Linear-Time Approximation Algorithms for Unit Disk Graphs

Guilherme D. da Fonseca
Vinícius G. Pereira de Sá
Celina M. H. de Figueiredo

Universidade Federal do Rio de Janeiro, Brazil

Abstract

Numerous approximation algorithms for unit disk graphs have been proposed in the literature, exhibiting sharp trade-offs between running times and approximation ratios. We propose a method to obtain linear-time approximation algorithms for unit disk graph problems. Our method yields linear-time $(4 + \varepsilon)$-approximations to the maximum-weight independent set and the minimum dominating set, as well as a linear-time approximation scheme for the minimum vertex cover, improving upon all known linear- or near-linear-time algorithms for these problems.

1 Introduction

A *unit disk graph* is the intersection graph of unit disks in the plane. Unit disk graphs are often represented using the coordinates of the disk centers instead of explicit adjacency information. In this geometric setting, two vertices are adjacent if the corresponding points (the disk centers) are within Euclidean distance at most 2 from one another.

Owing to their applicability in wireless networks [10, 13], numerous approximation algorithms for unit disk graphs have been proposed in the literature. Such approximations are either *graph-based* algorithms, when they receive as input solely the adjacency representation of the graph, or *geometric* algorithms, when the input consists of a geometric representation of the graph. While the edges of a graph $G(V, E)$, with $n = |V|$ and $m = |E|$, can be obtained from the vertices' coordinates in $O(n + m)$ time under the real-RAM model with floor function and constant-time hashing [3], obtaining a geometric representation of a given unit disk graph is NP-hard [4]. We note that, when the goal is to design $O(n)$-time algorithms, the geometric representation is required, since the number m of edges in a unit disk graph can be as high as $\Theta(n^2)$.

The *shifting strategy* gave rise to geometric PTASs for several problems for unit disk graphs [5, 8]. Essentially, the shifting strategy reduces the original problem to a set of subproblems of constant diameter. Such reduction takes $O(n)$ time and yields a $(1 + \varepsilon)$-approximation to the original problem, given the exact solutions to the subproblems. However, the running times of the PTASs are polynomials of high degree because each subproblem is solved exactly by exploiting the fact that the point set has constant diameter. For example, we can show by packing arguments that a set of diameter d has a dominating set with $c = O(d^2)$ vertices, and, consequently, exhaustive enumeration finds the minimum dominating set in roughly $O(n^c)$ time. Graph-based PTASs for these problems are also known [13]. While they do not use the shifting strategy, their running times are even higher than those of their geometric counterparts.

The minimum dominating set problem (MDSP) admits some PTASs [8, 13], the fastest of which is geometric and provides a $(1+\varepsilon)$-approximation in $n^{O(1/\varepsilon^2)}$ time and a 4-approximation in roughly $O(n^{10})$ time. Such high running times have motivated the study of faster constant-factor approximation algorithms. Examples of graph-based algorithms include a 44/9-approximation that runs in $O(n+m)$ time and a 43/9-approximation that runs in $O(n^2 m)$ time [6]. Among the geometric algorithms, we cite the original 5-approximation, which can be implemented in $O(n)$ time if the floor function and constant-time hashing are available [10]; a 44/9-approximation that uses local improvements and runs in $O(n \log n)$ time [6]; a 4-approximation that uses grids and runs in $O(n^8 \log n)$ time [7]; and a recent 4-approximation that uses hexagonal grids and runs in $O(n^6 \log n)$ time [9].

The maximum-weight independent set problem (MWISP) also admits some PTASs, the fastest of which attains a $(1+\varepsilon)$-approximation in $O(n^{4\lceil 2/\varepsilon\sqrt{3}\rceil})$ time, which gives $O(n^4)$ time for a 4-approximation [8, 11, 13]. A 5-approximation can be obtained in $O(n \log n)$ time by a greedy approach that considers the vertices in decreasing order of weights. However, no linear-time constant-approximation algorithm is known for this problem. In contrast, for the unweighted version, a simple greedy approach gives a 5-approximation in $O(n)$ time, while a greedy approach that considers the vertices from left to right gives a 3-approximation in $O(n \log n)$ time [10].

The minimum vertex cover problem (MVCP) admits a geometric PTAS [8], which attains a $(1+\varepsilon)$-approximation in $n^{O(1/\varepsilon)}$ time, but no graph-based PTAS is known. Also, a 1.5-approximation algorithm is presented in [10]. Its time complexity is dominated by that of the Nemhauser-Trotter decomposition [12], which can be implemented in $O(m\sqrt{n})$ time.

Our results. We introduce a novel method to obtain linear-time approximation algorithms for problems on unit disk graphs and other geometric intersection graphs (Section 2). By using our method, we obtain linear-time $(4+\varepsilon)$-approximation algorithms for the MWISP (Section 3) and the MDSP (Section 4). For the MVCP, an indirect application of our method yields a linear-time $(1+\varepsilon)$-approximation (Section 5). The proposed algorithms improve upon all known linear-time (or close to linear-time) approximations for these problems. Although our algorithms share the same basic idea, their analyses differ significantly. For example, the MWISP analysis applies the Four-Color Theorem for planar graphs [2], while the MDSP analysis applies packing arguments. We conclude with lower bounds to the approximation ratios of our algorithms and open problems (Section 6).

