On The Cover Time of Random Geometric Graphs

Chen Avin and Gunes Ercal
{avin, ercal}@cs.ucla.edu

UCLA CSD Technical Report TR040050

November 15, 2004

Abstract

The cover time of graphs has much relevance to algorithmic applications and has been extensively investigated. Recently, with the advent of ad-hoc and sensor networks, an interesting class of random graphs, namely *random geometric graphs*, has gained new relevance and its properties have been the subject of much study. A random geometric graph $\mathcal{G}(n,r)$ is obtained by placing n points uniformly at random on the unit square and connecting two points iff their Euclidean distance is at most r. The phase transition behavior with respect to the radius r of such graphs has been of special interest. We show that there exists a critical radius r_{opt} such that for any $r \geq r_{opt}$ $\mathcal{G}(n,r)$ has optimal cover time $\Theta(n \log n)$ with high probability, and, importantly, the critical radius guaranteeing optimal cover time is on the same order as the critical radius guaranteeing connectivity, namely $r_{opt} = \Theta(r_{con})$ where r_{con} denotes the critical radius guaranteeing asymptotic connectivity. Moreover, since a disconnected graph has infinite cover time, there is a phase transition and the corresponding threshold width is $O(r_{con})$. We are able to draw our results by giving a tight bound on the electrical resistance of $\mathcal{G}(n,r)$ via computing the power of certain constructed flows.

1 Introduction

The *cover time* C_G of a graph G is the expected time taken by a simple random walk on G to visit all nodes in G. This property has much relevance to algorithmic applications [25, 17, 35, 22, 21, 4], and methods of bounding the cover time of graphs have been thoroughly investigated [26, 2, 10, 9, 39, 3]. Several bounds on the cover times of particular classes of graphs have been obtained with many positive results [10, 9, 23, 24, 11].

A random geometric graph (RGG) is a graph $\mathcal{G}(n,r)$ resulting from placing n points uniformly at random on the unit square[1] and connecting two points iff their Euclidean distance is at most r. While these graphs have traditionally been studied in relation to subjects such as statistical physics and hypothesis testing [29], random geometric graphs have gained new relevance with the advent of ad-hoc and sensor networks [14, 30] as they are a model of such networks. Sensor networks have strict energy and memory constraints and in many cases are subject to high dynamics, created by failures, mobility and other factors. Thus, purely deterministic algorithms have disadvantages for such networks as they need to maintain data structures and have expensive recovery mechanism. Recently, questions regarding the random walk properties of such networks have been of interest

[1]We focus on the 2-dimensional case although we can also generalize our results to any d dimensions.

especially due to the locality, simplicity, low-overhead and robustness to failures of the process [18, 5, 7]. In particular random walk techniques have been proposed for gossiping in random geometric graphs [25], for information collection and query answering [32, 4] and even for routing [8, 33].

In ad-hoc and sensor networks, interference grows with increased communication radius. So, for a desirable property P of random geometric graphs, one wants to find a tight upper bound on the smallest radius r_p, the critical radius for which P exhibits a sharp threshold[2], such that P holds with high probability for all $\mathcal{G}(n,r)$ where $r \geq r_p$. Of special interest has been the critical radius r_{con} for connectivity, and it has been shown that if $\pi r^2 \geq \pi r_{con}^2 = \frac{\log n + \gamma_n}{n}$ then $\mathcal{G}(n,r)$ is connected with probability going to one as $n \to +\infty$ iff $\gamma_n \to +\infty$ [20, 36, 38].

Optimal cover time is cover time of $\Theta(n \log n)$ [16, 15]. In this paper, we study the property of optimal cover time for a simple random walk on random geometric graphs. We investigate the existence of a critical radius r_{opt} that will guarantee that $\mathcal{G}(n,r)$ with $r \geq r_{opt}$ has optimal cover time with high probability and show that such a threshold does exist, and, surprisingly, occurs at a radius $r_{opt} = \Theta(r_{con})$.

1.1 Discussion of Our Results and Techniques

Our main result can be formalized as follows:

Theorem 1 (Cover Time of RGG) *For $c > 1$, if $r^2 \geq \frac{c8 \log n}{n}$, then w.h.p.*[3] *$\mathcal{G}(n,r)$ has cover time $\Theta(n \log n)$. If $r^2 \leq \frac{\log n}{\pi n}$, then $\mathcal{G}(n,r)$ has infinite cover time with positive probability.*

The main contribution of this paper is in giving new tight theoretical bounds on the cover time and sharp threshold width associated with cover time for random geometric graphs. Aside from that, our results also have important implications for applications. Corollaries to our results are that both the *partial cover time* [4], which is the expected time taken by a random walk to visit a constant fraction of the nodes, and the *blanket time* [37], which is the expected time taken by a random walk to visit all nodes with frequencies according to the stationary distribution, are optimal for random geometric graphs. This demonstrates both the efficiency and quality of random walk approaches and certain token-management schemes for some ad-hoc and sensor networks [21, 25, 12, 9, 4].

In a recent related work Goel, Rai, and Krishnamachari have shown that any monotonic property of random geometric graphs has a sharp threshold [19] and have bounded the threshold width. While for general graphs optimality of cover time is not a monotonic property[4], it follows from our result that optimality of cover time is monotonic for random geometric graphs and has a threshold width of $O(r_{con})$, which is an order lower than the bounds on threshold width obtained by Goel *et al.* and supports their conjectured threshold width. Our results also improve upon bounds on the cover time obtained through bounding the mixing-time and spectral gap of random geometric graphs [31, 7, 5], as cover time can be bounded by the spectral gap [9]. In particular, the spectral gap method only guarantees optimal cover time of $G(n,r)$ for $r = \Theta(1)$.

The method that we used to derive our result is by bounding the electrical resistance of $\mathcal{G}(n,r)$, which bounds the cover time by the following result of Chandra *et al.* [10]: for any graph where R

[2]The threshold width is intuitively the difference between the smallest radius for which the property holds with high probability and the radius for which the property holds with low probability.

[3]Event $\mathcal{E}_n$ occurs with high probability if probability $P(\mathcal{E}_n)$ is such that $\lim_{n\to\infty} P(\mathcal{E}_n) = 1$.

