 (right). These parallelograms form a terrain over the region $\{(x, y) : -\lfloor k/2 \rfloor \le x, y \le \lceil k/2 \rceil\}$. Note that no two parallelograms are coplanar. Subdivide the space *below* this terrain by vertical planes $x = a$, $-\lfloor k/2 \rfloor \le a \le \lceil k/2 \rceil$. Similarly, subdivide the space *above* this terrain by planes $y = b$, $-\lfloor k/2 \rfloor \le b \le \lceil k/2 \rceil$. We obtain $2k$ interior-disjoint convex regions, k above and k below the terrain, such that the common boundary of a region above and a region below is a parallelogram of the terrain. The points in $\mathbb{R}^3$ that do not lie above or below the terrain can be covered by 4 convex wedges.

Enclose the terrain in a sufficiently large tetrahedron T. Clip the $2k$ convex regions and the 4 wedges into the interior of T. These $2k+4$ convex bodies tile T. Choose 3 noncollinear points of S in each of the k^2 parallelograms, such that no 4 points are coplanar and they are affine independent from the vertices of T. Let the point set S be the set of 4 vertices of the large tetrahedron T and the $3k^2$ points selected from the parallelograms.

4 Approximating the maximum empty convex body

Let S be a set of points in the unit cube $[0, 1]^d \subseteq \mathbb{R}^d$. Our task is to approximate the largest convex body $C \subseteq [0, 1]^d$ that contains no points of S in its interior. Let $C_{\text{opt}} = C_{\text{opt}}(S)$ denote this body, and let $\text{vol}_{\text{opt}}(S)$ denote its volume.

By Theorem 1, $\text{vol}(C_{\text{opt}}) \ge \Omega(1/n)$. The diameter of C is bounded from above by $\sqrt{d}$, and its width is bounded from below by c_d/n, where c_d is some constant that depends on the dimension. By John's ellipsoid theorem [Mat02], for any compact convex body C in $\mathbb{R}^d$ there exists an ellipsoid $\mathcal{E}$ such that $\mathcal{E} \subseteq C \subseteq u + d\mathcal{E}$ for some vector $u \in \mathbb{R}^d$, where we denote by $k\mathcal{E}$ the ellipsoid $\mathcal{E}$ scaled up by a factor of k and having the same center as $\mathcal{E}$. It follows that $\text{vol}(C)/d^d \le \text{vol}(\mathcal{E}) \le \text{vol}(C)$.

Lemma 1. *A set of at most $d^2 + d$ points in general position in $\mathbb{R}^d$ determine a unique ellipsoid passing through these points.*

Proof. Observe that the boundary of the ellipsoid is defined by a quadratic equation of the form $\sum_{i,j} \alpha_{i,j} x_i x_j + \sum_i \alpha_i x_i = 1$. As such, it is defined by $d^2 + d$ coefficients (notice, that not all possible coefficients correspond to legal ellipsoids). Requiring that a specific point lies on the boundary of the ellipsoid, corresponds to a linear equation defined by this point over $d^2 + d$ variables. Under a general position assumption, $d^2 + d$ such equations determine the solution to this linear system, hence the coefficients of the ellipsoid (and thus the ellipsoid itself). □

Our analysis in Lemma 1 is loose, as an ellipsoid has $d(d+3)/2$ degrees of freedom (indeed, fix its center, and pick its axes one by one, using the orthogonality of the axes to reduce the number of degrees of freedom being counted).

Lemma 2. *Assume that $\mathcal{E} \subseteq [0,1]^d$ is an ellipsoid, and $\mathrm{vol}(\mathcal{E}) \geq \rho$. Then, one can compute a set of Q of $O(1/\rho \log(1/\rho))$ points such that, with probability $\geq 1 - (\rho/2)^{O(1)}$, one of the points of Q lies in $\mathcal{E}/2$.*

Proof. Observe that ellipsoids in $\mathbb{R}^d$ have bounded VC dimension. Indeed, by Lemma 1 an ellipsoid has at most d^2+d points of S on its boundary. Now, given an ellipsoid we transform it continuously into an equivalent ellipsoid containing the same set of points, while having a maximum number of points on its boundary. As such, its shattering dimension is bounded by $O(d^2)$, thus its VC dimension is bounded by $d' = O(d^2 \log d)$. (It is likely that a better bound on the VC dimension is possible by being more careful.)

