



Minimum color spanning circle of imprecise points [☆]

Ankush Acharyya ^{a,*}, Ramesh K. Jallu ^{b,1}, Vahideh Keikha ^c, Maarten Löffler ^d,
 Maria Saumell ^{c,e}



^a Dept. of Mathematics and Computing, Indian Institute of Technology (ISM), Dhanbad, India

^b Dept. of Computer Science and Engineering, Indian Institute of Information Technology Raichur, India

^c The Czech Academy of Sciences, Institute of Computer Science, Czech Republic

^d Dept. of Information and Computing Sciences, Utrecht University, the Netherlands

^e Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague, Czech Republic

ARTICLE INFO

Article history:

Received 21 January 2022

Received in revised form 15 June 2022

Accepted 10 July 2022

Available online 18 July 2022

Keywords:

Color spanning circle

Imprecise points

Algorithms

Computational complexity

ABSTRACT

Let \mathcal{R} be a set of n colored imprecise points, where each point is colored by one of k colors. Each imprecise point is specified by a unit disk in which the point lies. We study the problem of computing the smallest and the largest possible minimum color spanning circle, among all possible choices of points inside their corresponding disks. We present an $O(nk \log n)$ time algorithm to compute a smallest minimum color spanning circle. Regarding the largest minimum color spanning circle, we show that the problem is NP-Hard and present a $\frac{1}{3}$ -factor approximation algorithm. We improve the approximation factor to $\frac{1}{2}$ for the case where no two disks of distinct color intersect.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Recognition of color spanning objects of optimum size, in the classical (precise) setting, is a well-studied problem in the literature [2,3,8,16]. The motivation of color spanning problems stems from facility location problems. Here facilities of type $i \in \{1, 2, \dots, k\}$ are modeled as points with color code i , and the objective is to identify the location of a desired geometric shape containing at least one facility of each type such that the desired measure parameter (width, perimeter, area, etc.) is optimized. Other applications of color spanning objects can be found in disk-storage management systems [5] and central-transportation systems [30].

The simplest type of two-dimensional problem considered in this setup is the *minimum color spanning circle* (MCSC) problem, defined as follows. Given a colored point set P in the plane, such that each point in P is colored with one of k possible colors, compute a circle of minimum radius that contains at least one point of each color (see Fig. 1a). The MCSC problem is well-understood: As observed by Abellanas et al. [1], a minimum color spanning circle can be computed in $O(nk \log n)$ time using results on the upper envelope of Voronoi surfaces obtained by Huttenlocher et al. [16].

A limitation of these studies on color spanning objects is that, in many real-life situations, the locations of the points are subject to errors and their exact coordinates are unknown. This is an issue that might arise every time that we try to apply

[☆] A preliminary version of this paper appeared in the Proceedings of the 27th International Conference on Computing and Combinatorics (COCOON 2021).

* Corresponding author.

E-mail addresses: ankush@iitism.ac.in (A. Acharyya), jallu@iiitr.ac.in (R.K. Jallu), keikha@cs.cas.cz (V. Keikha), M.Loffler@uu.nl (M. Löffler), saumell@cs.cas.cz (M. Saumell).

¹ The work was partially done while affiliated with Institute of Computer Science of the Czech Academy of Sciences, with institutional support RVO: 67985807.

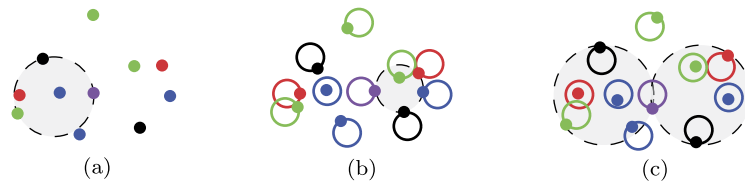


Fig. 1. (a) MCSC for a precise colored point set. (b,c) S-MCSC and L-MCSC for an imprecise colored point set. The representative for each disk is marked as a point of the corresponding color. In (c), we display two L-MCSC of the same radius. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

a geometric algorithm to real-world data. For this reason, in recent years there have been many studies aimed at describing how imprecision in the input data might affect some of the most fundamental geometric descriptors, such as the convex hull or the smallest bounding box. In these studies, the input is an *imprecise* or *uncertain* point set: the location of every point is not uniquely determined, but given by a set of possible locations called its *region* [29,33]. One of the fundamental questions to be studied in this setting is to compute the extreme values of the descriptors, that is, the smallest and biggest, e.g., convex hull of the point set, for all possible locations of every point in its region. In this paper, we consider this problem for the minimum color spanning circle.

1.1. Related work

Uncertainty in data is paramount in contemporary geometric computations. In the literature, different variations have been considered where the regions are modeled as simple geometric objects such as line segments, disks or squares [29,33]. Computing the smallest circle intersecting a set of disks or convex regions of total complexity n is called the *intersection radius problem*, and can be solved in $O(n)$ time [17]. Robert and Toussaint [32] studied the problem of computing the smallest width corridor intersecting a set of convex regions (disks and line segments) and proposed two $O(n \log n)$ time algorithms, where n is the number of convex regions. Löffler and van Kreveld [29] considered the problem of computing the smallest and largest possible axis-parallel bounding box and circle of a set of regions modeled as circles or squares. Their proposed algorithms have running times ranging from $O(n)$ to $O(n \log n)$. For a set of squares or line segments, computing a placement of points in the regions that maximizes or minimizes the area or the perimeter of the convex hull is studied by the same authors [28]. Some variants are shown to be NP-Hard, and the polynomial time algorithms have running times ranging from $O(n \log n)$ to $O(n^{13})$. If the input is a set of disks, a $(1 + \varepsilon)$ -approximation algorithm for this problem is given also by van Kreveld and Löffler [26]. Other problems that have been studied in the region-based model are computing a placement to maximize or minimize the diameter on a set of squares or disks [21,29], or the area of the largest or smallest triangle on a set of line segments [22].

We note that other model formulations have also been proposed for dealing with inaccuracies in geometric problems. These are epsilon geometry [33], probabilistic models [6,34], the aggregated uncertainty model [20], and the domain based models [11].

Regarding other color spanning objects in the precise setting, efficient algorithms are known for computing smallest color spanning squares [1], strips and rectangles [8], 2-intervals [18], equilateral triangles of fixed orientation [15], and axis-parallel squares [23]. Acharyya et al. [2] propose efficient algorithms to compute the narrowest color spanning annulus for circles, axis-parallel squares, rectangles, and equilateral triangles of fixed orientation. The minimum diameter color spanning set problem has also been studied [13,19,35]. Its general version is known to be NP-Hard in the L_p metric, for $1 < p < \infty$, while in the L_1 and L_∞ metrics the problem can be solved in polynomial time [13].