[4]see appendix

is the the electrical resistance of the graph and m is number of edges in the graph

$$mR \leq \textit{cover time} \leq O(mR \log n) \tag{1}$$

In turn, we bound the resistance of $\mathcal{G}(n, r)$ by bounding the power of a unit flow as permitted by Thomson's Principle which we formalize later. The crux of our method is in constructing such a flow for any pair of points in $\mathcal{G}(n, r)$. As we investigate $\mathcal{G}(n, r)$ for which cover time is *optimal*, for any pair of points we construct the flow c in such a manner that the power of the flow satisfies $P(c) = O(\frac{n}{m}) = O(\frac{1}{\delta_{avg}})$ where δ_{avg} denotes the average degree of a node in $\mathcal{G}(n, r)$.

To construct a flow from u to v, we partition nodes into contour layers based on distance from u and expanding outward until the midpoint between u and v, then from the midpoint line onward contracting towards v in a mirror fashion. The idea of using contour layers that expand with distance from a point is similar to the layering ideas used by Chandra *et al.* [10] for meshes [5] and originally by Doyle and Snell [13] for infinite grids. Layers in our case can be visualized as slices of an isosceles right triangle along the hypotenuse that connects u and v. The flow can thus be thought of as *moving through* consecutive layers, and the total flow on the edges connecting consecutive layers is 1. And, just as the variance of a probability function is minimized for the uniform distribution, for each layer l, we allocate flow almost uniformly along the set of edges used between layer l and layer $l + 1$ to minimize the power.

And, to this end, via a simple coupon collection argument, we shall utilize a certain density property that is exhibited by $\mathcal{G}(n, r)$ for $r \geq \sqrt{\frac{8c \log n}{n}}$ for some $c > 1$. Concisely, the property is that any $\frac{r}{2\sqrt{2}} \times \frac{r}{2\sqrt{2}}$ square bin has $\Theta(\frac{nr^2}{8})$ nodes. Since two nodes are connected iff they are within distance r of each other, we see how the property yields that every node v of $\mathcal{G}(n, r)$ has degree $\delta(v) = \Theta(nr^2)$. Each layer adds a bin to the previous layer so that each layer l has l bins, and touching bins form cliques. We use this clique property to pair adjacent bins in consecutive layers and allocate uniform flow along the edges used in the pairings. The flow that we obtain in this manner has power $O(\frac{1}{nr^2}) = O(\frac{1}{\delta_{avg}})$, which is exactly the desired flow for optimal cover time as mentioned above.

1.2 Related Work

There is a vast body of literature on cover times and on geometric graphs, and to attempt to summarize all of the relevant work would not do it justice. We have already mentioned some of the related results previously in our Introduction, however, here we would like to highlight the related literature that has been most influential to our result, namely that of Chandra *et al.* [10] and Doyle and Snell [13].

The work of Doyle and Snell [13] is a seminal work regarding the connection between random walks and electrical resistance. In particular,they proved that while the infinite 2-dimensional grid has infinite resistance, for any $d \geq 3$ the resistance of the d-dimensional grid is bounded from above, and these results were established to be sufficient in re-proving Pólya's beautiful result that a random walk on the 2-dimensional grid is recurrent whereas a random walk on the d-dimensional grid for any $d \geq 3$ is transient. In obtaining this result, essentially the authors bounded the power of a unit current flow from the origin out to infinity via a layering argument and found that the power diverges for the 2-dimensional case and converges for every dimension greater than 2. The

[5] To the best of our knowledge, there has been no previous work on the cover time of finite grids other than toroidal meshes.

authors used a layering argument, namely partitioning nodes into disjoint contour layers based on their distance from the origin, and the rate of growth of consecutive layers can be seen as the crucial factor yielding the difference between the properties of the different dimensions. Later, Chandra *et al.* [10] proved the tight relation between commute time and resistance $C_{ij} = 2mR$ and used that relationship to extend Doyle and Snell's result by bounding the cover time of the *finite* d-dimensional mesh by bounding the power and resistance also via an argument based on expanding contour layers. Together with the tight lower bound of Zuckerman [39], they showed that the 2-dimensional torus has cover time of $\Theta(n \log^2 n)$, and for $d \geq 3$ the d-dimensional torus has optimal cover time of $\Theta(n \log n)$.

While this paper deals with $\mathcal{G}(n, r)$ there are striking similarities between $\mathcal{G}(n, r)$ and a more familiar family of random graphs, the *Bernoulli* graphs $\mathcal{B}(n, p)$ in which each edge is chosen independently with probability p [6]. For example, for critical probability $p_{con} = \pi r_{con}^2 = \frac{\log n + \gamma_n}{n}$, $\mathcal{B}(n, p)$ is connected with probability going to one as $n \to +\infty$ iff $\gamma_n \to +\infty$, and both classes of graphs have sharp thresholds for monotone properties [6]. Regarding cover time, Jonasson [23] and Cooper and Frieze [11] gave tight bounds on the cover time and an interesting aspect of our result is that we add another similarity and both classes of graphs have optimal cover time around the same threshold for connectivity. Yet, despite the similarities between $\mathcal{G}(n, r)$ and $\mathcal{B}(n, p)$, *Bernoulli* random graphs are not appropriate models for connectivity in wireless networks since edges are introduced independent of the distance between nodes. In wireless networks the event of edges existing between i and j and between j and k is *not* independent of the event of an edge existing between k and i. There are other notable differences between $\mathcal{G}(n, r)$ and $\mathcal{B}(n, p)$ as well. For example, the proof techniques for the above results for $\mathcal{G}(n, r)$ are very different than the proof techniques for the respective results for $\mathcal{B}(n, p)$. Interestingly, whereas the proof of [11] for optimality of cover time in Bernoulli graphs of $\Theta(\log n)$ average degree depends on the property that Bernoulli graphs do *not* have small cliques (and, in particular that small cycles are sufficiently far apart), in the case of random geometric graphs the existence of many small cliques uniformly distributed over the unit square like bins is essential in our analysis.

Another recent result with a bin-based analysis technique for random geometric graphs is that of Muthukrishnan and Pandurangan [28]. However, as their technique uses large overlapping bins where the overlap is explicitly stated to be essential and there is no direct utilization of cliques, the relationship between our simple bin-based analysis and Muthukrishnan and Pandurangan's more sophisticated bin-covering technique is not yet clear to us. Nevertheless, as the intuition underlying both techniques seems similar, we wonder if we may obtain the same results with tighter constants using Muthukrishnan and Pandurangan's bin-covering technique.