By the ε-net theorem [Mat99, Ch. 5.2], a sample Q of size $O(d'/\rho \log 1/\rho)$ will hit any ellipsoid in the unit cube with volume $\geq \rho/2^d$ with high probability. In particular, this sample hits $\mathcal{E}/2$. □

4.1 Handling the fat case

In the following, assume that $m > 0$ is some integer, and consider the grid point set $\mathcal{G}(m) = \left\{ (i_1, \ldots, i_d)/m \;\middle|\; i_1, \ldots, i_d \in \{0, 1, \ldots, m\} \right\}$. Let $S \subseteq [0,1]^d$ be a point set such that $\mathrm{vol}_{\mathrm{opt}}(S) \geq \mu$, where μ is some constant, and let C_{opt} be the corresponding largest empty convex body in $[0,1]^d$. Given a grid $\mathcal{G}(m)$, we call $\mathrm{conv}(C_{\mathrm{opt}} \cap \mathcal{G}(m))$ the *discrete hull* of C_{opt}. (The discrete hull has some fascinating properties that are of no interest for us here; we refer the interested reader to [Har98].) We need the following easy lemma.

Lemma 3. *Let $C \subseteq [0,1]^d$ be a convex body and $D = \mathrm{conv}(C \cap \mathcal{G}(m))$. Then we have $\mathrm{vol}(C) - \mathrm{vol}(D) = O(1/m)$, where the constant of proportionality depends only on d.*

Proof. Consider a point $p \in C \setminus D$. Consider the set of $2d$ points $X = \{p \pm 2(d/m)e_i\}$, where e_i is the unit vector having one in the ith coordinate, and 0 everywhere else. If one the points of X is outside C, then the distance from p to the boundary of C is at most $2d/m$. Otherwise, the cube $p + [-2,2]^d/m$ is contained in the "diamond" $\mathrm{conv}(X)$, which is in turn contained in C. But then, the grid points of the grid cell of $\mathcal{G}(m)$ containing p are in C, and p can not be outside D. A contradiction.

It follows that all the points of the corridor $C \setminus D$ are at distance at most $2(d/m)$ from the boundary of C. The volume of the boundary of C is bounded by the volume of the boundary of the unit cube, namely $2d$. As such, the volume of this corridor is $\mathrm{vol}(\partial C)\, O(d/m) \leq (2d)(2d/m) = O(d^2/m)$. For a fixed d, this is $O(1/m)$, as claimed. □

Lemma 3 implies that if $\mathrm{vol}_{\mathrm{opt}}(S) \geq \mu$, where μ is some constant, then we can concentrate our search on convex polytopes that have their vertices at grid points in $\mathcal{G}(m)$, where $m = O(1/\varepsilon\mu)$.

4.2 Finding a large empty convex polygon

We first re-derive a result of Eppstein *et al.* [EORW92] concerning an exact algorithm for a related problem, with a simple proof.

Lemma 4. *Given a set S of n points and a set Q of m points in the plane, one can compute a convex polygon of the largest area with vertices in S that does not contain any point of Q in its interior in $O(n^3 m + n^4)$ time. The algorithm has the same running time if Q is a set of m forbidden rectangles.*

Proof. The algorithm works by dynamic programming. First, we compute for all triangles with vertices from S whether they contain a forbidden point inside them; trivially this can be done in $O(n^3 m)$ time. We then build a directed graph G on the allowable triangles, connecting two triangles Δ and Δ' if they share their left endpoint, are interior disjoint, share an edge, and their union forms a convex quadrilateral. We orient the edge from the triangle that is most counterclockwise (around the common vertex) to the other triangle. All edges are oriented "upwards", so G is a directed acyclic graph (DAG). Observe that G has $O(n^3)$ vertices (allowable triangles) and the maximum out-degree in G is bounded from above by n.

The weight of a vertex corresponding to a triangle is equal to its area. Clearly, a convex polygon corresponds to a path in G, namely the triangulation of the polygon from its leftmost vertex, and its weight is the area of the polygon. Finding the maximum weight path can be done in linear time in the size of the DAG; see *e.g.*, [DPV08, Section 4.7]. G has $O(n^3)$ vertices and $O(n^4)$ edges, and as such the overall running time is $O(n^3 m + n^4)$. □

Lemma 5. *Given a set* $S \subseteq [0,1]^2$ *of* n *points, such that* $\mathrm{vol}_{\mathrm{opt}}(S) \geq \rho$*, and a parameter* $\varepsilon > 0$*, one can compute an empty convex body* $C \subseteq [0,1]^2$ *such that* $\mathrm{vol}(C) \geq (1-\varepsilon)\mathrm{vol}_{\mathrm{opt}}(S)$*. The running time of the algorithm is* $O\big(n + 1/(\varepsilon\rho)^8\big)$*.*

Proof. Consider the grid $\mathcal{G}(m)$. By Lemma 3 we can restrict our search to a grid polygon. Going a step further, we mark all the grid cells containing points of S as forbidden. Arguing as in Lemma 3, one can show that the area of the largest convex grid polygon avoiding the forbidden cells is at least $\mathrm{vol}_{\mathrm{opt}}(S) - c/m$, where c is some constant.