Colored variations of other geometric problems have also been studied in the context of imprecise points [7,9,31]. Given a set of colored clusters, the problem of computing the minimum-weight color spanning tree (*generalized MST problem*) is APX-Hard [9]. Even when each cluster contains exactly two points the problem remains NP-Hard [14]. The problem admits a 2δ -approximation, where δ is the maximum size of the cluster for any imprecise vertex of the MST [31]. In the generalized TSP problem (GTSP), the imprecision is defined by neighborhoods (which are either continuous or discrete) and the goal is to find the shortest tour that visits all neighborhoods. It is known that GTSP with neighborhoods defined by subsets of cardinality two is inapproximable [9].

Finally, we would like to notice that the problem of computing the largest minimum color spanning circle (see the formal description below) is closely related to the dispersion problem in unit disks, where for a given set of n unit disks the goal is to select n points, one from each disk, such that the minimum pairwise distance among the selected points is maximized. This problem was introduced by Fiala et al. [12], and the authors proved that it is NP-Hard. It is also known that the problem is APX-hard [10]. Constant factor approximation algorithms for this problem are given by Cabello [4], and by Dumitrescu and Jiang [10].

1.2. Problem definition and results

In this work, we are given a set $\mathcal{R} = \{R_1, R_2, \dots, R_n\}$ of n unit disks of diameter 1 in the plane, where each disk is colored with one of k possible colors. A colored point set P is a *realization* of \mathcal{R} if there exists a color-preserving bijection

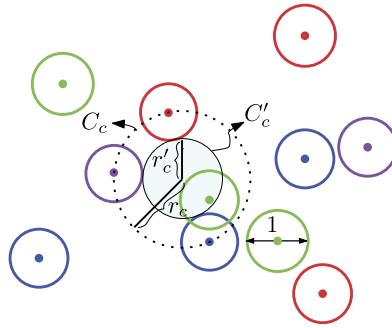


Fig. 2. Illustration of Lemma 1. The dotted circle C_c is the MCSC of the disk centers and the circle C'_c is obtained by decreasing C_c 's radius by $\frac{1}{2}$.

between P and \mathcal{R} such that each point in P is contained in the corresponding disk in \mathcal{R} . Each realization of \mathcal{R} gives a MCSC of certain radius. We are interested in finding realizations of \mathcal{R} such that the corresponding MCSC has the smallest (S-MCSC) and largest (L-MCSC) possible radius (see Fig. 1b and c).

We present the following results:

- The S-MCSC problem can be solved in $O(nk \log n)$ time.
- The L-MCSC problem is NP-Hard.
- A $\frac{1}{3}$ -factor approximation to the L-MCSC problem can be computed in $O(nk \log n)$ time. When no two distinct color disks intersect, the approximation factor becomes $\frac{1}{2}$.

To the best of our knowledge there is no prior result on the minimum color spanning circle problem for imprecise point sets.

2. The smallest MCSC (S-MCSC) problem

Given a set \mathcal{R} of n imprecise points modeled as unit disks, we present an algorithm that finds a S-MCSC, denoted by C_{opt} , and the realization of \mathcal{R} achieving it. Notice that the S-MCSC problem is equivalent to finding a smallest circle that intersects all color regions.

Let $\mathcal{C} = \{c_1, \dots, c_n\}$ be the set of center points of the disks in \mathcal{R} . Let C_c be a MCSC of the colored set \mathcal{C} , and let r_c be its radius. Finally, let r_{opt} be the radius of C_{opt} . The following relation holds:

Lemma 1. *If $r_c > \frac{1}{2}$, then $r_{opt} = r_c - \frac{1}{2}$.*

Proof. Consider a circle C'_c concentric with C_c with radius $r'_c = r_c - \frac{1}{2}$ (see Fig. 2). For every disk R_i such that c_i is contained in C_c , we have that C'_c contains c_i or the intersection between the boundary of R_i and the segment connecting c_i with the center of C_c . Thus, C'_c contains at least one point of each color and $r_{opt} \leq r'_c$.

If $r_{opt} < r'_c$, we would get a feasible solution for the MCSC problem of \mathcal{C} by increasing the radius of C_{opt} by $\frac{1}{2}$. Since such a solution would have radius $r_{opt} + \frac{1}{2} < r'_c + \frac{1}{2} = r_c$, we would get a contradiction with the fact that r_c is the radius of any MCSC of \mathcal{C} . \square

Using Lemma 1, we compute C_{opt} as described in Algorithm 1:

Algorithm 1: Algorithm for the S-MCSC problem

Input: A set \mathcal{R} of n unit disks

Output: A S-MCSC of \mathcal{R} with radius r_{opt}

- 1: compute C_c ;
- 2: **if** $r_c > 1/2$ **then**
- 3: C_{opt} is a circle concentric with C_c , $r_{opt} \leftarrow r_c - \frac{1}{2}$;
- 4: **else**
- 5: C_{opt} is a circle concentric with C_c , $r_{opt} \leftarrow 0$; $\triangleright C_{opt}$ is a point
- 6: **return** C_{opt}

It only remains to prove that Algorithm 1 is correct and efficient:

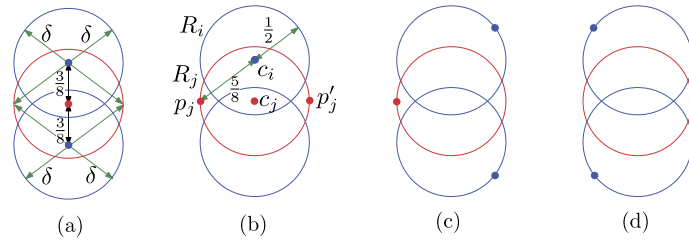


Fig. 3. (a) A stack of disks; (b) the distance from the left endpoint of the red disk to its farthest point in the top blue disk is $\frac{9}{8}$; (c,d) the two placements of points with red-blue distances equal to δ .

Theorem 1. A smallest minimum color spanning circle of \mathcal{R} can be computed in $O(nk \log n)$ time.

Proof. In Algorithm 1, we first compute a MCSC C_c of \mathcal{C} in $O(nk \log n)$ time using the technique proposed by Huttenlocher et al. [16]. If $r_c > \frac{1}{2}$, we shrink the radius of C_c by $\frac{1}{2}$ and return this circle as a solution to the S-MCSC problem. The optimality of the solution follows from Lemma 1. If $r_c \leq \frac{1}{2}$, we have the following property: If c_i is contained in C_c , then the center of C_c is contained in R_i . Hence, the center of C_c lies in the intersection of k distinct colored disks. In consequence, a circle of radius zero concentric with C_c is a solution to the S-MCSC problem. \square

3. The largest MCSC (L-MCSC) problem

In this section, we consider the L-MCSC problem, where for the given set \mathcal{R} the goal is to find a realization such that any MCSC is as large as possible. We show that the problem is NP-Hard, already for $k = 2$, using a reduction from planar 3-SAT [27]. Our reduction is inspired by those described by Fiala et al. [12], and by Knauer et al. [24].