2 Bounding The Cover Time via Resistance

Let $G = (V, E)$ be a graph where $|V| = n$ and $|E| = m$,and let $EN(G)$ denote the electrical network corresponding to G, where each edge $e \in E$ has a resistor of 1 Ω. So, when speaking of the electrical properties of G, we are really speaking of the electrical properties of $EN(G)$. The resistance R of G is the maximum effective resistance R_{uv} between any two nodes $u, v \in V$ [13].

Let H_{uv} be the *hitting time*, the expected time for a random walk starting at u to arrive to v for the first time, and let C_{uv} be the *commute time*, the expected time for a random walk starting at u to first arrive at v and then return to u. Chandra *et al.* [10] proved the following theorem:

Theorem 2 *For any two vertices u and v in G the commute time $C_{uv} = 2mR_{uv}$*

Using this direct relation between resistance and random walks, they introduced the bound of (1) on the cover time for G

Let H_{max} be the maximum hitting time over all pairs of nodes in G. Since $H_{uv} \leq C_{uv}$ it follows that $H_{max} \leq \max_{u,v \in V} C_{uv} = 2mR$. In [4] it has been shown that the partial cover time can be bounded by H_{max}, so combining:

$$\textit{partial cover time} \leq O(H_{max}) \leq O(mR) \tag{2}$$

Thus, by bounding the resistance R we may obtain tight bounds on the cover time C_G and on the partial cover time.

A powerful method used to bound resistance is by bounding the power of a current flow in the network. The following definitions and propositions from the literature [10, 13, 34] help to formalize that method.

Definition 1 (Power of a flow) *Given electrical network (V, E, ρ), with resistance $\rho(e)$ for each edge e, a flow c is a function from $V \times V$ to $\mathbb{R}$, having the property that $c(u, v) = 0$ unless $\{u, v\} \in E$, and c is anti-symmetric, i.e., $c(u, v) = -c(v, u)$. The net flow out of a node will be denoted $c(u) = \sum_{v \in V} c(u, v)$, and the flow along an edge $e = \{u, v\}$ is $c(u, v)$. A source (respectively, sink) is a node u with $c(u) > 0$ (respectively $c(u) < 0$). The power $P(c)$ in a flow is $P(c) = \sum_{e \in E} \rho(e)c^2(e)$. A flow is a current flow if it satisfies Kirchoff's voltage law, i.e., for any directed cycle $u_0, u_1, \ldots, u_{k-1}, u_0$, $\sum_{i=0}^{k-1} c(u_i, u_{i+1 mod k}) \cdot \rho(u_i, u_{i+1 mod k}) = 0$.*

Proposition 1 *[Thomson Principle [13, 34]] For any electrical network (V, E, ρ) and flow c with only one source u, one sink v, and $c(u) = -c(v) = 1$ (i.e a **unit** flow), we have $R_{uv} \leq P(c)$, with equality when the flow is a current flow.*

Finally,

Proposition 2 *[Rayleigh's Short/Cut Principle [13]] Resistance is never raised by lowering the resistance on an edge, e.g. by "shorting" two nodes together, and is never lowered by raising the resistance on an edge, e.g. by "cutting" it. Similarly, resistance is never lowered by "cutting" a node, leaving each incident edge attached to only one of the two "halves" of the node.*

3 The Cover Time and Resistance of Geometric Graphs

Before proving Theorem 1 about *random* geometric graphs we are going to prove a more general Theorem about geometric graphs. A **geometric graph** is a graph $G(n, r) = (V, E)$ with $n = |V|$ such that the nodes of V are embedded into the unit square with the property that $e = (u, v) \in E$ if and only if $d(u, v) \leq r$ (where $d(u, v)$ is the Euclidean distance between points u and v). Let $\mathcal{F}(n, r(n))$ be a class of geometric graphs[6], we say that such a class is **geo-dense** if every square bin of size at least $A = r^2/8$ (in the unit square) has $\Theta(nA) = \Theta(nr^2)$ nodes as $n \to \infty$.

Theorem 3 *Any class of geometric graphs $\mathcal{F}(n, r(n))$ that is* geo-dense *and has $r(n) = \Theta(\frac{\log n}{n})$ has cover time of $\Theta(n \log n)$, partial cover time of $O(n)$, and blanket time of $\Theta(n \log n)$.*

[6] either random or deterministic

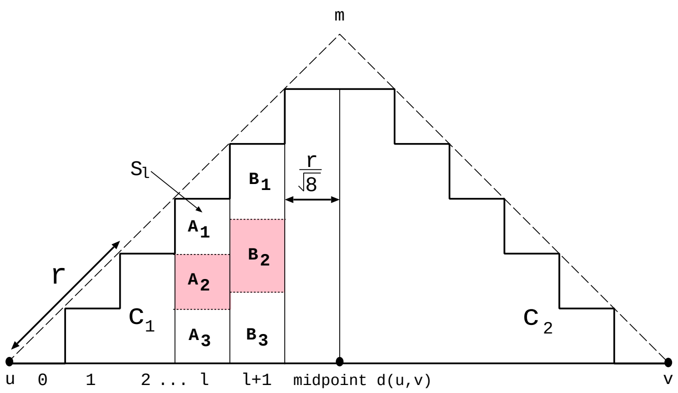

Figure 1: A flow c between u and v in $G(n, r)$

3.1 Proof of Theorem 3

Let $\mathcal{F}(n, r(n))$ be a class of geometric graphs that is *geo-dense*. We will prove the result using the bound on the cover time from (1) and by bounding the resistance between any two points u, v in $G = G(n, r) \in \mathcal{F}(n, r(n))$. Let V be the set of nodes of G and $\delta(v)$ denote the degree (i.e number of neighbors) of $v \in V$

Claim 1 $\forall v \in V \;\; \delta(v) = \Theta(nr^2)$

Proof: First note that the *geo-dense* property guarantees that if we divide the unit square into square bins of size $\frac{r}{\sqrt{2}} \times \frac{r}{\sqrt{2}}$ each, then the number of nodes in every bin will be $\Theta(nr^2)$. Since, for every bin, the set of nodes in the bin forms a clique, and every node $v \in V$ is in some bin, we have that $\delta(v) = \Omega(nr^2), \forall v \in V$. Similarly, when we divide the area into bins of size $r \times r$ every node may be connected to the nodes of at most nine bins (that is its own bin and the bordering bins), and we have that $\delta(v) = \Theta(nr^2), \forall v \in V$. □

Thus, since $m = \Theta(n^2r^2)$ if we can find r s.t. the resistance R of $G(n, r)$ is $O(\frac{n}{m}) = O(\frac{1}{nr^2})$ then we are done.