We now restrict our attention to the task of finding this largest polygon. We have a set Q of $O(m^2)$ grid points that might be used as vertices of the grid polygon, and a set of $O(m^2)$ grid cells that can not intersect the interior of the computed polygon. Using Lemma 4 finding the largest empty polygon takes $O(m^8)$ time. Setting $m = 1/\varepsilon\rho$, we get an algorithm with overall running time $O\big(n + 1/(\varepsilon\rho)^8\big)$. □

4.3 The higher dimensional case

Lemma 6. *Given a set* $S \subseteq [0,1]^d$ *of* n *points, such that* $\mathrm{vol}_{\mathrm{opt}}(S) \geq \mu$*, and a parameter* $\varepsilon > 0$*, one can compute an empty convex body* $C \subseteq [0,1]^d$*, such that* $\mathrm{vol}(C) \geq (1-\varepsilon)\mathrm{vol}_{\mathrm{opt}}(S)$*. The running time of the algorithm is* $O\Big(n + m^{2d} 2^{m^d}\Big)$*, where* $m = O(1/\varepsilon\mu)$*.*

Proof. Consider the grid $\mathcal{G}(m)$. Let X be the set of all the grid cells of $\mathcal{G}(m)$ that contain points from S. Observe that $|X| = O(m^d)$. Next, let S' be the set of all grid points of $\mathcal{G}(m)$. Enumerating overall possible subsets of the grid points S', generates 2^{m^d} candidate sets. Checking if such a convex hull intersects the interior of a specific forbidden cell of X can be done in linear time; that is $O(m^d)$. As such, checking if such a candidate set convex hull is a valid solution, takes $O\big(m^{2d}\big)$ time. Returning the largest such subset found yields the desired approximation. □

The running time in Lemma 5 can be somewhat reduced by being more careful about the details, for instance by using a result in [BP92].

4.4 A better approximation in the plane

Lemma 7. *Given a set* $S \subseteq [0,1]^2$ *of* n *points with* $\mathrm{vol}_{\mathrm{opt}}(S) \geq \rho$ *and parameters* $\delta, \varepsilon > 0$*, one can compute an empty convex body* $C \subseteq [0,1]^2$ *such that* $\mathrm{vol}(C) \geq (1-\varepsilon)\mathrm{vol}_{\mathrm{opt}}(S)$*. The running time* *of the algorithm is* $O\Big(\frac{n\log^2 n}{\varepsilon\rho}\big(\log n + \frac{1}{\varepsilon^8}\big)\log\frac{1}{\delta\rho}\Big)$. *The algorithm succeeds with probability* $\geq 1-\delta$. *For a fixed* δ, *the running time is* $O\big(n^2\log^3 n\,\varepsilon^{-1}\big(\log n + \varepsilon^{-8}\big)\big)$.

Proof. Let $\mathcal{E}$ be an ellipse of maximum area contained inside $C_{\text{opt}} = C_{\text{opt}}(S)$. As suggested by Lemma 2, let $\mathcal{R}$ be a random sample of $O(1/\rho\log(1/\delta\rho))$ points from $[0,1]^2$. With probability $\geq 1-\delta$ this sample hits $\mathcal{E}/2$. The intuitive idea is now to guess a copy of $\mathcal{E}/2$ and center it at one of the points of $\mathcal{R}\cap\mathcal{E}/2$. In particular, let $p\in\mathcal{R}$ be the guess for the desired center of this ellipse. Naturally, to guess the ellipse itself, we need to guess the lengths of the two axes of $\mathcal{E}/2$, and their orientation.

Since the shortest axis of $\mathcal{E}/2$ has length at least $1/8n$, and the maximum length axis has length at most $\sqrt{2}$, it follows that if we want to guess the lengths of the two axes, up to a factor of two, we need to consider only $O(\log^2 n)$ possibilities. Indeed, we consider the canonical lengths $\ell_i = 2^i/8n$, for $i = 0,\ldots,\lceil\log_2(8n)\rceil$.