Given a planar 3-SAT instance, we construct a set of colored unit disks with the following property: There exists a realization such that any MCSC has diameter δ if and only if the 3-SAT instance is satisfiable, where $\delta = \frac{9}{8}$. We use disks of colors red and blue. Thus, a point set having any MCSC of diameter at least δ is equivalent to saying that there is no red-blue pair of points at distance less than δ . We denote the family of realizations with this property by \mathcal{P}^δ . With some abuse of notation, sometimes we also use the terminology \mathcal{P}^δ to characterize realizations of (only) a subset of the disks.

We next describe our construction.

A *stack of disks* is a set of three vertically aligned unit disks of alternating colors such that the center of the middle disk is at distance $\frac{3}{8}$ from the centers of the two other disks. See Fig. 3a for an example of a blue-red-blue stack of disks.

Lemma 2. There exist two realizations in \mathcal{P}^δ of a stack of disks.

Proof. Let us denote by R_i and R_j the upper and middle disk of the stack, and let c_i and c_j denote their centers. Additionally, let the leftmost and rightmost points of R_j be denoted by p_j and p'_j (see Fig. 3b). Notice that, among all points which are on the upper half of R_j or on its horizontal diameter, p_j and p'_j are the furthest to c_i .

Let q_i, q_j be two points such that q_i lies in R_i and q_j lies in R_j . Without loss of generality, we assume that q_j lies on the upper half of R_j or on its horizontal diameter (otherwise, we repeat the same arguments taking the lower disk of the stack instead of the upper one). Then we have

$$\begin{aligned} d(q_j, q_i) &\leq d(q_j, c_i) + d(c_i, q_i) \leq d(p_j, c_i) + \frac{1}{2} = \sqrt{(d(p_j, c_j))^2 + (d(c_i, c_j))^2} + \frac{1}{2} \\ &= \frac{5}{8} + \frac{1}{2} = \frac{9}{8}. \end{aligned}$$

Notice that equality is only attained if q_j, c_i and q_i are aligned, and q_j is equal to p_j or p'_j . Thus, there exist only two realizations in \mathcal{P}^δ , shown in Fig. 3c and d. \square

3.1. Variable gadget

Our variable gadget (see Fig. 4) is an alternating chain of red and blue disks, whose centers lie on a hexagon, together with some stacks of disks. The distance between the centers of two consecutive red and blue disks along the same edge of the hexagon is δ . Each edge of the hexagon contains two stacks of disks placed near the endpoints, and every pair of consecutive edges is joined by a blue disk. In the following description, we say that p_i and p'_i are the *leftmost* and *rightmost* points of disk R_i if they are its leftmost and rightmost points after the hexagon has rotated so that the edge containing the center of R_i is horizontal and the center of the hexagon is below the edge.

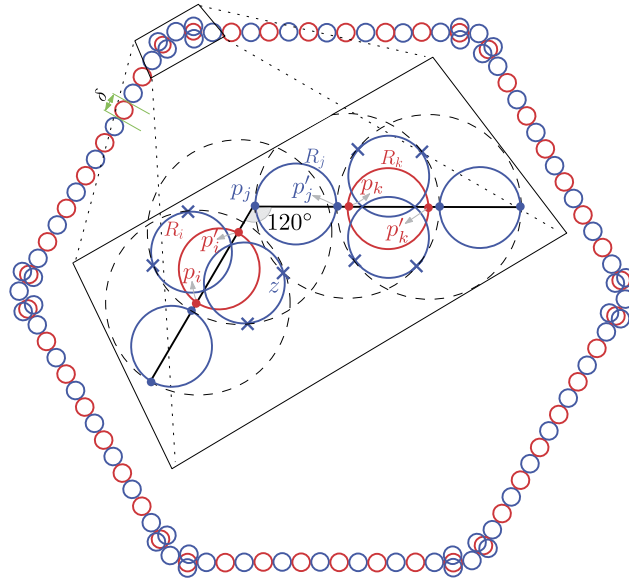


Fig. 4. A variable gadget with zoomed in view for the top-left corner. The dashed circles are centered at p'_k, p_k, p'_j, p_j and have radius δ .

At the top-left corner of the variable gadgets, the disks are placed as follows (the other corners are constructed similarly). Let R_i be the last disk in clockwise order along the top-left edge of the hexagon, and let R_j and R_k be the first and second disks along the top edge (see Fig. 4). The point p_j lies at the top left corner of the hexagon, and the centers of R_j and R_k are at distance δ . Regarding R_i , it is placed in such a way that the lower blue disk of its stack contains a point z which is at distance δ from both p_k and p_j (see Fig. 4). Notice that, if a realization in \mathcal{P}^δ chooses p'_j , the choice for R_j is not unique; however, none of the points in R_j at distance at least δ from p'_j is at distance at least δ from p_k . Therefore, the choice of p'_j forces the choice of p'_k , and clearly the choice of p_k forces the choice of p_j .

For a realization in \mathcal{P}^δ of a variable gadget, the following holds: By Lemma 2, the stack containing R_k is constrained to choose either p_k or p'_k . Let us assume that it chooses p'_k . This choice propagates to the right through the chain of disks in the top edge. The red disk of the stack on the right of the edge also chooses its rightmost point, and this forces the red disk of the first stack of the top right edge to choose its rightmost point too. Therefore, the choice of p'_k propagates through the whole hexagon. If R_k chooses p_k , the same phenomenon occurs. We conclude:

Lemma 3. For any realization in \mathcal{P}^δ of a variable gadget, either all the unit disks centered at the edges of the hexagon, except for the ones intersecting corners of the hexagon, choose their rightmost point, or they all choose their leftmost point.

3.2. Clause gadget

Clause gadgets are illustrated in Fig. 5a-c. We consider an equilateral triangle of side length $\frac{7}{2}$ and center g , and we place one red disk at every corner of the triangle in such a way that the center of the disk is aligned with g and its nearest corner of the triangle. Then we place a blue disk centered at g . Each red disk of a clause gadget is associated to one of the literals occurring in the clause, and is connected to the corresponding variable gadget via a connection gadget. Intuitively, to decide if there exists any realization in \mathcal{P}^δ , each red disk R_τ of the clause gadget has essentially two relevant placements, called t_{R_τ} and f_{R_τ} (see Fig. 5a). As we will see, when the associated literal is set to *true*, we can choose the placement t_{R_τ} , and when it is set to *false*, we are forced to choose f_{R_τ} . It is easy to see that any realization in \mathcal{P}^δ of the clause gadget does not choose f_{R_τ} for at least one of the disks R_τ (see Fig. 5b and c).

3.3. Connection gadget

A variable gadget is connected to each of its corresponding clause gadgets with the help of a *connection gadget*. A connection gadget consists of an alternating chain of red and blue disks together with some stacks of disks (see Fig. 6).

The chain in a connection gadget is formed by straight line portions of groups of disks equal to those along the edges of hexagons associated to variable gadgets. These straight line portions are connected at bends of 120° . The placement of disks at each bend is similar to the placements at the corners of a variable gadget; in particular, stacks of disks are used around the bends.