Theorem 4 *The resistance R_{uv} between $u, v \in V$ is $\Theta(\frac{1}{nr^2} + \frac{\log(d(u,v)/r)}{n^2r^4})$.*

Proof: The proof of the upper bound will be by bounding the power of a unit flow c that we construct between u and v.

Let $T(u, v)$ be an isosceles right triangle such that the line (u, v) is the hypotenuse. It is clear that such a triangle which lies inside the unit square must exist. We divide our flow c into two disjoint flows c_1 and c_2 where c_1 carries a unit flow from u up to the line perpendicular to the *midpoint* of $d(u, v)$ in increasing layer size, and c_2 forwards the flow in decreasing layer size up to v which is the only sink. By symmetry we can talk only about c_1 since the construction of c_2 mirrors that of c_1 and $P(c) = P(c_1 + c_2) = 2P(c_1)$ since the flows are disjoint.

To construct the flow in c_1 we divide the line $(u, midpoint(u, v))$ into $d(u, v)\sqrt{2}/r$ segments of size $r/\sqrt{8}$, and number them from 0 to $d(u, v)\sqrt{2}/r - 1$ (see Fig 1). [7] Let S_l be the largest rectangle of width $r/\sqrt{8}$ included in the intersection of the area perpendicular to the l^{th} segment

[7] Assume for simplicity the expression divides nicely, if not, the proof holds by adding one more segment that will end at the midpoint and overlap with the previous segment.

and $T(u,v)$. S_l will define the l^{th} layer in our flow. Note that the area of S_l is $lr^2/8$ and contains l squares of area $r^2/8$, each of them containing $\Theta(nr^2)$ nodes by the *geo-dense* property.

Let $V_l \subseteq V$ be the set of nodes in layer l. $V_0 = u$, and for $l > 0$ a node v is in layer l if and only if it is located inside S_l [8]. It follows that $|V_l| = \Theta(lnr^2)$. Edges in our flow are only among edges $e = (x,y)$ s.t. $x \in V_l$ and $y \in V_{l+1}$, and all other edges have zero flow. In particular the set of edges E_l that carries flow from layer l to layer $l+1$ in c_1 is defined as follows: For the case $l = 0$, E_0 contains all the edges from u to nodes in V_1, noting that $|E_0| = |V_1| = \Theta(nr^2)$ since $u \cup V_1$ is a clique (i.e the maximum $d(u,x), x \in V_1$ is r). This allows us to make the flow uniform such that each node in V_1 has incoming flow of $1/|V_1|$ and for each edge $e \in E_0$ $c_1(e) = 1/|V_l|$. For $l > 0$ (see again Fig. 1) we divide S_l into l equal squares $A_1, A_2, \dots A_l$ each of size $r^2/8$. Let V_{A_i} be the set of nodes contained in the area A_i. We then divide S_{l+1} into l equal sized rectangles $B_1, B_2 \dots B_l$ and define V_{B_i} similarly, with B_i touching A_i for each i.

Now let $E_l = \{(x,y) | x \in V_{A_i} \text{ and } y \in V_{B_i}\}$. Note again that since, for each i, the maximum $d(x,y)$ between nodes in A_i and nodes in B_i is r, $V_{A_i} \cup V_{B_i}$ is a clique (as the worst case distance occurs between the first two layers). So, the number of edges crossing from A_i to B_i is $|V_{A_i}||V_{B_i}| = \Theta(n^2r^4)$ by *geo-dense* property. The clique construction allows us to easily maintain the uniformity of the flow such that into each node in V_{B_i} the total flow is $1/l|V_{B_i}|$, and each edge carries a flow of $\Theta(1/ln^2r^4) = \Theta(1/E_l)$. All other edges have no flow. Now we compute the power of c:

$$
\begin{aligned}
R_{uv} &\le \sum_{e \in c} c(e)^2 = \sum_{e \in c_1} c_1(e)^2 + \sum_{e \in c_2} c_2(e)^2 = \\
&= 2 \sum_{l=0}^{\sqrt{2}d(u,v)/r} \sum_{e \in E_l} c_1(e)^2 \\
&= 2\frac{1}{|E_0|} + 2 \sum_{l=1}^{\sqrt{2}d(u,v)/r} \frac{1}{|E_l|} \\
&= 2O(\frac{1}{nr^2}) + 2O(\frac{1}{n^2r^4}) \sum_{l=1}^{\sqrt{2}d(u,v)/r} \frac{1}{l} \\
&= O(\frac{1}{nr^2} + \frac{\log(d(u,v)/r)}{n^2r^4})
\end{aligned}
$$

To prove the lower bound we again follow in the spirit of [13] and use the "Short/Cut" Principle. We partition the graph into $\lfloor d(u,v)/r \rfloor + 1$ partitions by drawing $\lfloor d(u,v)/r \rfloor$ squares perpendicular to the line (u,v), where the first partition P_0 is only u itself and the l^{th} partition P_l is the area of the l^{th} square excluding the $(l-1)^{\text{th}}$ square area. The last partition contains all the nodes outside the last square including v (See Fig 2.). We are shorting all vertices at the same partition, and following the reasoning of the upper bound, let m_l be the number of edges between partition l and