2 Our Method

The shifting strategy is the main idea behind the existing geometric PTASs for problems on unit disk graphs such as the minimum dominating set, maximum independent set, minimum vertex cover, and minimum connected dominating set problems [5, 8]. Generally, the shifting strategy reduces the original problem with n points to a set of subproblems whose inputs have constant diameter and the sum of the input sizes is $O(n)$. Such reduction is based on partitioning the points according to a number of iteratively shifted grids and takes $O(n)$ time (by using the floor function and constant-time hashing). Exploiting the inputs' constant diameter, each subproblem is solved exactly in polynomial time. The solutions to the subproblems are then combined appropriately (normally in $O(n)$ time) to yield feasible solutions to the original problem, the best of which is returned. The high complexities of these geometric PTASs are due to the exact algorithms that are employed to solve each subproblem.

We propose a method that is based on the shifting strategy. It presents, however, a crucial difference. Rather than obtaining exact, costly solutions for the subproblems, we solve each subproblem *approximately.* To do that, we employ the coresets paradigm [1], where only a subset with a constant number of input points is considered.

For a problem whose input is a set P of n points, our method can be briefly described as follows:

1. Apply the shifting strategy to construct a set of r subproblems with inputs $P_1, \ldots, P_r$ such that $\sum_{i=1}^{r} |P_i| = O(n)$ and $\text{diam}(P_i) = O(1)$ for all i.

2. For each subproblem instance P_i, obtain a coreset $Q_i \subseteq P_i$ with $|Q_i| = O(1)$, such that the optimal solution for instance Q_i is an α-approximation to the optimal solution for instance P_i.

3. Solve the problem exactly for each Q_i.

4. Combine the solutions into an $(\alpha + \varepsilon)$-approximation for the original problem.

Coresets for different problems must be devised appropriately. For the MWISP, we create a grid with cells of diameter 0.29 and consider only one point of maximum weight inside each cell. For the MDSP, we create a grid with cells of diameter 0.24 and consider only the (at most four) points, inside each cell, with minimum or maximum coordinate in either dimension (breaking ties arbitrarily). Finally, we solve the MVCP by breaking each subproblem into two cases. In the first one, the number of input points is already bounded by a constant. In the second one, we use the same coreset as in the MWISP.

We assume a real-RAM computation model with floor function and constant-time hashing (as in [3]), so we can partition the input points into grid cells efficiently, yielding an overall $O(n)$ running time for our method. Without these operations, the running time of our algorithms becomes $O(n \log n)$.

3 Maximum-Weight Independent Set

In this section, we show how to obtain a linear-time $(4 + \varepsilon)$-approximation to the MWISP. We start by presenting a 4-approximation for point sets of constant diameter, and then we use the shifting strategy to obtain the desired $(4 + \varepsilon)$-approximation.

Given a point p and a set S of points, let $w(p)$ denote the weight of p, and let $w(S) = \sum_{p \in S} w(p)$. We say two or more points are *independent* if their minimum distance is strictly greater than 2.

Theorem 1. *Given a set P of n points with real weights as input, with* $\text{diam}(P) = O(1)$, *the MWISP can be* 4*-approximated in $O(n)$ time in the real-RAM.*

Proof. Our algorithm proceeds as follows. First, we find the points of P with minimum or maximum coordinates in either dimension. That defines a bounding box of constant size for P. Within this bounding box, we create a grid with cells of diameter $\gamma = 0.29$ (any value $\gamma < (2 - \sqrt{2})/2$ suffices). Note that the number of grid cells is constant, and therefore the points of P can be partitioned among the grid cells in $O(n)$ time (even without using the floor function or hashing). Then, we build the subset $Q \subseteq P$ as follows. For each non-empty grid cell C, we add to Q a point of maximum weight in $P \cap C$. Afterwards, we determine the maximum-weight independent set I^* of Q. Since $|Q| = O(1)$, this can be done in constant time. We return the solution I^*.

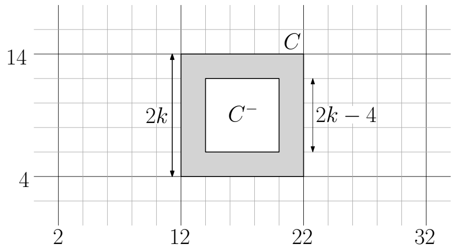

Figure 1: Grid rooted at $(2, 4)$ with $k = 5$ and the contraction of a cell

Next, we show that I^* is indeed a 4-approximation. We argue that, given an independent set $I \subseteq P$, there is an independent set $I' \subseteq Q$ with $4\ w(I') \geq w(I)$. Given a point $p \in P$, let $q(p)$ denote the point from Q that is contained in the same grid cell as p. Consider the set $S = \{q(p) : p \in I\}$. Note that $w(q(p)) \geq w(p)$ and $w(S) \geq w(I)$. The set S may not be independent, but since I is independent, the minimum distance in S is at least $2 - 2\gamma = 1.42 > \sqrt{2}$. We claim that the unit disk graph formed by S is a planar graph. To prove the claim, we show that a planar drawing can be obtained by connecting the points of S within distance at most 2 by straight line segments. Given a pair of points p_1, p_2 with distance $\|p_1p_2\| \leq 2$, the Pythagorean Theorem shows that a unit disk centered within distance greater than $\sqrt{2}$ from both p_1 and p_2 cannot intersect the segment p_1p_2. By the Four-Color Theorem [2], we can partition S into four independent sets $S_1, \ldots, S_4$. The set I' of maximum weight among $S_1, \ldots, S_4$ must have weight at least $w(I)/4$.