Consider now the bounding box of the guessed ellipse $\mathcal{F}$ (we do not know its orientation yet). Scale it up by a factor of 4 so that its contains C_{opt}. Let B be the resulting box fixed in the right orientation. We can apply Lemma 5 to B (as the unit square) to get the desired approximation. The polygon C_{opt} occupies a constant fraction of the area of B, and as such the resulting running time is $O(n + 1/\varepsilon^8)$. Note that the algorithm of Lemma 5 partitions B into a grid with $O(1/\varepsilon^2)$ cells. The approximation algorithm cares only about which cells are empty or not.

Since we do not know the orientation of B, perform a rotational sweeping algorithm [dBCvKO08], rotating B around p. Whenever a point of S moves form one grid cell to another in the grid of B, stop and recompute the optimal solution. We have $O(n/\varepsilon)$ such events during the sweeping process, and an update requires $O(1/\varepsilon^8)$ time to handle. Hence the running time for computing this polygon for p is $O((n/\varepsilon)\log n + n/\varepsilon^9)$.

Since we have to repeat this for all the points in the random sample $\mathcal{R}$, and all ellipse axes, the overall running time is

$$O\left(\frac{n\log^2 n}{\varepsilon\rho}\left(\log n + \frac{1}{\varepsilon^8}\right)\log\frac{1}{\delta\rho}\right).$$

Since $\rho = \Omega(1/n)$, for a fixed δ, the above expression is bounded by $O\big(n^2\log^3 n\,\varepsilon^{-1}\big(\log n + \varepsilon^{-8}\big)\big)$, as claimed. □

By doing an exponentially decreasing search for ρ, the running time increases only by a constant factor (this is a geometrically decreasing series, hence the term with the last value of ρ dominates the whole running time). We summarize our result for the plane in the following.

Theorem 4. *Given a set* $S\subseteq[0,1]^2$ *of* n *points and parameters* $\varepsilon,\delta>0$, *one can compute an empty convex body* $C\subseteq[0,1]^d$, *such that* $\text{vol}(C)\geq(1-\varepsilon)\text{vol}_{\text{opt}}(S)$. *The running time of the algorithm is* $O\Big(\frac{n\log^2 n}{\varepsilon\rho}\big(\log n + \frac{1}{\varepsilon^8}\big)\log\frac{1}{\delta\rho}\Big)$, *where* $\rho = \text{vol}_{\text{opt}}(S)$. *The algorithm succeeds with probability* $\geq 1-\delta$. *For a fixed* δ, *the running time is* $O\big(n^2\log^3 n\,\varepsilon^{-1}\big(\log n + \varepsilon^{-8}\big)\big)$.

Remark. If $\rho = \Omega(1)$ the running time of this algorithm in the plane is near linear in n.

4.5 A better approximation in higher dimensions

As before, consider an ellipsoid $\mathcal{E}$ of maximum volume contained inside $C_{\text{opt}}(S)$. The ellipsoid $\mathcal{E}$ is defined by a set of d orthogonal axes. Consider a situation where we replaced each one of these axes by an approximate axis that makes an angle at most $\alpha = 1/(cn)$ with its corresponding axis, where c is some absolute constant. Let $\mathcal{E}'$ to be a rotation of $\mathcal{E}$ such that it now uses these rotated axes. Since the length of the shortest axis of $\mathcal{E}$ is $\Omega(1/n)$, it is easy to verify that $\mathcal{E}$ and $\mathcal{E}'$ have an intersection which is quite large, which in particular includes a copy of $\mathcal{E}/2$. For our approximation algorithm, this implies that we can use $\mathcal{E}'/2$ as an approximation to the largest ellipsoid inside $C_{\text{opt}}(S)$.

We next enumerate all such possible approximate ellipsoids (guessing only the lengths of its axes and their orientations, not its center point), as follows:

(A) It suffices to guess every axis length up to a factor of 2. Since the minimum length of an axis in our case is $\Omega(1/n)$ and the maximum is $\sqrt{d}$, it follows that the number of possible lengths to be considered is $O\left(\log^d n\right)$.

(B) Guessing the axes of the ellipsoid can be done as follows. We spread a uniform grid on the sphere of directions, with angular distance at most $O(1/n)$ between any point on the sphere and its closest point. Clearly, this requires $O(n^{d-1})$ points. We try each one point as the direction of the first axis of the ellipsoid, and we generate the remaining axes directions on the orthogonal hyperplane for the chosen direction. Clearly, overall, this would generate $O(n^{d(d-1)/2})$ possibilities overall.