[8] breaking ties in some way

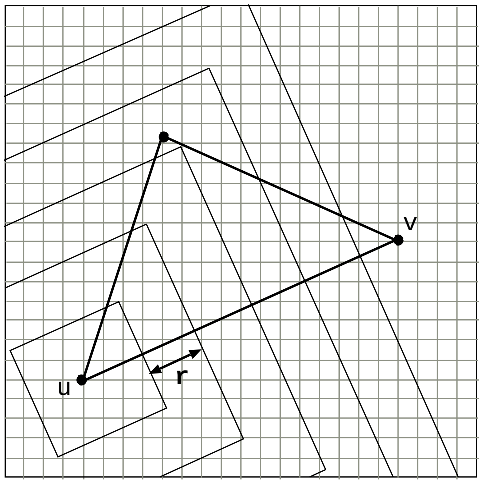

Figure 2: Lower bound for R_{uv} on the $G(n,r)$

$l+1$. m_0 is $\Theta(nr^2)$ and for $l>0$, $m_l = \Theta(ln^2r^4)$, so

$$R_{uv} \geq \sum_{l=0}^{\lfloor d(u,v)/r \rfloor} \frac{1}{m_l}$$
$$= \Omega(\frac{1}{nr^2}) + \sum_{l=1}^{\lfloor d(u,v)/r \rfloor} \Omega(\frac{1}{ln^2r^4})$$
$$= \Omega(\frac{1}{nr^2} + \frac{\log(d(u,v)/r)}{n^2r^4})$$

□

Corollary 2 *The resistance R of $G(n,r)$ is $\Theta(\frac{1}{nr^2} + \frac{\log(\sqrt{2}/r)}{n^2r^4})$.*

This follows directly from the fact that $\max d(u,v) \leq \sqrt{2}$. Now we can prove Theorem 3.

Proof of Theorem 3: Remember that $m = \Theta(n^2r^2)$, so all we need is $R = O(n/m) = O(1/nr^2)$ and then the cover time bound will follow by (1), the partial cover time bound will follow from (2), and the blanket time will follow from [37] and the $\log n$ order difference between the cover time and maximum hitting time. In order to have $R = \Theta(\frac{1}{nr^2})$ we want that $\frac{\log(\sqrt{2}/r)}{n^2r^4} = O(\frac{1}{nr^2})$, which means $\frac{\log(1/r)}{nr^2} \leq \alpha$ for some constant α. Taking $r^2 = (\frac{c\log n}{n})$ we get $\frac{\log(n/c\log n)}{c2\log n} = \frac{1}{2c} - \frac{\log(c\log n)}{2c\log n} \leq \frac{1}{2c}$. □

4 Cover Time and Resistance of $\mathcal{G}(n,r)$

After Proving Theorem 3, in order to prove Theorem 1 all we need to show is that for $c>1$, $r^2 = \frac{c8logn}{n}$ is sufficient to guarantee with high probability that $\mathcal{G}(n,r)$ is *geo-dense*. Note however that the second part of the theorem follows directly from [20] since if $\mathcal{G}(n,r)$ is disconnected with positive probability when $r^2 \leq \frac{\log n}{\pi n}$ then it has infinite cover time with positive probability.

To prove the *geo-dense* property for $\mathcal{G}(n,r)$ we utilize the two following Lemma which seems to be folklore [37] although we include a proof in the Appendix since we have not found a reference including a proof of the minimum condition.

Lemma 3 (Balls in Bins) *For a constant $c > 1$, if one throws $n \geq cB \log B$ balls uniformly at random into B bins, then w.h.p. both the minimum and the maximum number of balls in any bin is $\Theta(\frac{n}{B})$.*

And, the following lemma easily follows from the Balls in Bins Lemma:

Lemma 4 (Node Density) *For constants $c > 1$ and $a \geq 1$, if $r^2 = \frac{ca \log n}{n}$ then w.h.p. any area of size r^2/a in $\mathcal{G}(n,r)$ has $\Theta(c \log n)$ nodes.*

Proof: Let an area of r^2/a be a bin. If we divide the unit square into such equal size bins we have $B = \frac{n}{c \log n}$ bins. For the result to follow we check that Lemma (3) holds by showing that $n \geq c' B \log B$ for some constant $c' > 1$:

$$\begin{aligned} B \log B &= \frac{n}{c \log n} \log(\frac{n}{c \log n}) \\ &= \frac{n}{c \log n}(\log(n) - \log(c \log n)) \\ &= \frac{n}{c} - (\frac{n}{c \log n})(\log(c \log n)) \\ &\leq n/c \end{aligned}$$

□

Now combining the results of Lemmas (3) and (4) we can prove Theorem 1

Proof of Theorem 1: Clearly from Lemma 4 for $c > 1$, $r^2 = \frac{c8 \log n}{n}$ satisfies the *geo-dense* property w.h.p., and since r^2 is also $\Theta(\frac{\log n}{n})$ the result follows from Theorem 3. □

Corollary 5 *For $c > 1$, if $r^2 \geq \frac{c8 \log n}{n}$, then w.h.p. $\mathcal{G}(n,r)$ has partial cover time $O(n)$ and blanket time $\Theta(n \log n)$.*

5 Cover Time and Resistance of Deterministic Geometric Graphs

For an integer k, let the ***k*-fuzz** [13] of a graph G be the graph G_k obtained from G by adding an edge xy if x is at most k hops away from y in G. In particular, let $G_1(n)$ denote the 2-dimensional grid of n nodes, and let $G_k(n)$ be the k-fuzz of $G_1(n)$.

The question that we raise in this section is this: Given a finite 2−dimensional grid G of size n what is the minimum k s.t. G_k has cover time of $\Theta(n \log n)$?

Definition 2 *Let $\mathcal{D} = \mathcal{D}(n, r(n))$, where n is s.t. $\sqrt{n} \in \mathbb{Z}$ denote the class of r-**disk** geometric graphs, where the nodes of each instance of $\mathcal{D}(n,r)$ are placed on the unit square exactly as the 2-dimensional grid of n nodes. In other words, there is exactly one node at each position $(\frac{i}{\sqrt{n}}, \frac{j}{\sqrt{n}})$ $0 \leq i,j \leq \sqrt{n}, i,j \in \mathbb{Z}$.*

Note the following:

Corollary 6 *For $G_k(n)$, a $k-$fuzz of the 2-dimensional grid*

1. $G_1(n) = \mathcal{D}(n, \frac{1}{\sqrt{n}})$ *(i.e the 2-dimensional grid of n nodes).*

2. $G_k(n)$ *is a super-graph of* $\mathcal{D}(n, \frac{k}{\sqrt{2n}})$

3. $G_k(n)$ *is a sub-graph of* $\mathcal{D}(n, \frac{k}{\sqrt{n}})$

Claim 7 *For a constant k the resistance of $\mathcal{D}(n, \frac{k}{\sqrt{n}})$ is $\Theta(k^{-4} \log n)$.*