The precise connection of the variable gadget to the clause gadget by a connection gadget depends on whether the variable in the clause is positive or negative. For a positive variable, the connection is established through a pair of a red

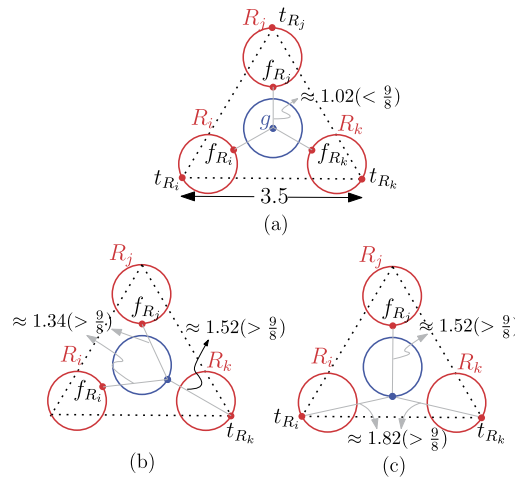


Fig. 5. (a-c) A clause gadget with distinct truth assignments. In the placement in (a), the red-blue pairs are at distance smaller than δ , while in (b) and (c) they are at distance greater than δ .

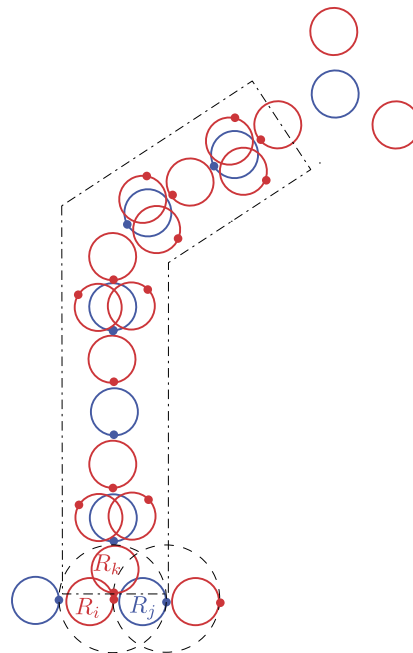


Fig. 6. Connection gadget for a positive variable in a clause; the variable has truth value T .

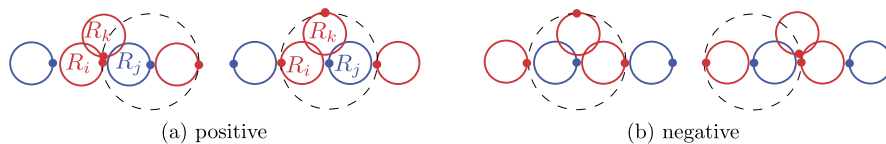


Fig. 7. Point placements corresponding to (a) positive and (b) negative variable in a clause, at the intersection of a connection gadget and a variable gadget. In both figures, the left subfigures correspond to the truth value T for the variable, and the right subfigures to the truth value F .

disk R_i and a blue disk R_j which appear consecutively along an edge of the hexagon, and such that none of them intersects a corner of the hexagon, and R_i comes before R_j in clockwise order (see Fig. 7a). Let R_k be the first red disk of the connection gadget. The top-most point of R_k is at distance δ from p_j , and its bottom-most point is at distance δ from p'_j . For a negative variable, the connection is established through a pair of blue-red disks in the variable gadget. The placement of the first red disk of the connection gadget is analogous to the one in the red-blue configuration (see Fig. 7b).

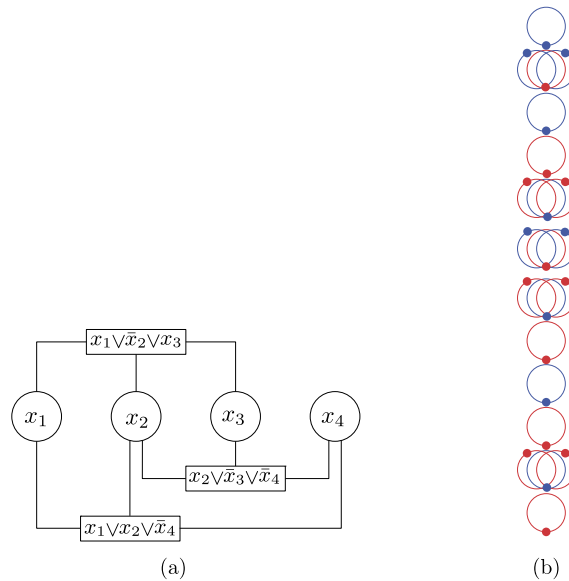


Fig. 8. (a) Planar embedding of $\mathcal{F} = (x_1 \vee x_2 \vee \bar{x}_4) \wedge (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_2 \vee \bar{x}_3 \vee \bar{x}_4)$. (b) A vertical portion of a connection gadget with increased gaps among three central disks, which have been replaced by stacks.

To ensure the desired propagation of the placement of points in the disks, stacks of disks are used next to the first disk R_k , and next to the red disk of the clause gadget associated to the literal (see Fig. 6).

The truth value T of a variable is associated with the choice of the rightmost points of the disks in the variable gadget. If the variable appears positive at a clause, this allows (for a realization in \mathcal{P}^δ) the choice of the bottom-most point of R_k , and this propagates through the connection gadget and eventually allows the choice of the associated point t_{R_τ} in the clause gadget (see Fig. 6). If the truth value is F , p_j is selected, which (for a realization in \mathcal{P}^δ) forces the choice of a point in a close vicinity of the top-most point of R_k . Since there is a stack next to R_k , the blue disk in the stack chooses the top-most point, and this eventually forces the choice of f_{R_τ} . The analysis of the other cases are similar.

3.4. Final construction

In the following lemma, we put all pieces of the construction together, and we prove that it has polynomial size.

Lemma 4. *Using the gadgets described above, we can construct in polynomial time a polynomial-size instance of the L-MCSC problem associated to the given planar 3-SAT formula.*

Proof. The dependency graph of the 3-SAT formula can be embedded so that all variables lie on a horizontal line, all clauses are on either side of them, and each edge connecting a clause to a variable is an orthogonal edge with at most one bend (see Fig. 8a for an example) [25]. We take an embedding with these properties which, additionally, is aligned with a grid of a resolution which is high enough to prevent interference once we place our gadgets; for the sake of simplicity, we choose resolution 20 (i.e., the distance between any pair of consecutive horizontal or vertical segments is at least 20). In the following lines, we explain how to modify this embedding to produce an instance of the L-MCSC problem. Our construction is illustrated in Fig. 9.

First, we replace the variables of the embedding by gadgets as described above. The size of a variable gadget is chosen so that the upper and lower side of the hexagon are long enough to fit all edges leading to clauses. Since any two such edges are at distance at least 20 and we place them at distance at least 20 from the extremes of the edges of the hexagon, the side of the hexagon is no longer than $20(m + 1)$, where m is the number of clauses in the formula. During this step, we might again need to horizontally “stretch” the initial embedding so that variable gadgets end up at distance at least 20 from each other.

