The Expected Size of the Rule k Dominating Set: II

Jennie C. Hansen Actuarial Math and Statistics Department Herriot-Watt University J.Hansen@ma.hw.ac.uk

Eric Schmutz Department of Mathematics Drexel University Philadelphia, Pa. 19104 Eric.Jonathan.Schmutz@drexel.edu

> Li Sheng Department of Mathematics Drexel University Philadelphia, Pa. 19104 Isheng@cs.drexel.edu

September 10, 2004

Abstract

Rule k (k=1,2,3,...) is a well known family of approximation algorithms that can be used to find connected dominating sets in a graph. They were originally proposed by Dai, Li, and Wu as the basis for efficient routing methods for ad hoc wireless networks. In this paper we study the asymptotic performance of Rules 1 and 2 on random unit disk graphs formed from n random points in an $\ell_n \times \ell_n$ square region of the plane, and we show that:

- Rule 1 does poorly in the following sense: if $\ell_n = o(\sqrt{n})$, then with asymptotic probability one, Rule 1 produces a connected dominating set consisting of n o(n) vertices.
- Rule 2 produces much smaller dominating sets on average: if $\ell_n = O(\sqrt{n/\log n})$, then Rule 2 produces a connected dominating set whose expected size is $O(n/(\log \log n)^{3/2})$.

These results complement our results in a companion paper [18] where we consider the asymptotic performance of Rule k in the case where $k \ge 3$. **keywords and phrases**: coverage process, dominating set, localized algorithm, performance analysis, probabilistic analysis, Rule 2, Rule k, unit disk graph

1 Introduction

In this paper we consider the problem of finding a small connected dominating set for a *unit disk graph* G = (V, E), where the vertex set, V, is a set of points in \Re^2 . Given the vertex set V, the edge set E is determined as follows: an *undirected* edge $e \in E$ connects vertices $u, v \in V$ (and in this case we say that uand v are adjacent) iff the Euclidean distance between them is less than or equal to one. Unit disk graphs have been used by many authors as simplified mathematical models for the interconnections between hosts in a wireless network, and random unit disk graphs have been used as stochastic models for these networks. e.g. [8],[13], [16],[17],[24],[25]. We particularly mention the work of the Hipercom Project, e.g. [20],[21], because it is closely related to our work.

A dominating set in any graph G = (V, E) is a subset $C \subseteq V$ such that every vertex $v \in V$ either is in the set C, or is adjacent to a vertex in C. We say C is a connected dominating set if C is a dominating set and the subgraph induced by C is connected. Obviously G cannot have a connected dominating set if G itself is not connected. We use the acronym "CDS" for a dominating set C such that the subgraph induced by C has the same number of components as G has. In this paper we consider a random unit disk graph model, \mathcal{G}_n , which is connected with asymptotic probability one. So, in this case, any CDS for \mathcal{G}_n will also be connected with high probability.

The identification of a small connected dominating set for the graph which represents a network is an important step in several routing methods. The general idea of CDS-based algorithms is to select a small CDS, and have only those nodes responsible for determining routes [9], [28], [30], [32]. However, it is an NPhard computational problem to find the *minimal* connected dominating set in a unit disk graph [22]. Hence there is considerable practical interest in designing good approximation algorithms for finding small connected dominating sets (see, for example [2], [6], [7], [15], [26], [30]). In addition, in the context of ad hoc wireless networks, it is particularly desirable to have *decentralized* algorithms which allow individual hosts (i.e. vertices) to determine their membership in the final CDS based on a very restricted set of information that is available locally[12]. In a companion paper [18] we considered the average-case performance of a family of such decentralized algorithms, called 'Rule k' (k > 1), which were proposed by Dai, Li, and Wu [10], [32] for determining a CDS. The Rule k algorithm is decentralized in the sense that membership of a vertex v in the CDS is determined solely by considering the two hop topology of the graph. In other words, the only information used to determine whether to include vin the CDS is the subgraph induced by vertices whose graph-distance from v is two or less. Note in particular that no information about vertex v's coordinates in \Re^2 or anything else about its global location within the network is used by the algorithm.

In order to describe Rule k, we introduce some notation. We assume that each vertex has a unique identifier taken from a totally ordered set. For convenience, when |V| = n, we will use the numbers $1, 2, \ldots, n$ as IDs, and will number the vertices accordingly. If x_i is any vertex, with ID given by i, let $N(x_i)$ be the set consisting of x_i and any vertices that are adjacent to x_i . The CDS constructed by the Rule k algorithm is denoted $C_k(V)$, and its cardinality is $C_k(V) = |\mathcal{C}_k(V)|$. The elements of $\mathcal{C}_k(V)$ are called "gateway nodes". $\mathcal{C}_k(V)$ consists of all vertices $x_i \in V$ that are not excluded under the following version of Rule k:

Rule k: Vertex x_i is excluded from $C_k(V)$ iff $N(x_i)$ contains at least one set of k vertices $x_{i_1}, x_{i_2}, \ldots x_{i_k}$ such that

- $i_1 > i_2 > \cdots > i_k > i$, and
- The subgraph induced by $\{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$ is connected, and

•
$$N(x_i) \subseteq \bigcup_{t=1}^k N(x_{i_t}).$$

Wu Li and Dai proved that $C_k(V)$ is a CDS, and they conjectured that $C_k(V)$ is, in some sense, small on average. In [18] we showed that, in the context of an appropriate probability model and for $k \geq 3$, their conjecture is true and in this case $E(C_k(V))$ is also of the same order as the expected size of the *minimal* CDS.

In this paper we give an average-case analysis of 'Rule k' for k = 1, 2 by considering its performance when it is applied to a random unit disk graph \mathcal{G}_n . Specifically, let $\ell_1 \leq \ell_2 \leq \ldots$ be a sequence of real numbers such that $\ell_n = O(\sqrt{n/\log n})$ as $n \to \infty$, but $\ell_n \geq \log n$ for all n. Let \mathcal{Q}_n be an $\ell_n \times \ell_n$ square region in \Re^2 . Select n points $\mathcal{V}_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$ independently and uniform randomly from an \mathcal{Q}_n , and use these n points as the vertex set for a unit disk graph \mathcal{G}_n . With this probabilistic model, $C_1(\mathcal{V}_n)$ and $C_2(\mathcal{V}_n)$ are random variables and we prove asymptotic estimates for $E(C_1(\mathcal{V}_n))$ and $E(C_2(\mathcal{V}_n))$ as $n \to \infty$.

To provide some perspective on our choice of growth rates for ℓ_n , we mention that it is known that the threshold for connectivity is $\ell_n = \Theta(\sqrt{n/\log n})$; if ℓ_n grows faster than this, then the random unit disk graph \mathcal{G}_n will be disconnected with probability 1 - o(1) as $n \to \infty$. In this case, with high probability, $\mathcal{C}_k(\mathcal{V}_n)$ will not be a *connected* dominating set for \mathcal{G}_n . More precise versions of these remarks are provided in the new book by Penrose[27] which gives an up to date survey of random geometric graphs.

We also note here that the bound obtained in [18] for $E(C_k(\mathcal{V}_n))$ $(k \geq 3)$ depends in a crucial way on the fact that for any unit disk D_1 in \Re^2 there are always three points $u, v, w \in D_1$ such that D_1 is contained in the union of the unit disks centered at u, v, and w. On the other hand, it is impossible to find two points $u, v \in D_1$ (other than the center of the disk) such that D_1 is contained in the union of the unit disks centered at u and v. This means that to analyze the performance of Rule 1 and Rule 2 we must also consider certain geometrical issues. In particular, the analysis involves some interesting problems in elementary geometry and geometric probability. The rest of this paper is organized as follows. In the next section we analyze Rule 1 and show that $E(C_1(\mathcal{V}_n)) = n - o(n)$ when $\ell_n = o(n^{1/2})$. In Section 3 we prove a geometric lemma that is needed in Section 4 to obtain a local coverage result. In the remainder of the paper, we use the local coverage result to obtain an upper bound for $E(C_2(\mathcal{V}_n))$ and we discuss optimality issues. Finally, throughout the remainder of this paper we adopt the following notation. For any points $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ in \Re^2 , let d(P, Q) denote the ordinary Euclidean distance between P and Q in \Re^2 . For any r > 0, and any $P \in \Re^2$, let $D_r(P) = \{W \in \Re^2 | d(P, Q) \le r\}$ be a closed disk of radius r, centered at P.