Proof: It is clear that $\mathcal{D}(n, \frac{k}{\sqrt{n}})$ satisfies the **geo-dense** property,[9] so the result follows directly from Theorem 4. □

Theorem 5 *For any constant k, the cover time of G_k is $\Theta(k^{-2} n \log^2 n)$.*

Proof: The upper bound follows directly from Corollary 6, Claim 7, and equation (1). To prove the lower bound note that the resistance R_{uv} is $\Theta(k^{-4} \log(\frac{d(u,v)\sqrt{n}}{k}))$. Letting $d'(u, v) = d(u, v)\sqrt{n}$ denote the non-normalized distance (hop distance), we have that $R_{uv} = \Omega(k^{-4} \log(d'(u, v)))$ where $1 \le d'(u, v) \le \sqrt{2n}$. Now we can use the method of Zuckerman [39] (specifically in Lemma 2 and Theorem 4 of that paper). And, by noting that it is known that the commute time $C_{uv} = 2mR_{uv}$, we have that hitting time ($E_u T_v$ in [39] notation) is $\Omega(k^{-2} n \log(d'(u, v))$. Then the result follows directly from the proof of Theorem 4 in [39]. □

Thus, we have the solution to our question:

Corollary 8 *$G_k(n)$ has Cover Time of $\Theta(n \log n)$ if $k = \gamma_n$ and $\lim_{n\to\infty} \dfrac{\log n}{\gamma_n^2} \le c$ for some constant c.*

6 Conclusions

We have shown that for a two dimensional random geometric graph $\mathcal{G}(n, r)$, if the radius r_{opt} is chosen just on the order of guaranteeing asymptotic connectivity then $\mathcal{G}(n, r)$ has optimal cover time of $\Theta(n \log n)$ for any $r \ge r_{opt}$. We present a similar proof for 1-dimensional random geometric graphs in the appendix. We find that the critical radius guaranteeing optimal cover time is $r_{opt} = \Omega(\frac{1}{\sqrt{n}})$ for such graphs, whereas the critical radius guaranteeing asymptotic connectivity is $r_{con} = \frac{\log n}{n}$. So, unlike the 2-dimensional case, we have $r_{opt} = \omega(r_{con})$.

Our proof techniques can be generalized to the d-dimensional random geometric graph $\mathcal{G}^d(n, r)$, yielding that for any given dimension d, $r_{opt} = \Theta(r_{con})$ with correspondingly optimal cover time. However, both grow exponentially with d which seems to be a consequence of a separation between average degree and minimum degree for higher dimensions rather than just an artifact of our method. We hope to present a tight result for this general case in the near future. Nevertheless, the case of dimension $d = 2$ is considered to be the hardest one [1]. This can intuitively be seen from the mesh results. The case for $d = 1$ (i.e the cycle) is easy to analyze. For $d > 2$ the cover time of the d-dimensional mesh is optimal [10], and we can show that for any k the cover time of the k-fuzz is also optimal. On the other hand, as we showed earlier, the cover time of the $k-$fuzz in 2 dimensions (i.e. $G_k(n)$) for constant k is not optimal and is $\Theta(n \log^2 n)$, making this the most interesting case.

[9] for large enough k

References

[1] Aldous, D., and Fill, J. *Reversible Markov Chains and Random Walks on Graphs.* Unpublished. http://stat-www.berkeley.edu/users/aldous/RWG/book.html.

[2] Aldous, D. J. Lower bounds for covering times for reversible Markov chains and random walks on graphs. *J. Theoret. Probab. 2*, 1 (1989), 91–100.

[3] Aleliunas, R., Karp, R. M., Lipton, R. J., Lovász, L., and Rackoff, C. Random walks, universal traversal sequences, and the complexity of maze problems. In *20th Annual Symposium on Foundations of Computer Science (San Juan, Puerto Rico, 1979).* IEEE, New York, 1979, pp. 218–223.

[4] Avin, C., and Brito, C. Efficient and robust query processing in dynamic environments using random walk techniques. In *Proceedings of the third international symposium on Information processing in sensor networks* (2004), ACM Press, pp. 277–286.

[5] Avin, C., and Ercal, G. Bounds on the mixing time and partial cover of ad-hoc and sensor networks. Tech. Rep. 040028, UCLA, June 2004. Under Review. ftp://ftp.cs.ucla.edu/tech-report/2004-reports/040028.pdf.

[6] Bollobás, B. *Random Graphs.* Academic Press, Orlando, FL, 1985.

[7] Boyd, S., Ghosh, A., Prabhakar, B., and Shah, D. Gossip and mixing times of random walks on random graphs. Unpublished, 2004. http://www.stanford.edu/ boyd/reports/gossip_gnr.pdf.

[8] Braginsky, D., and Estrin, D. Rumor routing algorthim for sensor networks. In *Proc. of the 1st ACM Int. workshop on Wireless sensor networks and applications* (2002), ACM Press, pp. 22–31.

[9] Broder, A., and Karlin, A. Bounds on the cover time. *J. Theoret. Probab. 2* (1989), 101–120.

[10] Chandra, A. K., Raghavan, P., Ruzzo, W. L., and Smolensky, R. The electrical resistance of a graph captures its commute and cover times. In *Proc. of the twenty-first annual ACM symposium on Theory of computing* (1989), ACM Press, pp. 574–586.

[11] Cooper, C., and Frieze, A. The cover time of sparse random graphs. In *Proceedings of the fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-03)* (Baltimore, Maryland, USA, 2003), ACM Press, pp. 140–147.

[12] Dolev, S., Schiller, E., and Welch, J. Random walk for self-stabilizing group communication in ad-hoc networks. In *Proceedings of the 21st IEEE Symposium on Reliable Distributed Systems (SRDS'02)* (2002), IEEE Computer Society, p. 70.

[13] Doyle, P. G., and Snell, J. L. *Random Walks and Electric Networks*, vol. 22. The Mathematical Association of America, 1984.

[14] Estrin, D., Govindan, R., Heidemann, J., and Kumar, S. Next century challenges: Scalable coordination in sensor networks. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking* (Seattle, Washington, USA, August 1999), ACM, pp. 263–270.

[15] Feige, U. A tight lower bound on the cover time for random walks on graphs. *Random Structures and Algorithms 6*, 4 (1995), 433–438.

[16] Feige, U. A tight upper bound on the cover time for random walks on graphs. *Random Structures and Algorithms 6*, 1 (1995), 51–54.