Next, we replace the clauses of the embedding by clause gadgets. We explain in detail the case where the clause is above the variables; the other case is solved symmetrically (in particular, clause gadgets lying below the variables will be rotated by 180°). Initially, the clause is incident to three edges, which we call the “left”, “middle” and “right” edges. Each of them is assigned to one of the red disks of the clause gadgets, i.e., the connection between the variable gadget and the connection gadget will be done through that disk. In particular, the left edge is assigned to the left-most red disk of the gadget, the middle edge is assigned to the right-most red disk of the gadget, and the right edge is assigned to the top red disk of the gadget (see Fig. 9).

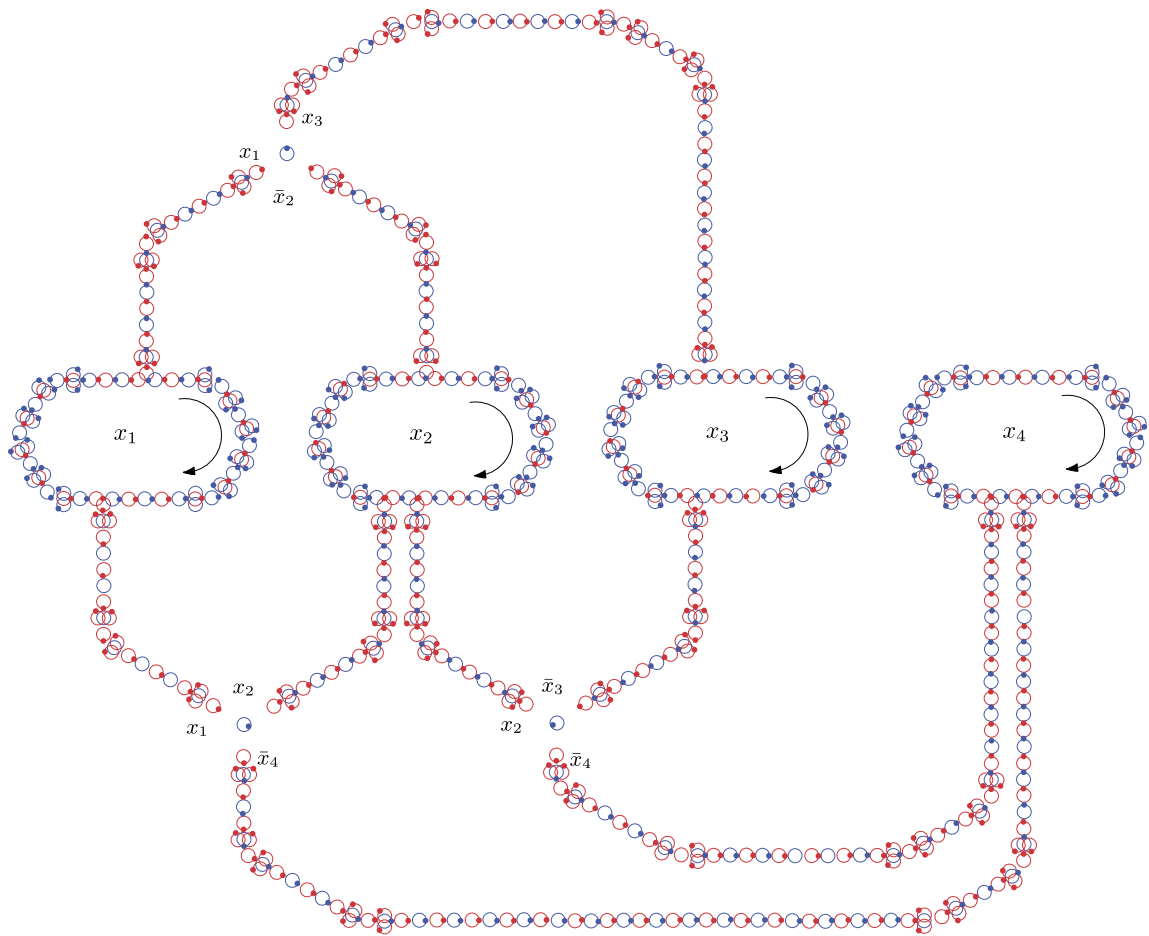


Fig. 9. Instance of the L-MCSC problem associated to the formula \mathcal{F} in Fig. 8a. The proportions among the sizes of the different objects have been altered a little, and the hexagons associated to the variables are not drawn regular. The placement of points corresponds to the satisfying assignment $x_1 = F, x_2 = T, x_3 = T, x_4 = F$.

Finally, we replace every edge connecting a clause to a variable gadget in the following way: The position of the first disk of the connection gadget (that is, the nearest to the clause gadget) is given by the previous assignment of edges to red disks of the clause. The position of the last disk (i.e., the disk of the connection gadget which intersects the variable gadget) is as near as possible to the original location of the edge, while respecting the requirements for each type of connection (which depend on whether the variable appears positive or negative in the clause; refer to Fig. 7). Since the left and middle edges become connection gadgets with one bend of 120° (see Fig. 9), the precise shape of these gadgets (including the position of the bend) is determined by their first and last disks. Notice also that, since the middle edge (which is vertical) is replaced by a chain of shape “ \backslash ”, it might be necessary to move the clause gadget slightly towards the left. Regarding the right edge, it is replaced by a chain of shape “ \lceil ” starting from the top red disk of the clause variable. In this case, there is some freedom to choose the positions of the bends, and we place them so that the final connection gadget is close enough to the original replaced edge.

After replacing orthogonal edges by connection gadgets with bends of 120° , some pairs of connection gadgets might have ended up too near (or even cross) and might interfere with each other. This can again be solved by stretching the embedding to separate them.

We finally point out that the straight portions of our connection gadgets are made of chains of unit disks at a certain fixed distance, and thus they can only have certain prescribed lengths of $\frac{9k}{8} + 1$ for some integer k , which we call *integer lengths* (e.g., for vertical portions of a connection gadget, the distance between the bottom-most and top-most points of the chain can only take values of the form $\frac{9k}{8} + 1$, for some integer k). Therefore, it might happen that the positions of a clause and variable gadget force the placement of a chain which does not have integer length. In this case, we proceed as follows: Suppose that we need a vertical chain such that the positions and colors of the first and last disk are prescribed, and either the distance l between the bottom-most and top-most points of the chain does not correspond to an integer length, or it does but it has the wrong parity (e.g., when we try to build the chain, the bottom-most disk turns up blue, but we would need it to be red). We take the longest chain of integer length having length smaller than l and having the

desired combination of colors of the first and last disk. If this chain has length $\frac{9k}{8} + 1$, for some integer k , then we have that $l < \frac{9(k+2)}{8} + 1$. Thus, the difference in lengths is at most $\frac{9}{4}$. To solve this discrepancy, we select two to four consecutive disks from the middle of the chain and increase the gap between them from $\frac{1}{8}$ to some appropriate value smaller than or equal to $\frac{7}{8}$ (the remaining gaps between consecutive disks of the chain remain of $\frac{1}{8}$). By replacing these two to four disks by stacks, we ensure the required propagation among them despite the increased gaps. See Fig. 8b for an example. Since we might increase up to three gaps by up to $\frac{3}{4}$, we can achieve the difference in lengths of at most $\frac{9}{4}$.