Since I^* is the maximum-weight independent set of Q, we have that I^* is a 4-approximation for the MWISP. □

The following theorem uses the shifting strategy to obtain a $(4 + \varepsilon)$-approximation for point sets of arbitrary diameter. The proof uses the ideas from [8], presented in a different manner and including details about an efficient implementation of the strategy.

Theorem 2. *Given a set P of n points in the plane as input, the MWISP can be $(4 + \varepsilon)$-approximated in $O(n)$ time on a real-RAM with constant-time hashing and the floor function. Without these operations, it can be done in $O(n \log n)$ time.*

Proof. Let k be the smallest integer such that

$$\left(\frac{k}{k-2}\right)^2 \geq 1 + \frac{\varepsilon}{4}. \tag{1}$$

Throughout this proof, we consider grids with square cells of side $2k$. We say a grid is *rooted at* a point (x, y) if there is a grid cell with corner at (x, y). Given a cell C, the square region $C^- \subset C$, called the *contraction* of C, is formed by removing from C the points within distance at most 2 from the boundary of C. Figure 1 illustrates these concepts.

The algorithm proceeds as follows. For i, j from 0 to $k-1$, we create a grid with cells of side $2k$ rooted at $(2i, 2j)$. For each cell C in the grid, we run the MWISP 4-approximation algorithm from Theorem 1 with point set $P \cap C^-$, obtaining a solution $I_{i,j}(C)$. Then, the independent set $I_{i,j}$ is constructed as the union of the independent sets $I_{i,j}(C)$ for all grid cells C. We return the maximum-weight set $I_{i,j}$ that is found, call it I^*.

To implement the algorithm efficiently, we create a subgrid of subcells of side 2, assigning each point to the subcell that contains it. In order to partition the n points into subcells, we use the floor function and constant-time hashing, taking $O(n)$ time. If these operations are not available, we determine the connected components of the graph (using the Delaunay triangulation, for example) and for each component we partition the points into subcells by sorting them by x coordinate, separating them into columns, and then sorting the points inside each column by y coordinate. The non-empty subcells are stored in a balanced binary search tree. This process takes $O(n \log n)$ time due to sorting, Delaunay triangulation, and binary search tree operations. Given the partitioning of the point set into subcells, each input to the MWISP algorithm can be constructed as the union of a constant number of subcells. Finally, the total size of the constant-diameter MWISP instances is $O(n)$, since each point from the original point sets appears in a constant number of such instances.

To prove that the returned solution I^* is indeed a $(4+\varepsilon)$-approximation, we use a probabilistic argument. Let i, j be picked uniformly at random from $0, \ldots, k-1$ and let OPT denote the optimal solution. For every cell C, we have

$$w(I_{i,j}(C)) \geq \frac{w(OPT \cap C^-)}{4}.$$

Consequently, by summing over all grid cells,

$$w(I_{i,j}) = \sum_C w(I_{i,j}(C)) \geq \frac{1}{4} \sum_C w(OPT \cap C^-).$$

We now bound $\mathrm{E}[w(I_{i,j})]$. Let $\rho(p)$ denote the probability that a given point p is contained in some contracted cell. Since $w(p)$ does not depend on the choice of i, j, we can write

$$4\, \mathrm{E}[w(I_{i,j})] \quad \geq \quad \mathrm{E}\left[\sum_C w\left(OPT \cap C^-\right)\right] \; = \sum_{p \in OPT} \rho(p) w(p).$$

Note that, for all $p \in P$, $\rho(p)$ corresponds to the ratio between the areas of C^- and C, namely

$$\rho(p) = \frac{\mathrm{area}(C^-)}{\mathrm{area}(C)} = \left(\frac{k-2}{k}\right)^2.$$

Therefore, by using inequality (1), we obtain

$$\mathrm{E}[w(I_{i,j})] \geq \frac{1}{4}\left(\frac{k-2}{k}\right)^2 \sum_{p \in OPT} w(p) \geq \frac{1}{4}\left(\frac{4}{4+\varepsilon}\right) w(OPT) = \frac{1}{4+\varepsilon}\, w(OPT).$$