2 Analysis of Rule 1

According to Rule 1, a vertex v is excluded from $\mathcal{C}_1(\mathcal{V}_n)$ only if it has a neighbor w such that $N(v) \subseteq N(w)$. We need not consider node IDs (the ordering of the vertices) in this section because we are using a *necessary* condition rather than a *sufficient* condition for a node to become a non-gateway under Rule 1. We prove that, with high probability in the random graph \mathcal{G}_n , most nodes do not have such a neighbor; most nodes end up as gateways and Rule 1 does not construct a 'small' CDS.

To this end, we now construct a random graph H_m , and use it to prove the crucial Lemma 1 below. Fix a point $x \in \Re^2$. Without loss of generality, assume x = (0, 0). Let $\hat{\mathcal{V}} = \{P_1, P_2, \ldots, P_m\}$ be m points sampled independently and uniform randomly from $D_1(x)$. Form a graph H_m with the m points as vertices, and two vertices adjacent iff the Euclidean distance between them is less than or equal to 1. (Note that H_m is not the same as \mathcal{G}_m since it has a small disk rather than a large square as the region.) It is possible that a single point P_i will be have degree m - 1 in H_m and therefore be a one-point dominating set. For example, if one of the P_i happens to coincide with x, the center of the unit disk, then P_i is adjacent to P_j for all $j \neq i$. However this is very rare: we shall prove that, with asymptotic probability one, there is no one point dominating set for H_m .

Let W_m be the size of the smallest dominating set in H_m , i.e. the smallest set \mathcal{C} of vertices with the property that all m vertices in H_m are within distance 1 of at least one of the points in \mathcal{C} .

Lemma 1 $Pr(W_m = 1) = O(m^{-1/2}).$

Proof. Let $\rho_m := \Pr(\hat{\mathcal{V}} \subseteq D_1(P_m))$ be the probability that all the m-1 points $P_1, P_2, \ldots, P_{m-1}$ are elements of $D_1(P_m)$. By Boole's inequality,

$$\Pr(W_m = 1) \le m\varrho_m. \tag{1}$$

Let R be the distance from P_m to x, and let A be the area of $D_1(x) \cap D_1(P_m)$. Using elementary geometry, we express the area A = A(R) in terms of R. Given that R = r, the conditional probability that the points $P_1, P_2, \ldots, P_{m-1}$ all lie $D_1(x) \cap D_1(P_m)$ is just $\left(\frac{A(r)}{\pi}\right)^{m-1}$. Since the density of R is 2r, it follows that

$$\varrho_m = \int_0^1 2r \left(\frac{A(r)}{\pi}\right)^{m-1} dr.$$
(2)

Let b_1 and b_2 be be the two points where the circles $\partial D_1(v)$ and $\partial D_1(P_m)$ intersect, i.e. the two points that lie at distance 1 from both v and P_m . By inclusion-exclusion, the area A of $D_1(x) \cap D_1(P_m)$ is

$$A = A_1 + A_2 - A_3 \tag{3}$$

where

- A_1 is the area of the sector of $D_1(x)$ that extends from b_1 to b_2 ,
- A_2 is the area of the sector of $D_1(P_m)$ that extends from b_1 to b_2 , and
- A_3 is the area of the parallelogram (x, b_1, P_m, b_2) .

Let θ be the acute angle between the line segments $\overline{x, b_1}$ and $\overline{x, P_m}$. Then

$$A_1 = A_2 = \theta. \tag{4}$$

Since $\frac{R}{2} = \cos(\theta)$, we have

$$A_3 = \sin(2\theta). \tag{5}$$

Putting (5) and (4) back into (3) and then into (2), we get

$$\varrho_m = \int_{\pi/3}^{\pi/2} 4 \left(\frac{2\theta - \sin 2\theta}{\pi}\right)^{m-1} \sin(2\theta) d\theta.$$
(6)

We estimate ρ_m by splitting the interval of integration into two parts. Let $u = 2\theta$, $\xi_m = \pi - \pi m^{-3/4}$, and define \mathcal{I}_1 and \mathcal{I}_2 as follows:

$$\varrho_m = \int_{2\pi/3}^{\pi} 2\left(\frac{u-\sin u}{\pi}\right)^{m-1} \sin u \ du = \int_{2\pi/3}^{\xi_m} + \int_{\xi_m}^{\pi} =: \mathcal{I}_1 + \mathcal{I}_2.$$
(7)

To estimate \mathcal{I}_1 , note that $u - \sin u$ is increasing, and that $\sin u \leq 1$. Therefore, for all u in $\left[\frac{2\pi}{3}, \xi_m\right]$, we have

$$2\left(\frac{u-\sin u}{\pi}\right)^{m-1}\sin u \le 2\left(\frac{\xi_m-\sin \xi_m}{\pi}\right)^{m-1} < 2\left(\frac{\xi_m}{\pi}\right)^{m-1}.$$

Hence

$$\mathcal{I}_1 \le 2\pi \left(\frac{\xi_m}{\pi}\right)^{m-1} \le 2\pi \left(1 - \frac{1}{m^{3/4}}\right)^{m-1} = O(e^{-m^{1/4}}).$$

We only need the cruder estimate

$$\mathcal{I}_1 = O(\frac{1}{m^{3/2}}).$$
 (8)

For the second integral, we use $\left(\frac{u-\sin(u)}{\pi}\right) \leq 1$:

$$\mathcal{I}_2 \le 2 \int_{\xi_m}^{\pi} 1^{m-1} \sin(u) du = 2 - 2 \cos(m^{-3/4}\pi) = O(\frac{1}{m^{3/2}}).$$
(9)

Putting (8) and (9) together, we get

$$\varrho_m = O(\frac{1}{m^{3/2}}).$$
(10)

The lemma now follows from (1) and (10).

The bound (10) is asymptotic; we cannot use it to estimate ρ_m for specific values of m. Equation (6) is exact, and $m\rho_m$ can be evaluated numerically to get upper bounds specific values of m. On the other hand, the inequality (1) is rather crude, so this method yields rather loose upper bounds for $\Pr(W_m = 1)$ for specific values of m.

Lemma 1 is the key tool that is needed to evaluate the asymptotic performance of Rule 1 on the random geometric graph \mathcal{G}_n . Let $S_n = n - C_1(\mathcal{V}_n)$ be the number of nodes that are *excluded* from $\mathcal{C}_1(\mathcal{V}_n)$ under Rule 1.

Theorem 2 If $\ell_n = o(n^{1/2})$, then $E(S_n) = O(\ell_n \sqrt{n}) = o(n)$.