[17] Gkantsidis, C., Mihail, M., and Saberi, A. Random walks in peer-to-peer networks. In *in Proc. 23 Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). to appear* (2004).

[18] Goel, A., Rai, S., and Krishnamachari, B. Monotone properties have sharp thresholds in random geometric graphs. STOC slides, 2004. http://www.stanford.edu/ sanat/slides/thresholdsstoc.pdf.

[19] Goel, A., Rai, S., and Krishnamachari, B. Sharp thresholds for monotone properties in random geometric graphs. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing* (2004), ACM Press, pp. 580–586.

[20] Gupta, P., and Kumar, P. R. Critical power for asymptotic connectivity in wireless networks. *In Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming, W. M. McEneaney, G. Yin, and Q. Zhang, Eds* (1998), 547–566.

[21] Israeli, A., and Jalfon, M. Token management schemes and random walks yield self-stabilizing mutual exclusion. In *Proceedings of the ninth annual ACM symposium on Principles of distributed computing* (1990), ACM Press, pp. 119–131.

[22] Jerrum, M., and Sinclair, A. The markov chain monte carlo method: an approach to approximate counting and integration. In *Approximations for NP-hard Problems, Dorit Hochbaum ed.* PWS Publishing, Boston, MA, 1997, pp. 482–520.

[23] Jonasson, J. On the cover time for random walks on random graphs. *Comb. Probab. Comput. 7*, 3 (1998), 265–279.

[24] Jonasson, J., and Schramm, O. On the cover time of planar graphs. *Electronic Communications in Probability 5* (2000), 85–90.

[25] Kempe, D., Dobra, A., and Gehrke, J. Gossip-based computation of aggregate information. In *Proc. of the 44th Annual IEEE Symposium on Foundations of Computer Science* (2003), pp. 482–491.

[26] Matthews, P. Covering problems for Brownian motion on spheres. *Ann. Probab. 16*, 1 (1988), 189–199.

[27] Motwani, R., and Raghavan, P. *Randomized algorithms.* Cambridge University Press, 1995.

[28] Muthukrishnan, S., and Pandurangan, G. The bin-covering technique for thresholding random geometric graph properties. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, to appear* (2005).

[29] Penrose, M. D. *Random Geometric Graphs*, vol. 5 of *Oxford Studies in Probability.* Oxford University Press, May 2003.

[30] Pottie, G. J., and Kaiser, W. J. Wireless integrated network sensors. *Communications of the ACM 43*, 5 (2000), 51–58.

[31] Rai, S. The spectrum of a random geometric graph is concentrated. http://arxiv.org/PS_cache/math/pdf/0408/0408103.pdf, September 2004.

[32] Sadagopan, N., Krishnamachari, B., and Helmy, A. Active query forwarding in sensor networks (acquire). *To appear Elsevier journal on Ad Hoc Networks* (2003).

[33] Servetto, S. D., and Barrenechea, G. Constrained random walks on random graphs: routing algorithms for large scale wireless sensor networks. In *Proc. of the first ACM Int. workshop on Wireless sensor networks and applications* (2002), ACM Press, pp. 12–21.

[34] Synge, J. L. The fundamental theorem of electrical networks. *Quarterly of Applied Math.*, 9 (1951), 113–127.

[35] Wagner, I. A., Lindenbaum, M., and Bruckstein, A. M. Robotic exploration, brownian motion and electrical resistance. *Lecture Notes in Computer Science 1518* (1998), 116–130.

[36] Wan, P.-J., and Yi, C.-W. Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing* (2004), ACM Press, pp. 1–8.

[37] Winkler, P., and Zuckerman, D. Multiple cover time. *Random Structures and Algorithms 9*, 4 (1996), 403–411.

[38] Xue, F., and Kumar, P. R. The number of neighbors needed for connectivity of wireless networks. *Wirel. Netw. 10*, 2 (2004), 169–181.

[39] Zuckerman, D. A technique for lower bounding the cover time. In *Proc. of the twenty-second annual ACM symposium on Theory of computing* (1990), ACM Press, pp. 254–259.

7 Appendix

7.1 Resistance Bounds for 1-dimensional RGGs

Let $\mathcal{G}^1(n,r)$ denote a 1-dimensional random geometric graph formed by placing n nodes uniformly at random on the line $[0,1]$. In this section we give the result for $\mathcal{G}^1(n,r)$. Note that for $\mathcal{G}^1(n,r)$ the critical radius for connectivity is $r_{con} = \frac{\log n}{n}$ [20]. As before we first prove a more general case.

Definition 3 *Let $\mathcal{F}^1(n, r(n))$ be a class of 1-dimensional geometric graphs*[10]*, we say that such a class is* ***geo-dense*** *if every 1-dimensional interval of length at least $A = \frac{r}{2}$ has $\Theta(nA) = \Theta(\frac{nr}{2})$ nodes as $n \to \infty$.*

[10] either random or deterministic

Lemma 9 *For any geo-dense 1-dimensional geometric graph $G = G^1(n, r)$,*

- *The resistance of G is $\Theta(\frac{1}{nr} + \frac{1}{n^2r^3})$.*
- *The hitting time between two points u, v in G is $\Theta(n + \frac{d(u,v)}{r^2})$.*
- *The cover time C_G of G is such that $C_G = \Omega(n + \frac{1}{r^2})$ and $C_G = O((n + \frac{1}{r^2})\log n)$.*

Proof of 9: Consider any two points u, v in $G = G^1(n, r)$. Assume w.l.o.g. that v is to the right of u, namely that it has a higher coordinate than u. Let $d(u, v)$ be the distance between u and v.

For the upper bound, we use Thomson's Minimal Power Principle and bound the power of a unit flow between u and v. The idea is similar to the layering idea used for the 2-dimensional case, except that in 1-dimension there is no expansion of layers, and so each layer is a single bin.