Since I^* has maximum weight among the independent sets $I_{i,j}$, it follows that $w(I^*)$ is at least as large as their average weight. Therefore, I^* satisfies

$$w(I^*) \geq \mathrm{E}[w(I_{i,j})] \geq \frac{1}{4+\varepsilon}\, w(OPT),$$

closing the proof. □

4 Minimum Dominating Set

In this section, we show how to obtain a linear-time $(4+\varepsilon)$-approximation to the MDSP (in fact, a generalization of it). We start by presenting a 4-approximation for point sets of constant diameter, and then we use the shifting strategy to obtain the desired $(4+\varepsilon)$-approximation. We say that a point p dominates a point q if $\|pq\| \leq 2$. Given two sets of

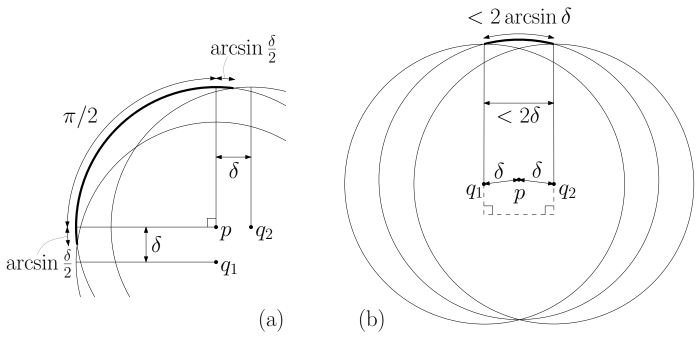

Figure 2: Proof of Lemma 3

points D and P', we say that D is a *P'-dominating set* if every point in P' is dominated by some point in D.

We now define a more general version of the MDSP, which we refer to as the *minimum partial dominating set problem (MPDSP)*. Such a generalization is necessary to properly apply the shifting strategy. In the MPDSP, we are given a set P of n points and also a subset $P' \subseteq P$. The goal is to find the smallest P'-dominating subset $D \subseteq P$.

In order to analyze our algorithm, we prove a geometric lemma that shows that the set-theoretic difference between a unit circle and two unit disks that are sufficiently close to it and form a sufficiently big angle consists of one or two "small" arcs. Given a point p, let $\bigcirc_p$ denote the unit disk centered at p, and $\partial\bigcirc_p$ denote its boundary circle.

Lemma 3. *Given $\delta > 0$ and three points $p, q_1, q_2 \in \mathbb{R}^2$ with (i) $\|pq_1\| \leq \delta$, (ii) $\|pq_2\| \leq \delta$, and (iii) the smallest angle $\angle q_1pq_2$ is greater than or equal to $\pi/2$, we have that:*

(1) the portion $T = (\partial\bigcirc_p) \setminus (\bigcirc_{q_1} \cup \bigcirc_{q_2})$ of the boundary $\partial\bigcirc_p$ consists of one or two circular arcs;

(2) if T consists of one circular arc, then the arc length is less than or equal to $\pi/2 + 2\arcsin(\delta/2)$; and

(3) if T consists of two circular arcs, then each arc length is less than $2\arcsin\delta$.

Proof. Statement (1) is clearly true. We start by proving statement (2). The arc length $\|T\|$ is maximized as the angle $\angle q_1pq_2$ decreases while the distances $\|pq_1\|, \|pq_2\|$ are kept constant, therefore it suffices to consider the case when $\angle q_1pq_2 = \pi/2$. The arc T centered at p can be decomposed into three arcs by rays in directions q_1p and q_2p, as shown in Figure 2(a). The central arc measures $\pi/2$, while each of the other two arcs measures $\arcsin(\delta/2)$, proving statement (2).

Next, we prove statement (3). Let T_1, T_2 denote the two arcs that form T with $\|T_1\| \geq \|T_2\|$. The arc length $\|T_1\|$ is maximized in the limit when $\|T_2\| = 0$, as shown in Figure 2(b). The rays connecting q_1 and q_2 to the two extremes of T_1 are parallel, and therefore $\|T_1\| < 2\arcsin\delta$. □

We are now able to prove the following theorem, which presents our 4-approximation algorithm for point sets of constant diameter.

Theorem 4. *Given two sets of points P and P' as input, with $P' \subseteq P$, $|P| = n$, and $\operatorname{diam}(P) = O(1)$, the MPDSP can be 4-approximated in $O(n)$ time in the real-RAM.*

Proof. First, we determine a bounding box of constant size for P, as we did in the algorithm for the MWISP. Within this bounding box, we create a grid with cells of diameter $\gamma = 0.24$. Note that the number of grid cells is constant, and therefore the points of P can be partitioned among the grid cells in $O(n)$ time (even without using the floor function or hashing). Then, we build the subset $Q \subseteq P$ as follows. For each non-empty grid cell, we add to Q the (at most four) extreme points inside the cell, i.e., those presenting minimum or maximum coordinate in either dimension. Ties are broken arbitrarily. Since there is a constant number of grid cells and we include in Q at most four points per cell, we have $|Q| = O(1)$. Afterwards, we determine the smallest P'-dominating subset $D^* \subseteq Q$. To do that, we examine the subsets of Q, from smallest to largest, verifying if all points of P' are dominated, until we find the dominating set D^*, which is returned as the approximate solution. Since Q has a constant number of points, this procedure takes $O(n)$ time.