Consider the guessed ellipsoid $\mathcal{E}''$ that is a good approximation to the ellipsoid $\mathcal{E}$. Consider the bounding box B'' of $\mathcal{E}''$ that is aligned with $\mathcal{E}''$ and is scaled by a constant, such that a random translation of a grid having B'' as grid cell, misses $C_{\text{opt}}(S)$ with constant probability (*i.e.*, $C_{\text{opt}}(S)$ is fully contained inside such a grid cell). It is then an easy matter to generate a constant number of candidate shifts such that one of them misses $C_{\text{opt}}(S)$.

Given such a candidate grid, we now compute the largest volume convex body avoiding the input points inside each of these grid cells (each such bounding box needs to be clipped to the unit cube $[0,1]^d$). Note, that inside the right grid cell, $C_{\text{opt}}(S)$ takes a constant fraction of the volume of the grid cell. As such, we can deploy Lemma 5 in this case, and get the desired $(1-\varepsilon)$-approximation. Putting everything together, we obtain the following.

Theorem 3. *Given a set S of n points in $[0,1]^d$, one can $(1-\varepsilon)$-approximate the maximum volume empty convex body in $[0,1]^d$. The running time of the approximation algorithm is*

$$O\left(n^{1+d(d-1)/2} 2^{O\left(1/\varepsilon^d\right)} \log^d n\right).$$

Remark. Consider a set S of n points in $\mathbb{R}^d$. The approximation algorithm we have presented can be modified to approximate the largest empty tile, *i.e.*, the largest empty convex body contained in $\text{conv}(S)$, rather than $[0,1]^d$. The running time is slightly worse, since we need to take the boundary of $\text{conv}(S)$ into account. We omit the details.

5 Conclusion

Interesting questions remain open regarding the structure of optimal convex Steiner partitions and the computational complexity of computing such partitions. Other questions relate to the problem of finding the largest empty convex body in the presence of points. We list some of them:

(1) Is there a polynomial-time algorithm for computing a minimum convex Steiner partition of a given set of n points in $\mathbb{R}^d$? Is there one for points in the plane?

(2) Is there a constant-factor approximation algorithm for the minimum convex Steiner partition of an arbitrary point sets in $\mathbb{R}^d$ (without the *affine independence* restriction)? Is there one for points in the plane?

(3) For $d > 2$, the running time of our approximation algorithm for the maximum empty polytope has a factor of the form $n^{O(d^2)}$. It seems natural to conjecture that this term can be reduced to $n^{O(d)}$. Furthermore, the current running time has a wild, doubly exponential dependency on d. It would be interesting to reduce it to something only exponential in d.

(4) Given n points in $[0, 1]^d$, the problem of finding the largest convex body in $[0, 1]^d$ that contains up to k (outlier) points naturally suggests itself and appears to be also quite interesting.

References

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

[AK01] O. Aichholzer and H. Krasser. The point set order type data base: A collection of applications and results. In *Proc. 13th Canad. Conf. Comput. Geom.*, pages 17–20, 2001.

[BC87] J. Beck and W. Chen. *Irregularities of Distributions*, volume 89 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1987.

[BF87] I. Bárány and Z. Füredi. Empty simplices in Euclidean space. *Canad. Math. Bull.*, 30:436–445, 1987.

[BK10] J. Backer and J. M. Keil. The mono- and bichromatic empty rectangle and square problems in all dimensions. In *Proc. 9th Latin Amer. Symp. Theo. Inf.*, pages 14–25, 2010.

[BP92] I. Bárány and J. Pach. On the number of convex lattice polygons. *Combin. Prob. Comput.*, 1:295–302, 1992.

[BW71] R. P. Bambah and A. C. Woods. On a problem of danzer. *Pacific J. Math.*, 37(2):295–301, 1971.

[CD79] B. Chazelle and D. P. Dobkin. Decomposing a polygon into its convex parts. In *Proc. 11th Annu. ACM Sympos. Theory Comput.*, pages 38–48, 1979.

[CFG91] H. P. Croft, K. J. Falconer, and R. K. Guy. *Unsolved Problems in Geometry*, volume II of *Unsolved Problems in Intuitive Mathematics*. Springer-Verlag, 1991.

[CKM+10] J. Cibulka, J. Kynčl, V. Mészáros, R. Stolař, and P. Valtr. On three parameters of invisibility graphs. In *Proc. 16th Annu. Int. Comput. Comb. Conf.*, pages 192–198, 2010.