Proof. In this section, let indicator variable $I_i = 1$ iff vertex *i* has a neighbor *j* such that every neighbor of *i* is a neighbor of *j*. Thus $S_n \leq \sum_{i=1}^n I_i$ and

$$E(S_n) \le nE(I_1). \tag{11}$$

Let \mathcal{A}_1 be the event that $D_1(1) \subseteq \mathcal{Q}_n$, i.e. that vertex 1 is not one of the exceptional vertices near the border of the region \mathcal{Q}_n . In order to apply Lemma 1, we need to take into account the variability in the degrees of the vertices. Let ρ_1 be the degree of vertex 1, i.e. number of nodes in $D_1(1)$ other than vertex 1 itself. Let $\mu = E(\rho_1|\mathcal{A}_1) = \frac{(n-1)\pi}{\ell_n^2}$. We can write ρ_1 as a sum of n-1 independent indicator variables: $\rho_1 = \sum_{j=2}^n I_j^{(1)}$ where $I_j^{(1)}$ is one iff the distance from vertex 1 to vertex j is less than one. Hence we can apply Chernoff's inequality to conclude that,

$$\Pr\left(\rho_1 < \frac{\mu}{2} | \mathcal{A}_1\right) < \exp(-\mu/8) = \exp\left(-\frac{\pi(n-1)}{8\ell_n^2}\right) = o(1). \tag{12}$$

(The asymptotic estimate uses the fact that $\ell_n = o(n^{1/2})$.) Now let \mathcal{B}_1 be the event that $\rho_1 > \mu/2$, and let $\mathcal{T}_1 = \mathcal{A}_1 \cap \mathcal{B}_1$. Then

$$\Pr(\mathcal{T}_1) = \Pr(\mathcal{A}_1) \Pr(\rho_1 > \mu/2 | \mathcal{A}_1) = 1 - o(1).$$
(13)

By Lemma 1,

n

$$\Pr(I_1 = 1 | \mathcal{T}_1) = O(\mu^{-1/2}) = O\left(\frac{\ell_n}{\sqrt{n}}\right).$$
(14)

But then

$$\Pr(I_1 = 1) = \Pr(I_1 = 1 | \mathcal{T}_1) \Pr(\mathcal{T}_1) + \Pr(I_1 = 1 | \mathcal{T}_1^c) \Pr(\mathcal{T}_1^c) = O\left(\frac{\ell_n}{\sqrt{n}}\right) = o(1),$$
(15)

and

$$-E(C_1(\mathcal{V}_n) = E(S_n) \le n \Pr(I_1 = 1) = O(\ell_n \sqrt{n}) = o(n)$$

3 A Geometric Lemma

As observed in [21], a unit disk centered at a point x cannot be completely covered with two unit disks having centers at points u and w ($u \neq x \neq w$): $(D_1(u) \bigcup D_1(w))^c \cap D_1(x) \neq \emptyset$. One might infer that a typical vertex $v \in \mathcal{V}_n$ is not likely to be be excluded from $\mathcal{C}_2(\mathcal{V}_n)$ under Rule 2 because no two points in $\mathcal{N}(v)$ will cover all the vertices in $\mathcal{N}(v)$. This reasoning suggests that, as in the case of Rule 1, Rule 2 does not construct a small CDS, but such reasoning is not sound. Typically there are points $u, w \in \mathcal{V}_n$ such that $D_1(u) \cup D_1(w)$ covers all but a negligible fraction of the disk $D_1(v)$ and the uncovered region is small enough so that it usually does not include any vertices. A more precise version of this statement is proved in the next section, but first we need to look carefully at the area of regions such as $(D_1(u) \bigcup D_1(w))^c \cap D_1(v)$. In particular, we need Lemma 3, which is the main result in this section.

To state Lemma 3 we adopt some notation. Throughout this section b > 1will be a parameter and in terms of b we let $L_b = \lfloor b^{1/3} (\log b)^2 \rfloor$, $\delta = \delta_b = \frac{1}{\sqrt[3]{\log b}}$, and $\theta_b = \pi/L_b$. We fix $o = (x_o, y_o) \in \Re^2$. We are going to partition the small disk $D_{\delta_b}(o)$ into $2L_b$ sectors as follows. Choose a new coordinate system centered at o, and for $0 \le i < L_b$, let Q_i be the sector consisting of those points $(x, y) = (r \cos \theta, r \sin \theta)$ whose polar coordinates satisfy $0 < r \le \delta$ and $(i - \frac{1}{2})\theta_b \le \theta \le (i + \frac{1}{2})\theta_b$. Similarly let R_i be the sector that is obtained by reflecting Q_i about o, namely the points with $0 < r < \delta$ and $(i - \frac{1}{2})\theta_b < \theta - \pi <$ $(i + \frac{1}{2})\theta_b$. Also, let \tilde{q}_i and \tilde{u}_i be the extreme points whose polar coordinates are respectively $(\delta, (i - \frac{1}{2})\theta_b)$ and $(\delta, (i + \frac{1}{2})\theta_b + \pi)$. Finally, for any points $u, w \in D_1(o)$, let X(u, w) denote the area of $(D_1(u) \bigcup D_1(w))^c \cap D_1(o)$, i.e. the area of the region in $D_1(o)$ that is not covered by $D_1(u) \sqcup D_1(w)$.

The analysis of Rule 2 depends on following geometric lemma about these sectors.

Lemma 3 There is a uniform constant C > 0 which is independent of the parameter b such that, for $0 \le i < L_b$, and for all $q_i \in Q_i, u_i \in R_i$, we have $X(q_i, u_i) \le X(\tilde{q}_i, \tilde{u}_i) \le \frac{C}{b \log^3 b}$.

Proof. We prove four facts which together imply Lemma 3. In the first fact, we observe that for any $v, w \in D_1(o)$ the omitted area X(v, w) increases if we move one (or both) of the two points v and w away from the origin along a radial line.

Fact 1 Let v_1, v_2 and w_1, w_2 be four points in $D_1(o)$ such that v_1 lies on the line segment $\overline{o, v_2}$ and w_1 lies on the line segment $\overline{o, w_2}$. Then $X(v_2, w_2) \ge X(v_1, w_1)$.

Proof. It suffices to show that $D_1(v_2) \cap D_1(o) \subseteq D_1(v_1) \cap D_1(o)$ and that $D_1(w_2) \cap D_1(o) \subseteq D_1(w_1) \cap D_1(o)$. Suppose $p \in D_1(v_2) \cap D_1(o)$. Since v_1 lies on the line segment from o to v_2 , we have $d(v_1, p) \leq \max(d(o, p), d(v_2, p)) \leq 1$. Hence $p \in D_1(v_1) \cap D_1(o)$. By a similar same argument, $D_1(w_2) \cap D_1(o) \subseteq D_1(w_1) \cap D_1(o)$.

Fact 2 Let a, b be the two points where the circles $\partial D_1(p), \partial D_1(q)$ intersect. Then, $\overline{a, b} \perp \overline{p, q}$, and the two line segments $\overline{a, b}$ and $\overline{p, q}$ intersect at their midpoints.

Proof. This follows immediately from the fact that d(p, a) = d(p, b) = d(q, a) = d(q, b) = 1.

Fact 3 Let o_1, o_2 be two points on the circle $x^2 + y^2 = \delta_b^2$. Then, $X(o_1, o_2)$ is a decreasing function of $\angle o_1 o o_2$.

Proof. For convenience, we will use polar coordinates. Without loss of generality, let o_1 be the point with polar coordinates $(r_{o_1}, \phi_{o_1}) = (\delta_b, \pi)$. Let o_2 be an arbitrary point on the circle with the polar coordinates (δ_b, ϕ_2) . By symmetry, we only need to consider the case when o_2 is in the first or second quadrant; we may, without loss of generality, assume that $0 \le \phi_2 \le \pi$. We will show that $X(o_1, o_2)$ is an increasing function of ϕ_2 , then the result follows from the fact that $\angle o_1 oo_2 = \pi - \phi_2$.

Let a_1, b_1 be the two points where the circles $\partial D_1(o_1)$ and $\partial D_1(o)$ intersect, with a_1 in the second quadrant and b_1 in the third quadrant.

Let o^* be a point on the circle $x^2 + y^2 = \delta_b^2$ so that $\partial D_1(o^*)$ meets with both $\partial D_1(o)$ and $\partial D_1(o_1)$ at a_1 . Let b^*, d^* be the other intersection points of $\partial D_1(o^*)$ with $\partial D_1(o)$ and $\partial D_1(o_1)$, respectively. For convenience, let's denote ϕ_{o^*} by ϕ^* . Figure 1 illustrates the position of $\partial D_1(o_1), \partial D(o)$, and $\partial D_1(o^*)$ and their intersections.