Now we show that the returned solution D^* is indeed a 4-approximation. We argue that, given a P'-dominating set $D \subseteq P$, there is a P'-dominating set $D' \subseteq Q$ with $|D'| \leq 4\,|D|$. To build the set D' from D, we proceed as follows. For each point $p \in D$, if $p \in Q$, we add p to D'. Otherwise, since the set Q contains points of extreme coordinates in both x and y axes, in the cell of p, there are two points $q_1, q_2 \in Q$ such that (i) $\|pq_1\| \leq \gamma$, (ii) $\|pq_2\| \leq \gamma$, and (iii) the smallest angle $\angle q_1pq_2$ is at least $\pi/2$. We add these two points q_1, q_2 to D'.

By Lemma 3, the portion $T = (\partial\bigcirc_p) \setminus (\bigcirc_{q_1} \cup \bigcirc_{q_2})$ of $\partial\bigcirc_p$ consists of one or two circular arcs. We first consider the case where T consists of one circular arc. Let R be the set of points from P' which are dominated by p, but not by q_1 or q_2. If R is empty, then no extra point needs to be added to D'. Otherwise, the line ℓ which contains p and bisects T separates R into two (possibly empty) sets R_1, R_2. If $R_1 \neq \emptyset$, let p_3 be an arbitrary point of R_1. Since Q contains a point in the same cell as p_3, there is a point q_3 with $\|p_3q_3\| \leq \gamma$. We add the point q_3 to D'. Analogously, if $R_2 \neq \emptyset$, let p_4 be an arbitrary point of R_2 and let $q_4 \in Q$ be a point with $\|p_4q_4\| \leq \gamma$. We add the point q_4 to D'.

We now show that the four points $q_1, q_2, q_3, q_4 \in Q$ dominate all points dominated by p. Consider a point v that is dominated by p but not by q_1 or q_2. The point v must be inside the circular crown sector depicted in Figure 3(a) and described as follows. Because v is dominated by p, we have $\|pv\| \leq 2$. By Lemma 3, the arc length $\|T\| < 1.82$. Also, $\|pv\| > 1$, because otherwise the unit circles centered at p and v would intersect forming an arc of length at least $2\pi/3$, which is greater than $\|T\|$, in which case v is dominated by q_1 or q_2. Finally, since v is closer to p than it is to q_1 or q_2, it follows that v must be between the lines that connect p to the endpoints of T. This circular crown sector is bisected by the line ℓ. Using the law of cosines, we calculate the diameter of each circular crown sector as $d = \sqrt{8 - 8\cos(\|T\|/2)} < 1.76$. Therefore, for any point v inside the circular crown sector, the point q_3 (or q_4, analogously) that is within distance at most γ from a point inside the same sector dominates v, as $\|vq_3\| \leq d + \gamma < 2$.

Finally, if T consists of two circular arcs T_1, T_2 centered in p, then we start by adding those same points q_1, q_2 to D', as if T consisted of only one arc. Then, if necessary, we add new points q_3, q_4 to D' as follows. The points that are dominated by p but not by q_1 or q_2 must be within distance 1 of either T_1 or T_2. Let p_3, p_4 be arbitrary points that are within distance 1 of T_1 or T_2, respectively, but are not dominated by q_1 or q_2. If such points p_3, p_4 exist, then there are two points q_3, q_4 in Q that are within distance at most γ from respectively p_3, p_4. By Lemma 3, the largest arc among T_1, T_2 measures at most 0.49. The proof that all points dominated by p are dominated by q_1, q_2, q_3, or q_4 is analogous to the case where T consists of a single arc, using the circular crown sector illustrated in Figure 3(b).

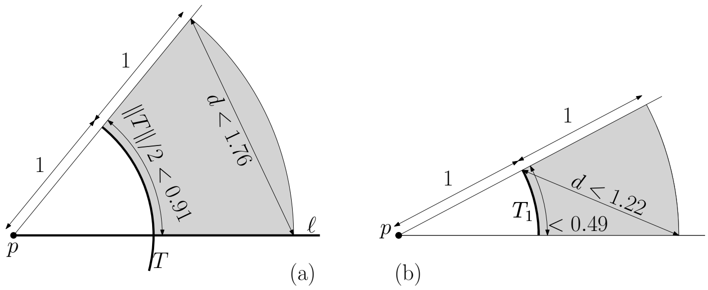

Figure 3: Proof of Theorem 4

Since D^* is minimum among all subsets of Q that are P'-dominating sets, D^* is a 4-approximation for the MPDSP. □

The following theorem uses the shifting strategy [8] to obtain a $(4+\varepsilon)$-approximation for point sets of arbitrary diameter.

Theorem 5. *Given two sets of points P and P' as input, with $P' \subseteq P$ and $|P| = n$, the MPDSP can be $(4+\varepsilon)$-approximated in $O(n)$ time on a real-RAM with constant-time hashing and the floor function. Without these operations, it can be done in $O(n \log n)$ time.*

Proof. Let k be the smallest integer such that

$$\left(\frac{k+2}{k}\right)^2 \leq 1 + \frac{\varepsilon}{4}. \tag{2}$$

Throughout this proof, we consider grids with square cells of side $2k$. We say a grid is *rooted at* a point (x, y) if there is a grid cell with corner at (x, y). Given a cell C, the square region C^+, called the *expansion* of C, is formed by C and all points within L_∞ distance at most 2 from C. Figure 4 illustrates these concepts.