[dBCvKO08] M. de Berg, O. Cheong, M. van Kreveld, and M. H. Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, 3rd edition, 2008.

[DJ09] A. Dumitrescu and M. Jiang. On the largest empty axis-parallel box amidst n points. *CoRR*, abs/0909.3127v2, 2009.

[DPV08] S. Dasgupta, C. H. Papadimitriou, and U. V. Vazirani. *Algorithms*. McGraw-Hill, 2008.

[DT11] A. Dumitrescu and Cs. D. Tóth. Minimum weight convex steiner partitions. *Algorithmica*, 60(3):627–652, 2011.

[EORW92] D. Eppstein, M. H. Overmars, G. Rote, and G. Woeginger. Finding minimum area k-gons. *Discrete Comput. Geom.*, 7:45–58, 1992.

[FMR01] T. Fevens, H. Meijer, and D. Rappaport. Minimum convex partition of a constrained point set. *Discrete Appl. Math.*, 109:95–107, April 2001.

[FP92] Z. Füredi and J. Pach. Traces of finite sets: extremal problems and geometric applications. In *Extremal Problems for Finite Sets*, volume 3 of *Bolyai Society Mathematical Studies*, pages 251–282. János Bolyai Mathematical Society, Budapest, 1992.

[GL04] M. Grantson and C. Levcopoulos. A fixed parameter algorithm for the minimum number convex partition problem. In J. Akiyama, M. Kano, and X. Tan, editors, *Japan Conf. Discrete. Comput. Geom.*, volume 3742 of *LNCS*, pages 83–94. Springer, 2004.

[GN06] J. García-López and C. Nicolás. Planar point sets with large minimum convex partitions. In *Proc. 22nd Euro. Workshop on Comput. Geom.*, pages 51–54, 2006.

[Har98] S. Har-Peled. An output sensitive algorithm for discrete convex hulls. *Comput. Geom. Theory Appl.*, 10:125–138, 1998.

[Hor83] J. D. Horton. Sets with no empty convex 7-gons. *Canad. Math. Bull.*, 26:482–484, 1983.

[HS92] J. Hershberger and S. Suri. Applications of a semi-dynamic convex hull algorithm. *BIT*, 32(2):249–267, 1992.

[Kei00] J. M. Keil. Polygon decomposition. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 491–518. Elsevier, 2000.

[Kir80] D. G. Kirkpatrick. A note on Delaunay and optimal triangulations. *Inform. Process. Lett.*, 10(3):127–128, 1980.

[KS02] J. M. Keil and J. Snoeyink. On the time bound for convex decomposition of simple polygons. *Internat. J. Comput. Geom. Appl.*, 12(3):181–192, 2002.

[KS06] C. Knauer and A. Spillner. Approximation algorithms for the minimum convex partition problem. In L. Arge and R. Freivalds, editors, *Proc. 10th Scand. Workshop Algorithm Theory*, volume 4059 of *LNCS*, pages 232–241. Springer, 2006.

[LM09] J. Lawrence and W. Morris. Finite sets as complements of finite unions of convex sets. *Discrete Comput. Geom.*, 42:206–218, May 2009.

[Mat99] J. Matoušek. *Geometric Discrepancy.* Springer, 1999.

[Mat02] J. Matoušek. *Lectures on Discrete Geometry.* Springer, 2002.

[NRU04] V. Neumann-Lara, E. Rivera-Campo, and J. Urrutia. A note on convex decompositions of a set of points in the plane. *Graphs Combin.*, 20(2):223–231, 2004.

[PT11] J. Pach and G. Tardos. Piercing quasi-rectangles: On a problem of danzer and rogers. In F. Dehne, J. Iacono, and J.-R. Sack, editors, *Proc. 12th Workshop Algorithms Data Struct.*, volume 6844 of *LNCS*, page 654. Springer, 2011.

[She92] T. C. Shermer. Recent results in art galleries. *Proc. IEEE*, 80(9):1384–1399, September 1992.

[Spi08] A. Spillner. A fixed parameter algorithm for optimal convex partitions. *J. Discrete Alg.*, 6:561–569, December 2008.

[SU09] T. Sakai and J. Urrutia. Convex decompositions of point sets in the plane. In *Proc. 7th Japan Conf. on Comput. Geom. and Graphs*, 2009.

[Val92] P. Valtr. Sets in $\mathbb{R}^d$ with no large empty convex subsets. *Discrete Appl. Math.*, 108:115–124, 1992.