Since the final construction has a polynomial number of disks and all the steps in the construction can be performed in polynomial time, the lemma follows. \square

Given a planar 3-SAT formula \mathcal{F} , let us denote by $\mathcal{R}^{\mathcal{F}}$ an associated instance of the L-MCSC problem constructed following the steps in the proof of Lemma 4. We prove that $\mathcal{R}^{\mathcal{F}}$ has the desired property:

Lemma 5. *The planar 3-SAT formula \mathcal{F} has a satisfying assignment if and only if there exists a realization of $\mathcal{R}^{\mathcal{F}}$ in \mathcal{P}^δ .*

Proof. Suppose that \mathcal{F} has n variables x_1, x_2, \dots, x_n and m clauses C_1, C_2, \dots, C_m .

Let us first assume that \mathcal{F} is satisfiable. Thus, every clause contains a literal whose truth value is T . We describe a realization of $\mathcal{R}^{\mathcal{F}}$ in \mathcal{P}^δ .

Let C_j be a clause and let ℓ_i be a literal making the clause true. Then we pick t_{R_i} from the associated disk R_i in the clause. This choice propagates through the connection gadget and eventually forces one of the realizations for the variable gadget described in Lemma 3: the one choosing the rightmost points of the disks, if ℓ_i 's truth value is T and it appears in the clause positively; or the one choosing the leftmost points of the disks, if ℓ_i 's truth value is F and it appears in the clause negatively. Notice that a variable might make true several clauses, but the obtained realizations of the variable gadget are consistent with each other.

Let us consider a variable whose realization has been fixed in the paragraph above. If its truth value is T and it appears in a clause negatively, or if its truth value is F and it appears in a clause positively, the placement described above forces the realizations of the corresponding connection gadget and associated disk in the clause gadget.

For the variables whose realization has not been fixed yet (if any), we pick, say, the realization choosing the leftmost points of the disks. For their connection gadgets to the clauses, we pick the realization choosing the top-most (respectively, bottom-most) points of the disks, if the clause is above (respectively, below) the variable. The described realizations belong to \mathcal{P}^δ .

After following the steps above, for each clause gadget at least one of the three red disks has its representative at the position t_{R_τ} . In consequence, it is possible to find a realization of the blue disk such that the realization of the clause gadget is in \mathcal{P}^δ .

Notice that we have described realizations for all gadgets in the construction. Clearly, the global realization is in \mathcal{P}^δ .

We next prove the second implication. Suppose that there exists a realization R of $\mathcal{R}^{\mathcal{F}}$ in \mathcal{P}^δ . Let C_j be a clause. Since the realization of the clause is in \mathcal{P}^δ , at least one of the red disks R_i of the clause does not have the representative at the position f_{R_i} . If the corresponding literal appears positive in the clause, we set the associated variable to T . Otherwise, we set it to F . In this way, we ensure that every clause is true, but we need to argue that we did not assign T and F simultaneously to the same variable due to two distinct clauses.

Since in the connection gadget there is a stack of disks next to R_i , the fact that R_i does not have the representative at the position f_{R_i} forces the choice of the representative of the blue disk in the stack. The choice propagates to the connection gadget and eventually to the variable gadget: If the literal associated to R_i appears positive in C_j , the variable has been set to T and the realization in the variable gadget is the one choosing the rightmost points of the disks. If the literal appears negative, the variable has been set to F and the realization in the variable gadget chooses the leftmost points of the disks. Since in R the realization of the variable gadget is either the rightmost or the leftmost, the variable has either been set to T or to F . \square

Lemmas 4 and 5 imply the following:

Theorem 2. *The problem of finding a largest minimum color spanning circle of \mathcal{R} is NP-Hard.*

Remark 1. Given a yes-instance, we can verify in $O(nk \log n)$ time whether the given realization is correct and the radius of its MCSC is at least δ . Therefore, the decision version of the problem is NP-Complete.

4. Approximation algorithms

Given that the L-MCSC problem is NP-Hard, in this section we turn our attention to approximation algorithms.

Let \tilde{r}_{opt} denote the radius of a largest possible minimum color spanning circle of \mathcal{R} . We first prove bounds on \tilde{r}_{opt} .

Lemma 6. $\tilde{r}_{opt} \geq 1/4$.

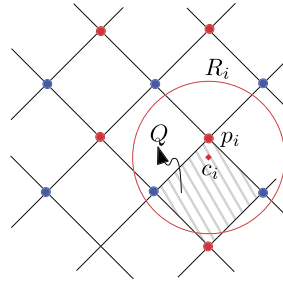


Fig. 10. The tilted grid in the proof of Lemma 6. We choose p_i as the red corner of Q contained in R_i .

Proof. First, we show that it is enough to prove the result for the case where $k = 2$: Suppose that $k > 2$. We select two of the k colors, say red and blue, and the set $\mathcal{R}' \subseteq \mathcal{R}$ of red and blue disks. Suppose that the radius of any L-MCSC of \mathcal{R}' is greater than or equal to $1/4$. Then there exists a realization of \mathcal{R}' whose MCSC has radius greater than or equal to $1/4$. We complete this realization of \mathcal{R}' to a realization of \mathcal{R} by choosing any point in each disk of the remaining $k - 2$ colors. Since any MCSC of this realization contains at least one red and one blue point, its radius is greater than or equal to $1/4$.

Therefore, we only show that the bound holds when $k = 2$ by providing a realization P whose MCSC achieves the bound. Consider a regular square grid rotated by $\pi/4$ such that the side of every cell of the grid has length $1/2$. We color the corners of the cells in red or blue in such a way that all corners lying in some vertical line are colored red, all corners lying in the next vertical line are colored blue, and so on (see Fig. 10). Now let $R_i \in \mathcal{R}$ have red color, and let Q be the cell of the grid containing the center of R_i (if the center of R_i lies on an edge or vertex of the grid, we assign it to any of the adjacent cells). Notice that at least one of the two red corners of Q lies inside R_i . We choose such a corner as $p_i \in P$. Similarly, for every $R_j \in \mathcal{R}$ of blue color, P contains one of the blue corners of a cell containing the center of R_j . We obtain that P is a subset of the grid corners. Since the distance between any pair of red and blue corners is at least $1/2$, the radius of any MCSC is at least $1/4$. \square

Next, we establish a relation between \tilde{r}_{opt} and r_c . We recall that r_c is the radius of a MCSC (denoted C_c) of the colored set \mathcal{C} containing the center points of the disks in \mathcal{R} .

Lemma 7. $\tilde{r}_{opt} \leq r_c + \frac{1}{2}$.