The algorithm proceeds as follows. For i, j from 0 to $k-1$, we create a grid with cells of side $2k$ rooted at $(2i, 2j)$ and, for each cell C in the grid, we use Theorem 4 to 4-approximate the MPDSP with point sets $P \cap C^+, P' \cap C$, obtaining a solution $D_{i,j}(C)$. The dominating set $D_{i,j}$ is constructed as the union of the dominating sets $D_{i,j}(C)$ for all grid cells C. We return the smallest dominating set $D_{i,j}$ that is found, call it D^*.

To prove that the returned solution is indeed a $(4+\varepsilon)$-approximation, we use a probabilistic argument. Let i, j be picked uniformly at random from $0, \ldots, k-1$ and let OPT denote the optimal solution. For every cell C, we have

$$|D_{i,j}(C)| \leq 4 \, \left|OPT \cap C^+\right|.$$

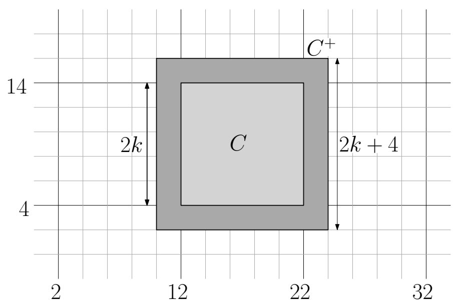

Figure 4: Grid rooted at $(2, 4)$ with $k = 5$ and the expansion of a cell

Consequently, by summing over all grid cells,

$$|D_{i,j}| = \sum_{C} |D_{i,j}(C)| \leq 4 \sum_{C} \left| OPT \cap C^+ \right| .$$

We now bound $\mathrm{E}[|D_{i,j}|]$. To do that, we define $\mathcal{C}^+(p)$ as the collection of all cell expansions containing a point p, so we can write

$$\frac{\mathrm{E}[|D_{i,j}|]}{4} \leq \mathrm{E}\left[\sum_{C} \left| OPT \cap C^+ \right|\right] = \mathrm{E}\left[\sum_{p \in OPT} |\mathcal{C}^+(p)|\right] = \sum_{p \in OPT} \mathrm{E}\left[|\mathcal{C}^+(p)|\right]$$

by the linearity of expectation. Note that the expected size of $\mathcal{C}^+(p)$, for all $p \in P$, corresponds to the ratio between the areas of C^+ and C, namely

$$\mathrm{E}\left[|\mathcal{C}^+(p)|\right] = \frac{\mathrm{area}(C^+)}{\mathrm{area}(C)} = \left(\frac{k+2}{k}\right)^2 .$$

Therefore, by using inequality (2), we obtain

$$\mathrm{E}[|D_{i,j}|] \leq 4 \left(\frac{k+2}{k}\right)^2 |OPT| \leq 4 \left(1 + \frac{\varepsilon}{4}\right) |OPT| = (4+\varepsilon)\ |OPT|.$$

Since the smallest among the dominating sets $D_{i,j}$ has no more than their average number of elements, the set D^* returned by the algorithm satisfies

$$|D^*| \leq \mathrm{E}[|D_{i,j}|] \leq (4+\varepsilon)\ |OPT|,$$

closing the proof. □

The MDSP is the special case of the MPDSP in which $P' = P$, and thus it can be $(4+\varepsilon)$-approximated in linear time by the same algorithm.

5 Minimum Vertex Cover

In this section, we show how to obtain a linear-time approximation scheme to the MVCP. We start by presenting an approximation scheme for point sets of constant diameter, and then we use the shifting strategy to generalize the result to arbitrary diameter. Differently than in the previous two problems, the size of a minimum vertex cover for a point set of constant diameter is not upper bounded by a constant. Therefore, strictly speaking, a coreset for the problem does not exist. Nevertheless, it is possible to use coresets to approach the problem indirectly.

Given a graph $G = (V, E)$ with n vertices, it is well known that I is an independent set if and only if $V \setminus I$ is a vertex cover. While a maximum independent set corresponds to a minimum vertex cover, a constant approximation to the maximum independent set does not necessarily correspond to a constant approximation to the minimum vertex cover. However, in certain cases, an even stronger correspondence holds, as we show in the following proof.

Theorem 6. *Given a set P of n points as input, with $\mathrm{diam}(P) = O(1)$, the MVCP can be $(1+\varepsilon)$-approximated in $O(n)$ time in the real-RAM, for constant $\varepsilon > 0$.*

Proof. Our algorithm considers two cases, depending on the value of n. If

$$n < \left(1 + \frac{3}{4\varepsilon}\right) \frac{(\mathrm{diam}(P) + 2)^2}{4},$$

then n is constant, and we can solve the MVCP optimally in constant time.