Figure 1: The position of the circle $\partial D_1(o^*)$

As in the proof of Fact 2, we have $\overline{a_1, d^*} \perp \overline{o_1, o^*}, \overline{a_1, b^*} \perp \overline{o, o^*}$. Notice also that o is on the line segment $\overline{a_1, d^*}$. So,

$$\angle b^* a_1 o = \angle oo^* o_1 = \angle o^* o_1 o = \frac{\phi^*}{2}.$$
 (16)

It follows that

$$0 < \phi^*/2 < \pi/2$$
, and, $\sin \frac{\phi^*}{2} = \frac{\delta_b}{2}$ (17)

Now, for the point o_2 with polar coordinates (δ_b, ϕ_2) , let a_2, b_2 denote the two points where $\partial D_1(o_2)$ and $\partial D_1(o)$ intersect, and let c_2, d_2 denote the two points where $\partial D_1(o_2)$ and $\partial D_1(o_1)$ intersect. There are two cases to consider: $\phi_2 \leq \phi^*$, and $\phi_2 \ge \phi^*$ Case 1. $\phi_2 \le \phi^*$.



Figure 2: The case when $\phi_2 \leq \phi^*$

Notice that a_1, b_1 partitions the circle $\partial D_1(o)$ into two arcs: the right section and the left section. When, $\phi_2 \leq \phi^*$, as illustrated in Figure 2, a_2, b_2 are both on the right section of the circle $\partial D_1(o)$ between a_1, b_1 . Similarly, c_2, d_2 are both on the right section of the circle $\partial D_1(o_1)$ between a_1, b_1 . Clearly,

$$X(o_1, o_2) = B_1 - (B_2 - B_3) = B_1 - B_2 + B_3,$$

where

- $B_1 = area(D_1(o_1)^c \cap D_1(o))$
- $B_2 = area(D_1(o) \cap D_1(o_2))$
- $B_3 = area(D_1(o_1) \cap D_1(o_2))$, the shaded area in Figure 2

Notice that B_3 is the only area that depends on ϕ_2 . We shall now give an expression for B_3 . Let's denote $\angle c_2 o_1 o_2 = y$. Since $\angle o_2 o_1 o = \frac{\phi_2}{2}$, we have

$$0 < y < \frac{\pi}{2}$$
, and, $\cos y = \delta_b \cos \frac{\phi_2}{2}$ (18)

By symmetry, one can see that the shaded region is partitioned equally by the line $\overline{c_2, d_2}$. So,

$$B_3 = 2\left(\frac{2y}{2\pi}\pi - \frac{1}{2}(2\sin y)(\cos y)\right) = 2y - \sin 2y.$$

Here, the first term is the area of the sector $D_1(o_1)$ that extends from c_2 to d_2 , and the second term is the area of the triangle (c_2, o_1, d_2) .

¿From the above two equations, we have

$$\frac{dX(o_1, o_2)}{d\phi_2} = \frac{dB_3}{d\phi_2} = \frac{dB_3}{dy} \cdot \frac{dy}{d\phi_2} = (1 - \cos 2y) \cdot \frac{\delta_b \sin \frac{\phi_2}{2}}{2\sin y} > 0.$$

Here the last inequality follows from the fact that $0 < \frac{\phi_2}{2}, y < \frac{\pi}{2}$. Thus $X(o_1, o_2)$ is an increasing function in ϕ_2 .

Case 2. $\phi_2 > \phi^*$.



Figure 3: The case when $\phi_2 > \phi^*$

One can see from Figure 3 that

$$X(o_1, o_2) = B_1 - (B_2 - B_3) = B_1 - B_2 + B_3$$

Where B_1, B_2 are defined the same as those in the case 1, but

 $B_3 = area(D_1(o_1) \cap D_1(o_2) \cap D_1(o))$, the shaded area in Figure 3

Again, B_3 is the only area that depends on ϕ_2 . We will now give an expression of B_3 . We show first that $\angle c_2 o a_1 = \angle a_2 o c_2$ by showing that $\phi_{c_2} - \phi_{a_1} = \phi_{a_2} - \phi_{c_2}$. Then, it follows that the region with area B_3 is split in half by the line segment $\overline{c_2, d_2}$.

¿From Figure 1, one can see that

$$\phi_{a_1} = \phi^* + \left(\frac{\pi}{2} - \angle b^* a_1 o\right) = \phi^* + \left(\frac{\pi}{2} - \frac{\phi^*}{2}\right) = \frac{\pi}{2} + \frac{\phi^*}{2} \tag{19}$$

To find ϕ_{a_2} , observe that, as in the proof of Fact 2, we have $\overline{a_2, b_2} \perp \overline{o, o_2}$. So, $\sin \angle b_2 a_2 o = \frac{\delta_b}{2}$. Comparing with (17), we see that $\sin \angle b_2 a_2 o = \sin \frac{\phi^*}{2}$. This implies that $\angle b_2 a_2 o = \frac{\phi^*}{2}$. Thus,

$$\phi_{a_2} = \phi_2 + \left(\frac{\pi}{2} - \angle b_2 a_2 o\right) = \phi_2 + \left(\frac{\pi}{2} - \frac{\phi^*}{2}\right) \tag{20}$$

Lastly, using the fact that $\overline{c_2, o} \perp \overline{o_1, o_2}$, we have

$$\phi_{c_2} = \pi - \left(\frac{\pi}{2} - \angle o_2 o_1 o\right) = \pi - \left(\frac{\pi}{2} - \frac{\phi_2}{2}\right) = \frac{\pi}{2} + \frac{\phi_2}{2} \tag{21}$$

It follows that $\phi_{c_2} - \phi_{a_1} = \phi_{a_2} - \phi_{c_2} = \frac{\phi_2}{2} - \frac{\phi^*}{2}$. Using that the circle $\partial D_1(o_1)$ in the polar system is

$$r = \sqrt{1 - \delta_b^2 \sin^2 \phi} - \delta_b \cos \phi$$

and that

$$\phi_{d_2} = -(\pi - \phi_{c_2}) = -(\frac{\pi}{2} - \frac{\phi_2}{2}) \tag{22}$$

we get

$$B_{3} = 2\left(\int_{-\left(\frac{\pi}{2} - \frac{\phi_{2}}{2}\right)}^{\frac{\pi}{2} + \frac{\phi^{*}}{2}} \int_{0}^{\sqrt{1 - \delta_{b}^{2} \sin^{2}\phi} - \delta_{b} \cos\phi} r \, dr d\phi + \frac{\frac{\phi_{2}}{2} - \frac{\phi^{*}}{2}}{2\pi} \cdot \pi\right)$$
$$= \int_{-\left(\frac{\pi}{2} - \frac{\phi_{2}}{2}\right)}^{\frac{\pi}{2} + \frac{\phi^{*}}{2}} 1 - \delta_{b}^{2} \sin^{2}\phi + \delta_{b}^{2} \cos^{2}\phi - 2\delta_{b} \cos\phi \sqrt{1 - \delta_{b}^{2} \sin^{2}\phi} d\phi + \frac{\phi_{2} - \phi^{*}}{2}$$

Thus,

$$\frac{dX(o_1,o_2)}{d\phi_2} = \frac{dB_3}{d\phi_2} = -\frac{1}{2} \left[1 - \delta_b^2 \sin^2 \left(-\frac{\pi}{2} + \frac{\phi_2}{2} \right) + \delta_b^2 \cos^2 \left(-\frac{\pi}{2} + \frac{\phi_2}{2} \right) \right] \\ -2\delta_b \cos \left(-\frac{\pi}{2} + \frac{\phi_2}{2} \right) \sqrt{1 - \delta_b^2 \sin^2 \left(-\frac{\pi}{2} + \frac{\phi_2}{2} \right)} \right] + \frac{1}{2} \\ = \frac{1}{2} \left[\delta_b^2 \cos^2 \frac{\phi_2}{2} - \delta_b^2 \sin^2 \frac{\phi_2}{2} + 2\delta_b \sin \frac{\phi_2}{2} \sqrt{1 - \delta_b^2 \cos^2 \frac{\phi_2}{2}} \right] \\ = \frac{1}{2} \left[- \left(\delta_b \sin \frac{\phi_2}{2} - \sqrt{1 - \delta_b^2 \cos^2 \frac{\phi_2}{2}} \right)^2 + 1 \right] \\ \ge 0$$

The last inequality follows because $0 \le \delta_b \sin \frac{\phi_2}{2} \le 1$, $0 \le \sqrt{1 - \delta_b^2 \cos^2 \frac{\phi_2}{2}} \le 1$, and thus $(\delta_b \sin \frac{\phi_2}{2} - \sqrt{1 - \delta_b^2 \cos^2 \frac{\phi_2}{2}})^2 < 1$.