Proof. Without loss of generality we can assume that C_c contains the centers of the disks in $\mathcal{R}' = \{R_1, R_2, \dots, R_k\}$. A circle of radius $r_c + \frac{1}{2}$, concentric with C_c , contains the k disks in \mathcal{R}' . Therefore, for any placement of points inside disks of \mathcal{R}' , we can always get a color spanning circle with radius at most $r_c + \frac{1}{2}$. Thus, $\tilde{r}_{opt} \leq r_c + \frac{1}{2}$. \square

We use the bounds above to design a simple approximation algorithm, presented in Algorithm 2. Let P^g denote the realization of \mathcal{R} described in the proof of Lemma 6.

Algorithm 2: $\frac{1}{3}$ -factor approximation algorithm for the L-MCSC problem

Input: A set \mathcal{R} of n unit disks
Output: A MCSC of a realization of \mathcal{R} with radius at least $\tilde{r}_{opt}/3$

- 1: compute C_c ;
- 2: **if** $r_c \geq 1/4$ **then**
- 3: **return** C_c ;
- 4: **else**
- 5: **return** a MCSC of P^g ;

It only remains to prove that Algorithm 2 indeed gives a $\frac{1}{3}$ -factor approximation:

Theorem 3. A $\frac{1}{3}$ -factor approximation for the L-MCSC problem can be computed in $O(nk \log n)$ time. If no two distinct colored unit disks of \mathcal{R} intersect, the approximation factor becomes $\frac{1}{2}$.

Proof. If $r_c \geq 1/4$, then, by Lemma 7, $\tilde{r}_{opt} \leq 3r_c$ and C_c gives a $\frac{1}{3}$ -approximation for the problem. If $r_c < 1/4$, then, by Lemma 7, $\tilde{r}_{opt} < 3/4$. By Lemma 6, any MCSC of P^g has radius at least $1/4$, so such a circle gives a $\frac{1}{3}$ -approximation for the problem.

Computing P^g takes $O(n)$ time, and computing a MCSC of P^g or the set of disk centers can be done in $O(nk \log n)$ time.

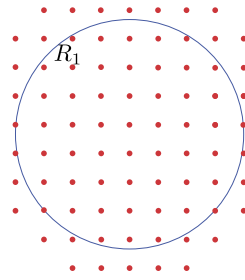


Fig. 11. Construction in the proof of Remark 2. For the sake of clarity, the red disks centered at the grid points are not displayed.

Finally, if no two distinct colored disks intersect, $r_c \geq \frac{1}{2}$. This, combined with Lemma 7, gives $r_c \geq \frac{\tilde{r}_{opt}}{2}$. \square

Since Lemma 6 is one of the key ingredients of our approximation algorithm, we conclude this section by showing that it is the best possible.

Remark 2. Given any $\varepsilon > 0$, there exists a set \mathcal{R} of unit disks for which $\tilde{r}_{opt} < 1/4 + \varepsilon$.

Proof. We start with a blue disk R_1 . We overlay a square grid of red points over the area covered by R_1 enlarged a little around the boundary (see Fig. 11). The length of the sides of the grid cells is set to $2\sqrt{2}\varepsilon$. Then, for every red point of the grid, we place a red disk centered at that point. Finally, we place the remaining blue disks far away from the construction.

Let p_1 be the representative for R_1 in a realization P giving a minimum color spanning circle of radius \tilde{r}_{opt} . Since p_1 lies inside some cell of the grid, there exists some red grid point q such that $d(p_1, q) \leq 2\varepsilon$. Let R_j be the red disk centered at q , and p_j be its representative in the realization P . Then, $d(p_1, p_j) \leq d(p_1, q) + d(q, p_j) \leq 2\varepsilon + 1/2$. Since p_1, p_j is a blue-red pair at distance at most $1/2 + 2\varepsilon$, the radius of any minimum color spanning circle is at most $1/4 + \varepsilon$. \square

5. Open problems

Naturally, the main open problems are related to the L-MCSC problem and the existence of better approximation algorithms for it. So far we have not succeeded in finding a PTAS for this problem, and we suspect that the problem might be APX-Hard. If this is the case, it would also be interesting to improve the approximation factor of our approximation algorithm.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ankush Acharyya, Ramesh Jallu, Vahideh Keikha, Maria Saumell reports financial support was provided by Czech Science Foundation and institutional support from the Institute of Computer Science of the Czech Academy of Sciences, Czechia. Maarten Loffler reports financial support was provided by Netherlands Organisation for Scientific Research Division Humanities.

Acknowledgements

The authors would like to thank Irina Kostitsyna for key discussions on the hardness reduction and Hans Raj Tiwary for the proof of Lemma 6. A.A., R.J., V.K., and M.S. were supported by the Czech Science Foundation, grant number GJ19-06792Y, and by institutional support RVO: 67985807. M.L. was partially supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 614.001.504. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 734922.

References

- [1] M. Abellanas, F. Hurtado, C. Icking, R. Klein, E. Langetepe, L. Ma, B. Palop, V. Sacristán, Smallest color-spanning objects, in: F.M. auf der Heide (Ed.), Algorithms - ESA 2001, 9th Annual European Symposium, Proceedings, Aarhus, Denmark, August 28–31, 2001, in: Lecture Notes in Computer Science, vol. 2161, Springer, 2001, pp. 278–289.
- [2] A. Acharyya, S.C. Nandy, S. Roy, Minimum width color spanning annulus, Theor. Comput. Sci. 725 (2018) 16–30, <https://doi.org/10.1016/j.tcs.2017.11.031>.
- [3] M. de Berg, J. Gudmundsson, M.J. Katz, C. Levcopoulos, M.H. Overmars, A.F. van der Stappen, TSP with neighborhoods of varying size, J. Algorithms 57 (1) (2005) 22–36, <https://doi.org/10.1016/j.jalgor.2005.01.010>.
- [4] S. Cabello, Approximation algorithms for spreading points, J. Algorithms 62 (2) (2007) 49–73, <https://doi.org/10.1016/j.jalgor.2004.06.009>.
- [5] M.E. Consuegra, G. Narasimhan, Geometric avatar problems, in: A. Seth, N.K. Vishnoi (Eds.), IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2013, December 12–14, 2013, Guwahati, India, in: LIPIcs, vol. 24, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2013, pp. 389–400.