Otherwise, we use Theorem 1 to obtain a 4-approximation I to the maximum independent set. We now show that $V = P \setminus I$ is a $(1+\varepsilon)$-approximation to the minimum vertex cover. Let I_{OPT}, V_{OPT} respectively be the maximum independent set and the minimum vertex cover. Note that $|V| = n - |I|$ and $|V_{OPT}| = n - |I_{OPT}|$. By a simple packing argument, dividing the area of a disk of diameter $\mathrm{diam}(P) + 2$ by the area of a unit disk,

$$|I_{OPT}| \leq \frac{(\mathrm{diam}(P) + 2)^2}{4},$$

and consequently

$$n \geq \left(1 + \frac{3}{4\varepsilon}\right) |I_{OPT}| = \left(1 + \frac{3}{4\varepsilon}\right) (n - |V_{OPT}|).$$

Manipulating the previous inequality, we obtain

$$n \leq \frac{4\varepsilon + 3}{3} |V_{OPT}|. \tag{3}$$

Since I is a 4-approximation to I_{OPT},

$$|V| = n - |I| \leq n - \frac{|I_{OPT}|}{4} = \frac{4n - |I_{OPT}|}{4} = \frac{3n + |V_{OPT}|}{4}. \tag{4}$$

Combining (3) and (4), we can write $|V| \leq (1 + \varepsilon)|V_{OPT}|$, as desired. □

Using the shifting strategy we obtain the following result. The proof is similar to that of Theorem 2.

Theorem 7. *Given a set P of n points in the plane as input, the MVCP can be $(1 + \varepsilon)$-approximated in $O(n)$ time on a real-RAM with constant-time hashing and the floor function, for constant $\varepsilon > 0$. Without these operations, it can be done in $O(n \log n)$ time.*

Proof. Let k be the smallest integer such that

$$\left(\frac{k+2}{k}\right)^2 \leq \frac{1 + \varepsilon}{1 + \frac{\varepsilon}{2}}. \tag{5}$$

Throughout this proof, we consider grids with square cells of side $2k$. We say a grid is *rooted at* a point (x, y) if there is a grid cell with corner at (x, y). Given a cell C, the square region C^+, called the *expansion* of C, is formed by C and all points within L_∞ distance at most 2 from C.

The algorithm proceeds as follows. For i, j from 0 to $k - 1$, we create a grid with cells of side $2k$ rooted at $(2i, 2j)$ and, for each cell C in the grid, we use Theorem 6 to $(1 + \varepsilon/2)$-approximate the MVCP for $P \cap C^+$, obtaining a solution $V_{i,j}(C)$. The vertex cover $V_{i,j}$ is constructed as the union of the vertex covers $V_{i,j}(C)$ for all grid cells C. We return the smallest vertex cover $V_{i,j}$ that is found, call it V^*.

To prove that the returned solution is indeed a $(1+\varepsilon)$-approximation, we use a probabilistic argument. Let i, j be picked uniformly at random from $0, \dots, k-1$ and let OPT denote the optimal solution. For every cell C, we have

$$|V_{i,j}(C)| \leq \left(1 + \frac{\varepsilon}{2}\right) \; |OPT \cap C^+| \, .$$

Consequently, by summing over all grid cells,

$$|V_{i,j}| = \sum_C |V_{i,j}(C)| \leq \left(1 + \frac{\varepsilon}{2}\right) \sum_C |OPT \cap C^+| \, .$$

We now bound $\mathrm{E}[|V_{i,j}|]$. To do that, we define $\mathcal{C}^+(p)$ as the collection of all cell expansions containing a point p, so we can write

$$\frac{\mathrm{E}[|V_{i,j}|]}{1+\frac{\varepsilon}{2}} \leq \; \mathrm{E}\left[\sum_C |OPT \cap C^+|\right] = \mathrm{E}\left[\sum_{p \in OPT} |\mathcal{C}^+(p)|\right] = \sum_{p \in OPT} \mathrm{E}\left[|\mathcal{C}^+(p)|\right]$$

by the linearity of expectation. The expected size of $\mathcal{C}^+(p)$, for all $p \in P$, corresponds to the ratio between the areas of C^+ and C, namely

$$\mathrm{E}\left[|\mathcal{C}^+(p)|\right] = \frac{\mathrm{area}(C^+)}{\mathrm{area}(C)} = \left(\frac{k+2}{k}\right)^2 .$$

Therefore, by using inequality (5), we obtain

$$\mathrm{E}[|V_{i,j}|] \leq \left(1 + \frac{\varepsilon}{2}\right)\left(\frac{k+2}{k}\right)^2 |OPT| \leq (1+\varepsilon) \; |OPT|.$$

Since the smallest among the vertex covers $V_{i,j}$ has no more than their average number of elements, the set V^* returned by the algorithm satisfies

$$|V^*| \leq \mathrm{E}[|V_{i,j}|] \leq (1+\varepsilon) \; |OPT|,$$

closing the proof. □

6 Conclusion

We introduced a method to obtain linear-time approximation algorithms for problems on unit-disk graphs and other geometric intersection graphs as well. The central idea of the method is a technique to obtain approximate solutions when the inputs are point sets of constant diameter. The proposed method yielded linear-time $(4+\varepsilon)$-approximation algorithms for the MWISP and the MDSP, and a linear-time approximation scheme for the MVCP.