Fact 4 Uniformly for all *i*, we have $X(\tilde{q}_i, \tilde{u}_i) = O(\frac{1}{b \log^3 b})$.

Proof. Without loss of generality, let i = 0 and v = (0, 0). To simplify notation, define $x_b = \delta_b \cos(-\frac{1}{2}\theta_b)$, $y_b = \delta_b \sin(-\frac{1}{2}\theta_b)$. Let (ξ, η) be the point in the first quadrant where the circles $x^2 + y^2 = 1$ and $(x - x_b)^2 + (y - y_b)^2 = 1$ meet. Then

$$X(\tilde{q}_0, \tilde{u}_0) \le 4 \int_0^{\xi} \sqrt{1 - x^2} - (y_b + \sqrt{1 - (x - x_b)^2}) dx$$
$$= -4y_b \xi + 4 \int_0^{\xi} \frac{-2xx_b + x_b^2}{\sqrt{1 - x^2} + \sqrt{1 - (x - x_b)^2}} dx$$

Hence we have

$$X(\tilde{q}_0, \tilde{u}_0) = O(\xi y_b) + O(x_b \xi^2) + O(x_b^2 \xi).$$
(23)

Note that $x_b^2 + y_b^2 = \delta_b^2 = \frac{1}{b^{2/3} \log^2 b}$, that $\xi^2 + \eta^2 = 1$, that $(\xi - x_b)^2 + (\eta - y_b)^2 = 1$, that $x_b = \delta_b (1 + O(\theta_b^2))$, and that $y_b = \frac{-\delta_b \theta_b}{2} (1 + O(\theta_b^2))$. Combining these equations, we get $\xi = O(\delta_b)$. Putting this estimate back into (23), we get

$$X(\tilde{q}_0, \tilde{u}_0) = O(\frac{1}{b \log^3 b}).$$
 (24)

In the analysis of Rule 2 it is necessary to consider vertices in \mathcal{G}_n which are close to the boundary of the square \mathcal{Q}_n . For this reason we define, for $o \in \Re^2_+$, the "truncated unit disk" $\hat{D}_1(o) := D_1(o) \cap \Re^2_+$ and we note that $\hat{D}_1(o) \subseteq D_1(o)$, and $\hat{D}_1(o) = D_1(o)$ iff $x_o, y_o \ge 1$. Then for L_b and δ_b as defined above, we have the following corollary to Lemma 3:

Corollary 4 There is a uniform constant C > 0, independent of the parameter b, such that, for all $o \in \Re^2_+$ such that $D_{\delta_b}(o) \subseteq \Re^2_+$, for $0 \le i < L_b$, and for all $q_i \in Q_i, u_i \in R_i$, we have $\hat{X}(q_i, u_i) \le X(\tilde{q}_i, \tilde{u}_i) \le \frac{C}{b \log^3 b}$, where $\hat{X}(q, u)$ is the area of $(D_1(q) \cap D_1(u))^c \cap \hat{D}_1(o)$.

Proof. Clearly $\hat{X}(q_i, u_i) \leq X(q_i, u_i)$ since $\hat{D}_1(o) \subseteq D_1(o)$. So the result follows from Lemma 3 (since $\tilde{q}_i, \tilde{u}_i \in D_{\delta_b}(o) \subseteq \Re^2_+$).

4 Local Coverage by Two Discs

Recall that under Rule 2 a vertex v_i is excluded from $C_2(\mathcal{V}_n)$ if there are two adjacent vertices, $v_{i_1}, v_{i_2} \in \mathcal{N}(v_i)$, with higher IDs than v_i which also 'cover' v_i , i.e. $\mathcal{N}(v_i) \subseteq \mathcal{N}(v_{i_1}) \cup \mathcal{N}(v_{i_2})$. In the analysis of Rule 2 we will distinguish vertices in $\mathcal{N}(v_i)$ with higher ID than v_i by coloring them blue; all other vertices in $\mathcal{N}(v_i)$ are colored white. With this in mind, we consider in this section a two-colored random unit disk graph and prove a local coverage result.

Let w and b be positive integers such that $w < b(\log b)^2$ and, as before, let $L_b = \lfloor b^{1/3} (\log b)^2 \rfloor$ and $\delta_b = \frac{1}{b^{1/3} \log b}$. Fix $o \in \Re^2_+$ such that $D_{\delta_b}(o) \subseteq \Re^2_+$ and select w + b points independently and uniform randomly from the truncated disk $\hat{D}_1(o)$. Color the first w points white, and the remaining b points blue. Form a random (improperly colored) unit disk graph $\hat{\mathcal{H}}_{w,b}$ by putting an edge between two of the w + b colored points iff the distance between them is one or less. Our goal in this section is to prove that, with high probability, $\hat{\mathcal{H}}_{w,b}$ contains a dominating set consisting of two blue vertices that are adjacent to each other.

For $0 \leq i < L_b$, let Q_i, R_i denote the sectors of $D_{\delta_b}(o)$ as defined in the previous section and let $N(Q_i), N(R_i)$ respectively be the number of blue vertices of $\hat{\mathcal{H}}_{w,b}$ that lie in Q_i and R_i . Let $\tau_b = \sum_{i=0}^{L_b-1} I_i$ where, in this section only, the indicator variable $I_i = 1$ if and only if $N(R_i) = N(Q_i) = 1$ (and otherwise $I_i = 0$.) We note that the distribution of τ_b depends on the position of o and we indicate this dependence by using the notation $\Pr_o(\tau_b \in \cdot)$. Provided o is not too close to the boundary of \Re^2_+ , we can obtain uniform bounds on the tail of the distribution of τ_b :

Lemma 5 $\Pr_o\left(\tau_b < \frac{b^{1/3}}{16\log^6 b}\right) = O(\frac{\log^6 b}{b^{1/3}})$ uniformly for all $o \in \Re^2_+$ such that $D_{\delta_b}(o) \subseteq \Re^2_+$.