- [6] G. Cormode, A. McGregor, Approximation algorithms for clustering uncertain data, in: M. Lenzerini, D. Lembo (Eds.), Proceedings of the Twenty-Seventh ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2008, June 9–11, 2008, Vancouver, BC, Canada, ACM, 2008, pp. 191–200.
- [7] O. Daescu, W. Ju, J. Luo, NP-completeness of spreading colored points, in: W. Wu, O. Daescu (Eds.), Combinatorial Optimization and Applications - 4th International Conference, COCOA 2010, Proceedings, Part I, Kailua-Kona, HI, USA, December 18–20, 2010, in: Lecture Notes in Computer Science, vol. 6508, Springer, 2010, pp. 41–50.
- [8] S. Das, P.P. Goswami, S.C. Nandy, Smallest color-spanning object revisited, *Int. J. Comput. Geom. Appl.* 19 (5) (2009) 457–478, <https://doi.org/10.1142/S0218195909003076>.
- [9] M. Dror, J.B. Orlin, Combinatorial optimization with explicit delineation of the ground set by a collection of subsets, *SIAM J. Discrete Math.* 21 (4) (2008) 1019–1034, <https://doi.org/10.1137/050636589>.
- [10] A. Dumitrescu, M. Jiang, Dispersion in disks, *Theory Comput. Syst.* 51 (2) (2012) 125–142, <https://doi.org/10.1007/s00224-011-9331-x>.
- [11] A. Edalat, A. Lieutier, E. Kashefi, The convex hull in a new model of computation, in: Proceedings of the 13th Canadian Conference on Computational Geometry, CCCG, Waterloo, Canada, August 13–15, 2001, 2001, pp. 93–96.
- [12] J. Fiala, J. Kratochvíl, A. Proskurowski, Systems of distant representatives, *Discrete Appl. Math.* 145 (2) (2005) 306–316, <https://doi.org/10.1016/j.dam.2004.02.018>.
- [13] R. Fleischer, X. Xu, Computing minimum diameter color-spanning sets is hard, *Inf. Process. Lett.* 111 (21–22) (2011) 1054–1056, <https://doi.org/10.1016/j.ipl.2011.07.015>.
- [14] R. Fraser, Algorithms for geometric covering and piercing problems, Ph.D. thesis, University of Waterloo, 2013.
- [15] J. Hasheminejad, P. Khanteimouri, A. Mohades, Computing the smallest color spanning equilateral triangle, in: Proceedings of 31st European Workshop on Computational Geometry, EuroCG, Ljubljana, Slovenia, March 15–18, 2015, 2015, pp. 32–35.
- [16] D.P. Huttenlocher, K. Kedem, M. Sharir, The upper envelope of Voronoi surfaces and its applications, *Discrete Comput. Geom.* 9 (1993) 267–291, <https://doi.org/10.1007/BF02189323>.
- [17] S. Jadhav, A. Mukhopadhyay, B.K. Bhattacharya, An optimal algorithm for the intersection radius of a set of convex polygons, *J. Algorithms* 20 (2) (1996) 244–267, <https://doi.org/10.1006/jagm.1996.0013>.
- [18] M. Jiang, H. Wang, Shortest color-spanning intervals, *Theor. Comput. Sci.* 609 (2016) 561–568, <https://doi.org/10.1016/j.tcs.2015.01.039>.
- [19] W. Ju, C. Fan, J. Luo, B. Zhu, O. Daescu, On some geometric problems of color-spanning sets, *J. Comb. Optim.* 26 (2) (2013) 266–283, <https://doi.org/10.1007/s10878-012-9458-y>.
- [20] V. Keikha, S. Aghamolaei, A. Mohades, M. Ghodsi, Clustering geometrically-modeled points in the aggregated uncertainty model, *Fundam. Inform.* 184 (3) (2021) 205–231, <https://doi.org/10.3233/FI-2021-2097>.
- [21] V. Keikha, M. Löffler, A. Mohades, A fully polynomial time approximation scheme for the smallest diameter of imprecise points, *Theor. Comput. Sci.* 814 (2020) 259–270, <https://doi.org/10.1016/j.tcs.2020.02.006>.
- [22] V. Keikha, M. Löffler, A. Mohades, Largest and smallest area triangles on imprecise points, *Comput. Geom.* 95 (2021) 101742, <https://doi.org/10.1016/j.comgeo.2020.101742>.
- [23] P. Khanteimouri, A. Mohades, M.A. Abam, M.R. Kazemi, Computing the smallest color-spanning axis-parallel square, in: L. Cai, S. Cheng, T.W. Lam (Eds.), Algorithms and Computation - 24th International Symposium, ISAAC 2013, Proceedings, Hong Kong, China, December 16–18, 2013, in: Lecture Notes in Computer Science, vol. 8283, Springer, 2013, pp. 634–643.
- [24] C. Knauer, M. Löffler, M. Scherfenberg, T. Wolle, The directed Hausdorff distance between imprecise point sets, *Theor. Comput. Sci.* 412 (32) (2011) 4173–4186, <https://doi.org/10.1016/j.tcs.2011.01.039>.
- [25] D.E. Knuth, A. Raghunathan, The problem of compatible representatives, *SIAM J. Discrete Math.* 5 (3) (1992) 422–427, <https://doi.org/10.1137/0405033>.
- [26] M. van Kreveld, M. Löffler, Approximating largest convex hulls for imprecise points, *J. Discret. Algorithms* 6 (4) (2008) 583–594, <https://doi.org/10.1016/j.jda.2008.04.002>.
- [27] D. Lichtenstein, Planar formulae and their uses, *SIAM J. Comput.* 11 (2) (1982) 329–343, <https://doi.org/10.1137/0211025>.
- [28] M. Löffler, M. van Kreveld, Largest and smallest convex hulls for imprecise points, *Algorithmica* 56 (2) (2010) 235–269, <https://doi.org/10.1007/s00453-008-9174-2>.
- [29] M. Löffler, M. van Kreveld, Largest bounding box, smallest diameter, and related problems on imprecise points, *Comput. Geom.* 43 (4) (2010) 419–433, <https://doi.org/10.1016/j.comgeo.2009.03.007>.
- [30] R. Manzini, R. Gamberini, Design, management and control of logistic distribution systems, *Int. J. Adv. Robot. Syst.* (2008) 263–290.
- [31] P.C. Pop, The generalized minimum spanning tree problem: an overview of formulations, solution procedures and latest advances, *Eur. J. Oper. Res.* 283 (1) (2020) 1–15, <https://doi.org/10.1016/j.ejor.2019.05.017>.
- [32] J.M. Robert, G. Toussaint, Computational geometry and facility location, in: Proceedings of International Conference on Operations Research and Management Science, 1990, pp. 11–15.
- [33] D. Salesin, J. Stolfi, L.J. Guibas, Epsilon geometry: building robust algorithms from imprecise computations, in: K. Mehlhorn (Ed.), Proceedings of the Fifth Annual Symposium on Computational Geometry, Saarbrücken, Germany, June 5–7, 1989, ACM, 1989, pp. 208–217.
- [34] S. Suri, K. Verbeek, H. Yildiz, On the most likely convex hull of uncertain points, in: H.L. Bodlaender, G.F. Italiano (Eds.), Algorithms - ESA 2013 - 21st Annual European Symposium, Proceedings, Sophia Antipolis, France, September 2–4, 2013, in: Lecture Notes in Computer Science, vol. 8125, Springer, 2013, pp. 791–802.
- [35] D. Zhang, Y.M. Chee, A. Mondal, A.K.H. Tung, M. Kitsuregawa, Keyword search in spatial databases: towards searching by document, in: Y.E. Ioannidis, D.L. Lee, R.T. Ng (Eds.), Proceedings of the 25th International Conference on Data Engineering, ICDE 2009, March 29–April 2, 2009, Shanghai, China, IEEE Computer Society, 2009, pp. 688–699.