While the approximation ratio for the MWISP and the MDSP is 4 (for constant diameter inputs), we only know that the analysis is tight for the MDSP. Figure 5(a) shows an MDSP instance where our algorithm does not achieve an approximation ratio better than 4, even if we reduce the grid size and search for extreme points in a larger number of directions.

In contrast, for the MWISP, the best lower bound we are aware of is 3.25, as shown in the following example. Let P_1 be the weighted point set from Figure 5(b), where all adjacent vertices are at distance exactly 2. Create another set P_2 by multiplying the

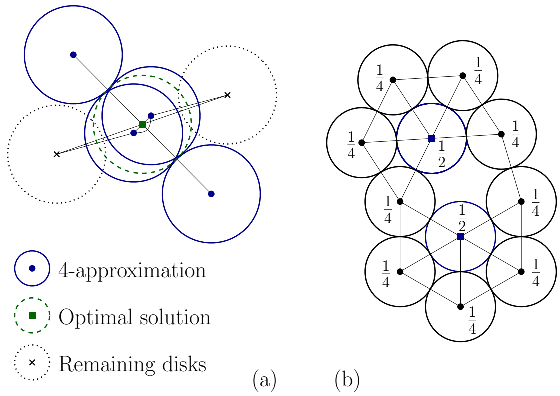

Figure 5: (a) Example where the approximation ratio for the MDSP is exactly 4 (b) Coin graph used in the example where the approximation ratio for the MWISP is 3.25

coordinates of the points in P_1 by $1 + \varepsilon$, while multiplying their weights by $1 - \varepsilon$, for arbitrarily small $\varepsilon > 0$. The set P_2 forms an independent set of weight just smaller than 3.25, while the maximum independent set in P_1 has weight 1. Since each vertex in P_2 has a smaller weight and is arbitrarily close to a vertex of P_1, the vertices of P_2 will be disregarded by the algorithm for the input instance $P_1 \cup P_2$.

Several open problems remain. Can we obtain an approximation ratio better than 4 in (close to) linear time for the MWISP, or at least for its unweighted version? Can the linear-time approximation scheme for the MVCP be generalized for the weighted version? Are the point coordinates really necessary, or is it possible to devise similar graph-based algorithms? Also, can we use our method to obtain better linear-time approximations to related problems on unit disk graphs such as finding the minimum-weight dominating set or the minimum connected dominating set?

References

[1] P. K. Agarwal, S. Har-Peled, and K. R. Varadarajan. Geometric approximation via coresets. In J. E. Goodman, J. Pach, and E. Welzl, editors, *Combinatorial and Computational Geometry*. Cambridge Univ. Press, 2005.

[2] K. Appel and W. Haken. Solution of the four color map problem. *Scientific American*, 237(4):108–121, 1977.

[3] J. Bentley, D. Stanat, and E. H. Williams Jr. The complexity of finding fixed-radius near neighbors. *Information Processing Letters*, 6(6):209–212, 1977.

[4] H. Breu and D. G. Kirkpatrick. Unit disk graph recognition is NP-hard. *Computational Geometry*, 9(1–2):3–24, 1998.

[5] X. Cheng, X. Huang, D. Li, W. Wu, and D.-Z. Du. A polynomial-time approximation scheme for the minimum-connected dominating set in ad hoc wireless networks. *Networks*, 42:202–208, 2003.

[6] G. D. da Fonseca, C. M. H. de Figueiredo, V. G. P. Pereira de Sá, and R. C. S. Machado. Efficient sub-5 approximations for minimum dominating sets in unit disk graphs. *Theoretical Computer Science*, arXiv:1204.3488, to appear.

[7] M. De, G. Das, and S. Nandy. Approximation algorithms for the discrete piercing set problem for unit disk. *International J. of Computational Geometry and Applications*, to appear.

[8] H. B. Hunt III, M. V. Marathe, V. Radhakrishnan, S. Ravi, D. J. Rosenkrantz, and R. E. Stearns. NC-approximation schemes for NP- and PSPACE-hard problems for geometric graphs. *Journal of Algorithms*, 26:238–274, 1998.

[9] R. K. Jallu, P. R. Prasad, and G. K. Das. Minimum dominating set for a point set in $\mathbb{R}^2$. *preprint*, arXiv:1111.2931, 2014.

[10] M. V. Marathe, H. Breu, H. B. Hunt III, S. S. Ravi, and D. J. Rosenkrantz. Simple heuristics for unit disk graphs. *Networks*, 25(2):59–68, 1995.

[11] T. Matsui. Approximation algorithms for maximum independent set problems and fractional coloring problems on unit disk graphs. In *JCDCG*, volume 1763 of *Lecture Notes in Computer Science*, pages 194–200, 1998.

[12] G. L. Nemhauser and L. E. Trotter Jr. Vertex packings: Structural properties and algorithms. *Mathematical Programming*, 8(1):232–248, 1975.

[13] T. Nieberg, J. Hurink, and W. Kern. Approximation schemes for wireless networks. *ACM Transactions on Algorithms*, 4(4):49:1–49:17, 2008.