Proof. Let $|\hat{D}_1(o)|$ denote the area of $\hat{D}_1(o)$, let $\hat{\lambda} = \hat{\lambda}(o) = \frac{\pi}{|\hat{D}_1(o)|}$, and define

$$\hat{p} = \frac{\text{Area}(Q_i)}{|\hat{D}_1(o)|} = \frac{\pi \delta_b^2 / 2L_b}{|\hat{D}_1(o)|} = \frac{\hat{\lambda}}{2b \log^4 b} \left(1 + O(\frac{1}{b^{1/3} \log^2 b})\right).$$
(25)

The expected value of I_i depends on o:

$$E_o(I_i) = b(b-1)\hat{p}^2(1-2\hat{p})^{b-2} = \frac{\hat{\lambda}^2}{4\log^8 b} \left(1 + O(\frac{1}{\log^4 b})\right).$$
(26)

Hence

$$E_o(\tau_b) = L_b E_o(I_i) = \frac{b^{1/3} \hat{\lambda}^2}{4 \log^6 b} \left(1 + O(\frac{1}{\log^4 b}) \right).$$
(27)

We likewise have, for $i \neq j$,

$$E_o(I_iI_j) = b(b-1)(b-2)(b-3)\hat{p}^4(1-4\hat{p})^{b-4} = \frac{\hat{\lambda}^4(o)}{16\log^{16}b} \left(1 + O(\frac{1}{\log^4 b})\right).$$
(28)

Note that

$$\frac{\pi}{4} \le |\hat{D}_1(o)| \le \pi,$$
 (29)

and therefore

$$1 \le \hat{\lambda}(o) \le 4. \tag{30}$$

Therefore we have uniformly for all $o \in \Re^2_+$ such that $D_{\delta_b}(o) \subseteq \Re^2_+$

$$Var_o(\tau_b) = O\left(\frac{b^{1/3}}{\log^6 b}\right).$$
(31)

Observe that

$$\Pr_o\left(\tau_b < \frac{b^{1/3}}{16\log^6 b}\right) \le \Pr_o\left(\tau_b \le \frac{1}{2}E_o(\tau_b)\right) \le \Pr_o\left(|\tau_b - E_o(\tau_b)| > \frac{1}{2}E_o(\tau_b)\right).$$
(32)

The lemma now follows from (27), (30)-(32) and Chebyshev's inequality.

Recall our assumptions that $w < b(\log b)^{3/2}$, $\delta_b = \frac{1}{b^{1/3} \log b}$, and $x_o, y_o \ge \delta_b$. With these assumptions, we have:

Theorem 6 There is a constant c > 0, independent of the position of o, such that with probability at least $1 - \frac{c}{(\log b)^{3/2}}$, the random graph $\hat{\mathcal{H}}_{w,b}$ has a connected dominating set that consists of two blue vertices in $D_{\delta_b}(o)$.

Proof.

Let $\mathcal{T}_b \subseteq \{0, 1, 2, 3, \dots, L_b - 1\}$ be the random subset of indices such that $i \in \mathcal{T}_b$ iff $N(Q_i) = N(R_i) = 1$. If $\mathcal{T}_b \neq \emptyset$, define $Y = \min \mathcal{T}_b$ to be the smallest of the indices in \mathcal{T}_b ; otherwise, if $\mathcal{T}_b = \emptyset$, set Y = -1.

Define the random variable X_b as follows: If $\tau_b = |\mathcal{T}_b| = 0$ then $X_b = 0$; otherwise, if $\mathcal{T}_b = \{i_1, i_2, \dots, i_{\tau_b}\}$ and $i_1 < i_2 < \dots < i_{\tau_b}$, then $X_b = 1$ iff $Q_{i_1} \cup R_{i_1}$ contains a blue connected dominating set for $\hat{\mathcal{H}}_{w,b}$.

Let $\mathcal{B} = \{g_1, g_2, \dots, g_b\}$ be the set of blue nodes, selected independently and uniform randomly from $\hat{D}_1(o)$. Define $\mathcal{Z} = \mathcal{B} \bigcap D_{\delta_b}(o)$ to be set of blue points that fall near the origin o, and let $Z = |\mathcal{Z}|$ be the number of these points. Then

$$\Pr_{o}(X_{b}=0) \leq \Pr_{o}\left(X_{b}=0, \tau_{b} \neq 0, Z \leq \frac{2\hat{\lambda}b^{1/3}}{(\log b)^{2}}\right) + \Pr_{o}(\tau_{b}=0) + \Pr_{o}\left(Z > \frac{2\hat{\lambda}b^{1/3}}{(\log b)^{2}}\right)$$
(33)

Note that Z has a binomial distribution: $Z \stackrel{d}{=} Bin(b, \hat{\lambda} \delta_b^2)$ where $\hat{\lambda}$ is as defined in the proof of Lemma 5. If $\beta = \frac{2\hat{\lambda}b^{1/3}}{(\log b)^2}$, then by Chernoff's inequality,

$$\Pr_o(Z \ge \beta) \le \exp(-b^{1/3}/4(\log b)^2).$$
 (34)

By Lemma 5, $\Pr_o(\tau_b = 0) = O(\frac{\log^6 b}{b^{1/3}})$. Therefore

$$\Pr_o(X_b = 0) \le \Pr_o(X_b = 0, \tau_b \neq 0, Z \le \beta) + O(\frac{\log^6 b}{b^{1/3}}).$$
(35)

Now we decompose the first term on the right side of (35) according to the value of Y.

$$\Pr_{o}(X_{b} = 0, \tau_{b} \neq 0, Z \leq \beta) = \sum_{k=0}^{L_{b}-1} \Pr_{o}(X_{b} = 0 | Y = k, Z \leq \beta) \Pr_{o}(Y = k, Z \leq \beta).$$
(36)

(The redundant condition $\tau_b \neq 0$ need not be included on the right side of (36) because it a consequence of the condition $Y \geq 0$.) We have

$$\Pr_o(X_b = 0|Y = k, Z \le \beta) = \sum_S \Pr_o(X_b = 0|\mathcal{Z} = S, Y = k) \Pr_o(\mathcal{Z} = S|Y = k, Z \le \beta)$$
(37)

where the sum is over subsets $S \subseteq [b]$ such that $2 \leq |S| \leq \beta$.

$$\Pr(X_b = 0 | \mathcal{Z} = S, Y = k) = 1 - \Pr(X_b = 1 | \mathcal{Z} = S, Y = k),$$
(38)

so it is enough to find a lower bound for $Pr(X_b = 1 | \mathcal{Z} = S, Y = k)$.

To simplify notation, let $\gamma = X(\tilde{q}_0, \tilde{u}_0)$, and recall that $\gamma = O(\frac{1}{b \log^3 b})$. In this section of the paper, define $|D_{\delta_b}(o)| = \frac{\pi}{b^{2/3}(\log b)^2}$ to be the area of the disk $D_{\delta_b}(o)$, and let $|\hat{D}_1(o)| = Area(\hat{D}_1(o))$. An important observation is that, once we have specified b - |S| = the number of blue points that fall *outside* $D_{\delta_b}(o)$, the locations in $D_{\delta_b}(o)^c \cap \hat{D}_1(o)$ of these b - |S| points are independent of the locations of the |S| blue points in $D_{\delta_b}(o)$, and are also independent of the locations of the white points. Hence

$$\Pr_{o}(X_{b} = 1 | \mathcal{Z} = S, Y = k) \geq \frac{\left(1 - \frac{|D_{\delta_{b}}(o)|}{|\hat{D}_{1}(o)|} - \frac{\gamma}{|\hat{D}_{1}(o)|}\right)^{b-|S|}}{\left(1 - \frac{|D_{\delta_{b}}(o)|}{|\hat{D}_{1}(o)|}\right)^{b-|S|}} \begin{pmatrix} 1 - \frac{\gamma}{|\hat{D}_{1}(o)|} \end{pmatrix}^{w} \\ \geq \left(1 - \frac{C}{b(\log b)^{3}}\right)^{b-|S|+w}$$
(39)

for some constant C that is independent of o. With our assumption $w < b(\log b)^{3/2}$ we get, for all sufficiently large b, the lower bound

$$\Pr_o(X_b = 1 | \mathcal{Z} = S, Y = k) \ge \left(1 - \frac{C'}{b(\log b)^3}\right)^{b(\log b)^{3/2}} \ge 1 - \frac{C''}{(\log b)^{3/2}} \quad (41)$$

for some constants C' and C'' which are independent of \mathcal{Z}, Y , and o. Hence

$$\Pr_o(X_b = 0) \le \frac{c}{(\log b)^{3/2}} \tag{42}$$

for some constant c that is *independent of the point o*.

5 Analysis of Rule 2

For $n \geq 1$, U_n be the number of vertices that are excluded from $C_2(\mathcal{V}_n)$ when Rule 2 is applied to the random graph \mathcal{G}_n : $U_n = \sum_i I_i$ where (in this section) the indicator variable $I_i = 1$ iff the node with ID *i* becomes a non-gateway under Rule 2. Assume that there is a positive constant \bar{c} such that, for all n > 1, $\log n \leq \ell_n \leq \bar{c} \sqrt{\frac{n}{\log n}}$. Let $\xi_n = \frac{\alpha_n}{\ell_n^2}$, where $\langle \alpha_n \rangle$ is any sequence of real numbers satisfying the following three conditions:

- $\alpha_n = o(n)$ as $n \to \infty$.
- $\xi_n = \frac{\alpha_n}{\ell_n^2} \to \infty$ as $n \to \infty$.
- For all sufficiently large n, $\frac{16n}{\log^{3/2} \xi_n} < \alpha_n$.

For example, if $\ell_n = \Theta(\sqrt{n/\log n})$, then the sequence $\alpha_n = \frac{32n}{(\log \log n)^{3/2}}$ satisfies the three conditions. On the other hand, if $\ell_n = \Theta((n/\log n)^t)$ for some fixed positive t < 1/2, then $\alpha_n = \frac{n}{\log n}$ satisfies the three conditions above. With these three assumptions, our goal is to prove

Theorem 7 $E(U_n) \ge n - O(\alpha_n).$

Proof. The idea of the proof is to use Theorem 6 to bound the probability that a typical vertex v_i is pruned by Rule 2. In this case the blue vertices correspond to nodes in $D_1(v_i)$ with IDs *higher* than i, and the white vertices correspond to nodes in $D_1(v_i)$ with lower IDs. Let $r = \frac{1}{\log^{3/2} \xi_n}$, and let \mathcal{A}_i be the event that $D_r(v_i) \subseteq \mathcal{Q}_n$. Then

$$\Pr(\mathcal{A}_i) = \frac{(\ell_n - 2r)^2}{\ell_n^2} \ge 1 - \frac{4r}{\ell_n}.$$
(43)

Let $\hat{D}_1(v_i) = D_1(v_i) \cap \mathcal{Q}_n$ be the set of points in \mathcal{Q}_n whose distance from v_i is one or less, and let $|\hat{D}_1(v_i)|$ be the area of $\hat{D}_1(v_i)$. Let $\rho_i^{(b)}$ denote the number of nodes in $\hat{D}_1(v_i)$ having a label that is larger than i, and let $\rho_i^{(w)}$ be the number of nodes in $\hat{D}_1(v_i)$ having a label that is smaller than i. Then, given the location of the *i*'th vertex v_i , $\rho_i^{(b)}$ has a Binomial $\left(n - i, \frac{|\hat{D}_1(V_i)|}{\ell_n^2}\right)$ distribution. Define $\mu_b = \mu_b(i)$ to be the expected value of $\rho_i^{(b)}$ given the location of the *i*'th point:

$$\mu_b = E(\rho_i^{(b)}|v_i) = \frac{(n-i)|\hat{D}_1(v_i)|}{\ell_n^2}.$$
(44)

Similarly $\rho_i^{(w)}$ has a Binomial $\left(i-1, \frac{|\hat{D}_1(v_i)|}{\ell_n^2}\right)$ distribution, and we define $\mu_w = \mu_w(i)$ to be the expected value:

$$u_w = E(\rho_i^{(w)}|v_i) = \frac{(i-1)|\hat{D}(v_i)|}{\ell_n^2}.$$
(45)

If \mathcal{A}_i occurs, then by Chebyshev's inequality,

ļ

$$\Pr(|\rho_i^{(b)} - \mu_b(i)| < \frac{\mu_b}{2} |\mathcal{A}_i| \ge 1 - \frac{16\ell_n^2}{n-i},\tag{46}$$

and similarly for $\rho_i^{(w)}$.

If we let \mathcal{D}_i be the event that *both* of the inequalities $|\rho_i^{(b)} - \mu_b(i)| < \frac{\mu_b}{2}$ and $|\rho_i^{(w)} - \mu_w(i)| < \frac{\mu_w}{2}$ are satisfied, then

$$\Pr(\mathcal{D}_i|\mathcal{A}_i) \ge 1 - \frac{16\ell_n^2}{n-i} - \frac{16\ell_n^2}{i-1}.$$
(47)

Combining (47) and (43), we get

$$\Pr(\mathcal{D}_i \cap \mathcal{A}_i) \ge \left(1 - \frac{16\ell_n^2}{n-i} - \frac{16\ell_n^2}{i-1}\right) \left(1 - \frac{4r}{\ell_n}\right). \tag{48}$$

Now let $\lambda_n = n - \alpha_n$, then clearly

$$E(U_n) \ge \sum_{i=\alpha_n}^{\lambda_n} \Pr(I_i=1) \ge \sum_{i=\alpha_n}^{\lambda_n} \Pr(I_i=1|\mathcal{D}_i \cap \mathcal{A}_i) \Pr(\mathcal{D}_i \cap \mathcal{A}_i)$$
(49)

To obtain a lower bound for the right hand side of inequality (49), we prove

Lemma 8 There is a constant $\tilde{c} > 0$ such that for all sufficiently large n and all $\alpha_n \leq i < \lambda_n$, $\Pr(I_i = 1 | \mathcal{D}_i \cap \mathcal{A}_i) \geq 1 - \frac{\tilde{c}}{(\log \xi_n)^{3/2}}$.

Proof. We begin by noting that given the event $\mathcal{D}_i \cap \mathcal{A}_i$ and $\alpha_n \leq i < \lambda_n = n - \alpha_n$, we have

$$\rho_i^{(w)} < \frac{3}{2}\mu_w(i) = \frac{3(i-1)|D_1(v_i)|}{2\ell_n^2} \le \frac{3\pi n}{2\ell_n^2}.$$
(50)

Similarly

$$\rho_i^{(b)} > \frac{1}{2}\mu_b(i) = \frac{(n-i)|\hat{D}_1(v_i)|}{2\ell_n^2} > \frac{\alpha_n\pi}{8\ell_n^2} = \frac{\xi_n\pi}{8}$$
(51)

It follows from inequalities (50) and (51) and from the conditions on the sequences $\langle \xi_n \rangle$ and $\langle \alpha_n \rangle$ that, given $\mathcal{D}_i \cap \mathcal{A}_i$ and $\alpha_n \leq i < \lambda_n$,

$$\rho_i^{(b)} (\log \rho_i^{(b)})^{3/2} \ge \rho_i^{(w)}.$$
(52)

Next we consider the conditional probability $\Pr(I_i = 1 | \rho_i^{(b)}, \rho_i^{(w)}, v_i, \mathcal{D}_i \cap \mathcal{A}_i)$ where the values of $\rho_i^{(b)}$ and $\rho_i^{(w)}$ and the location of v_i are consistent with the event $\mathcal{D}_i \cap \mathcal{A}_i$. In this case, it follows from inequality (51) that

$$\delta_b(\rho_i^{(b)}) := \frac{1}{(\rho_i^{(b)})^{1/3} \log(\rho_i^{(b)})} \le \frac{1}{(\xi_n/3)^{1/3} \log(\xi_n/3)} \le \frac{1}{(\log(\xi_n))^{3/2}} = r.$$
(53)

Since the event \mathcal{A}_i implies $D_r(v_i) \subseteq \Re^2_+$, it follows from (53) that $D_{\delta_b(\rho_i^{(b)})}(v_i) \subseteq \Re^2_+$. Finally, it follows from Theorem 4 that for some fixed positive constant \tilde{c}

$$\Pr(I_i = 1 | \rho_i^{(b)}, \rho_i^{(w)}, v_i, \mathcal{D}_i \cap \mathcal{A}_i) \ge 1 - \frac{c}{(\log(\rho_i^{(b)}))^{3/2}} \ge 1 - \frac{\tilde{c}}{(\log(\xi_i^{(b)}))^{3/2}}$$
(54)

for all sufficiently large n and all $\alpha_n \leq i < \lambda_n$. The lemma now follows from (54).

Recall that $\lambda_n = n - \alpha_n$, that $\alpha_n = o(n)$, that $\xi_n = \frac{\alpha_n}{\ell_n^2} \to \infty$ as $n \to \infty$, and that for all sufficiently large $n, \alpha_n > \frac{16n}{(\log \xi_n)^{3/2}}$. So it follows from Lemma 8 and (48) and (49), that

$$E(U_n) \ge n - 2\alpha_n + o(\alpha_n).$$

6 Discussion

The results in this paper show that if $\log n \leq \ell_n \leq c_{\sqrt{\frac{n}{\log n}}}$ then on average, for large $n, C_2(\mathcal{V}_n)$ is small relative to n but $C_1(\mathcal{V}_n)$ is not. However, Theorem 7 only gives us an upper bound on $E(C_2(\mathcal{V}_n)$ (i.e. $E(C_2(\mathcal{V}_n)) = O(\alpha_n)$). In the companion paper [18] we have shown that necessarily, for all large n and $k \geq 1$, $E(C_k(\mathcal{V}_n) \geq \ell^2/4$, and, in addition, for all $k \geq 3$ we have $E(C_k(\mathcal{V}_n)) = \Theta(\ell^2)$. So, in the case k = 2, there remains a gap between the lower bound for $E(C_2(\mathcal{V}_n)$ and the $O(\alpha_n)$ upper bound from Theorem 7. For example, if $\ell_n = \Theta(\sqrt{n/\log n})$, the lower and upper bounds for the expected size of the Rule 2 dominating set are respectively $\Theta(n/\log n)$ and $\Theta(n/(\log \log n)^{3/2})$. If we take $\ell_n = O((n/\log n)^t)$ with t < 1/2 the gap is even wider: the lower and upper bounds are respectively $\Theta((\frac{1}{\log n})^{2t})$ and $\Theta(\frac{n}{\log n})$. It remains an open problem to close this gap. We conjecture that, in fact, the expected size of the Rule 2 dominating set is $\Theta(\ell_n^2)$.

Acknowledgement We thank Jie Wu for introducing us to this problem and for his helpful suggestions regarding an early version of this paper. We also thank Jerzy Jaworski for carefully reading the paper and making many helpful suggestions that improved our exposition.

References

- C.Adjih, P.Jacquet, L.Viennot, Computing Connected Dominating Sets with Multipoint Relays, INRIA RR-4597 (2002).
- [2] K.M.Alzoubi, P.J.Wan, and O.Frieder, Distributed Heuristics for Connected Dominating Sets in Wireless Ad Hoc Networks, *Journal of Communications* and Networks, 4 (1) (2002) 1–8.
- [3] M.J.B.Appel and R.P.Russo, The minimum vertex degree of a graph on uniform points in [0, 1]^d, Advances in Appl.Probab 29 (3) (1997) 582–594.
- [4] E.Baccelli and P.Jacquet, Flooding Techniques in Mobile Ad Hoc Networks, INRIA RR-5002 (2003).
- [5] T.Camp and B.Williams, Comparison of Broadcasting Techniques for Mobile Ad Hoc Networks, Proceedings of the 3'rd ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2002 (2002) 194– 205.
- [6] M.Cardei, X.Cheng, X.Cheng, D.Du, Connected Domination in Multihop Ad Hoc Wireless Networks, JCIS (2002) 251–255.
- [7] B. Chen, K. Jamieson, H. Balakrishnan, R. Morris, Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks *Proc. of the 6th ACM MOBICOM Conf.*, Rome, Italy, July 2001.
- [8] B.N. Clark, C.J. Colburn, and D.J.Johnson, Unit Disk Graphs, Discrete Mathematics 86(1-3) (1990) 165–177.
- B.Das and V. Bharghavan Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets, International Conference on Communications 1 (1997) 376–380.
- [10] F.Dai and J.Wu, An Extended Localized Algorithm for Connected Dominating Set Formation in Ad Hoc Wireless Networks, *IEEE Transactions on Parallel and Distributed Systems*, **15** (10) 2004.
- [11] F.Dai and J.Wu, Performance analysis of broadcast protocols in ad hoc networks based on self pruning, *IEEE Trans.on Parallel and Distributed* Systems 15 (11) (2004).

- [12] D.Estrin, R.Govindan, J.Heidemann, Next century challenges: scalable coordination in sensor networks, MOBICOM (1999) 263–270.
- [13] E.N.Gilbert, Random Plane Networks, J.Soc.Indust.Appl.Math. 9 (1961) 533.
- [14] A.Godbole and B.Wielund, On the Domination Number of a Random Graph, Electronic Journal of Combinatorics 8 #R37 (2001)
- [15] S.Guha and S.Khuller, Approximation algorithms for connected dominating sets, Algorithmica 20 (4) (1998) 374–387.
- [16] P.Gupta and P.R.Kumar, "Critical power for asymptotic connectivity in wireless networks", in Stochastic Analysis, Control, Optimization and Applications, Birkhauser (1999) 547–566.
- [17] W.K.Hale, Frequency Assignment: Theory and Applications, Proc. IEEE 68 (1980) 1497–1514.
- [18] J.C.Hansen and E.Schmutz, The Expected size of the Rule k dominating set: I, submitted.
- [19] P.Hall, Introduction to the Theory of Coverage Processes, Wiley (1988).
- [20] P.Jacquet, Analytical Results on Connected Dominating Sets in Mobile Ad Hoc Networks, INRIA RR-5173 (2004).
- [21] P.Jacquet, A.Laouiti, P.Minet, L.Viennot, Performance of Multipoint Relaying in Ad Hoc Mobil Routing Protocols, In "Networking 2002" Lecture Notes in Computer Science 2345 (2002) 387–398.
- [22] D.Lichtenstein, Planar formulae and their uses, SIAM J.Comput. 11(2) (1982) 329–343.
- [23] M.V.Marathe, H.Breu, H.B.Hunt, S.S.Ravi, and D.J.Rosenkrantz, Simple Heuristics for Unit Disk Graphs, Networks 25 no.2 (1995) 59–68.
- [24] C.McDiarmid, Discrete mathematics and radio channel assignment, Recent Advances in algorithms and combinatorics, 27–63, CMS Books Math (2003).
- [25] C.McDiarmid, Random channel assignment in the plane, Random Structures and Algorithms 22 (2) 187–212.
- [26] W.Peng and X.Lu,On the reduction of broadcast redundancy in mobile ad hoc networks, Proceedings of the 1'st ACM International Symposium on Mobile Ad Hoc Networking (2000)129–130.
- [27] Random Geometric Graphs, Oxford Studies in Probability 5, Oxford University Press, (2003) ISBN 0-19-850626-0.
- [28] R.Sivakumar, B.Das, and V. Bharghavan, Spine-based routing in ad hoc networks, *Cluster Computing* 1 (2) (1998) 237–248.

- [29] Solomon, Geometric Probability, CBMS-NSF Conference Series 28 SIAM, 1978.
- [30] I. Stojmenovic, M. Seddigh, J. Zunic, Dominating sets and neighbor elimination based broadcasting algorithms in wireless networks, *IEEE Transac*tions on Parallel and Distributed Systems, Vol. 13, No. 1,(2002), 14-25.
- [31] Y.C.Tseng,S.Y.Ni, Y.S.Chen, J.P.Sheu, The broadcast storm problem in a mobile ad hoc network, Wireless Networks 8(2-3) (2002) 153-167.
- [32] J. Wu and H.Li, On calculating connected dominating set for efficient routing in ad hoc wireless networks, Workshop on Discrete Algorithms and Methods for MOBILE Computing and Communications (1999) 7–14.
- [33] A.C.C.Yao, On constructing spanning trees in k dimensional spaces and related problems, SIAM J.Computing 11 (4) (1982) 721–736.