Partition the segment between u and v into $h = \frac{2d(u,v)}{r}$ smaller segments, say bins, of length $\frac{r}{2}$ each, thus with $\Theta(\frac{nr}{2})$ nodes in each bin by geo-dense property. And, number the bins as B_i with increasing index $0 \le i \le h$ from left to right, letting the corresponding node sets be denoted by V_{B_i}. Note that each V_{B_i} is a clique, and moreover that each pair of adjacent segments $V_{B_i} \cup V_{B_{i+1}}$ is also a clique. Now we construct our flow c as follows: For every edge of the form $e = (u, z), z \in V_{B_0}$, $c(e) = \frac{1}{|V_{B_0}|} = \Theta(\frac{2}{nr})$. For every edge of the form $e = (z, v), z \in V_{B_h}$, $c(e) = \frac{1}{|V_{B_h}|} = \Theta(\frac{2}{nr})$. For every pair of consecutive bins $V_{B_i}, V_{B_{i+1}}$, the total flow between the bins is 1, and the flow between the bins is allocated uniformly along the $|V_{B_i}||V_{B_{i+1}}| = \Theta(\frac{n^2r^2}{4})$ edges between the bins in the direction of increasing bin index. Letting E_l denote the set of edges between bin B_l and bin B_{l+1}, this means that $|E_l| = \Theta(\frac{n^2r^2}{4})$. Moreover, letting $E_0 = \{(u, z)|z \in V_{B_0}\}$ and $E_{h+1} = \{(z, v)|z \in V_{B_h}\}$, we have $|E_0| = |E_{h+1}| = \Theta(\frac{2}{nr})$. So, now we may bound the power and thus the resistance of c in a straightforward manner by utilizing the uniformity of the flow along the edges between consecutive layers, obtaining:

$$\begin{aligned} R_{uv} \le P(c) = \sum_{e\in E} c^2(e) = \sum_{l=0}^{h+1} \sum_{e\in E_l} \frac{1}{|E_l|^2} = \\ = \sum_{l=0}^{h+1} \frac{1}{|E_l|} = \Theta(2\frac{2}{nr} + h\frac{4}{n^2r^2}) \\ = O(\frac{1}{nr} + \frac{d(u,v)}{n^2r^3}) \end{aligned}$$

For the lower bound, we use Rayleigh's Short/Cut Principle. Partition all nodes to the right of u into r length segments, which have $\Theta(nr)$ nodes each by geo-dense property, numbering the segments as B_i with increasing index i from left to right, and letting the corresponding node sets be denoted by V_{B_i}. Create a new graph G' by shorting nodes of G as follows: Short all nodes to the left of u with $V_{B_0}/\{u\}$ to create node b_0 in G'. Short all nodes of $\cup_{i\ge j-1} V_{B_i}/\{v\}$ together to create node b_{j-1} in G'. Now for each $0 < i < j$, short all the nodes in V_{B_i} together to create node b_i in G'.

Note that in G, a node in segment B_i can only talk to nodes of segments B_{i-1}, B_i, or B_{i+1}, namely itself or adjacent segments. Note also that every consecutive set of nodes all within distance r of each other is a clique. So, the set of edges of G', not counting multi-edges yet, is

$$E' = \{(b_i, b_{i+1})|0 \le i \le j-2\} \cup \{(u, b_0), (b_{j-1}, v)\}$$

Moreover, the multiplicity of each edge (b_i, b_{i+1}) is $\Theta(n^2r^2)$, and the multiplicity of edge (u, b_0) is the same as the multiplicity of edge (b_{j-1}, v) which is the same as $\Theta(nr)$. So summing the series of reciprocals of the edge multiplicities we get:

$$\begin{aligned} R_{uv}(G) \geq R_{uv}(G') &= \Omega(\frac{1}{nr} + (j-1)\frac{1}{n^2r^2}) \\ &= \Omega(\frac{1}{nr} + \frac{d(u,v)}{r}\frac{1}{n^2r^2}) \\ &= \Omega(\frac{1}{nr} + \frac{d(u,v)}{n^2r^3}) \end{aligned}$$

So, combining the upper and lower bounds we get $R_{uv}(G) = \Theta(\frac{1}{nr} + \frac{d(u,v)}{n^2r^3})$. By consider the maximum distance of 1 then, we get the resistance bound of our lemma. Then, since $m = \Theta(n^2r)$, from Theorem 2 we get the hitting time result (as the commute time and hitting time are on the same order). And, finally by Equation (1), we get the cover time result. □

It is a straightforward result of Balls in Bins lemma to show the following:

Corollary 10 *If $r \geq c2\frac{\log n}{n}$ for some $c > 1$, then $\mathcal{G}^1(n, r)$ is* geo-dense *w.h.p..*

Note immediately last corollary that this yields negative results for $\mathcal{G}^1(n, r)$ in that optimality of cover time for one-dimensional geometric graphs requires a radius of order strictly greater than the order for connectivity. In particular:

Corollary 11 *For any $r \geq c2\frac{\log n}{n}$, we have the following w.h.p. for $G = \mathcal{G}^1(n, r)$:*

- *The resistance of G is $\Theta(\frac{1}{nr} + \frac{1}{n^2r^3})$.*
- *The hitting time between two points u, v in G is $\Theta(n + \frac{d(u,v)}{r^2})$.*
- *The cover time C_G of G is such that $C_G = \Omega(n + \frac{1}{r^2})$ and $C_G = O((n + \frac{1}{r^2})\log n)$.*

As can be seen, the maximum hitting time between any two points is only optimal for $r = \Omega(\frac{1}{\sqrt{n}})$, and the cover time can only possibly be optimal for $r_{opt} = \Omega(\frac{1}{\sqrt{n}})$, so $r_{con} = o(r_{opt})$.

Definition 4 *Let $G_1^1(n)$ denote the 1-dimensional grid of n nodes, and let $G_k^1(n)$ be the k-fuzz of $G_1^1(n)$.*

And, finally, since any 1-dimensional k-fuzz $G_1^1(n)$ is a type of regular and geo-dense geometric graph $G^1(n, \frac{k}{n})$:

Corollary 12 *For any $1 \leq k \leq n$*

- *The resistance of $G_k^1(n)$ is $\Theta(\frac{1}{k} + \frac{n}{k^3})$.*
- *The hitting time between two points u, v in $G_k^1(n)$ is $\Theta(n + \frac{n^2 d(u,v)}{k^2})$.*
- *The cover time C of $G_k^1(n)$ is such that $C = \Omega(n + \frac{n^2}{k^2})$ and $C = O((n + \frac{n^2}{k^2})\log n)$.*

7.2 Optimal Cover Time is not Monotone

An immediate and well-known corollary to Rayleigh's Short/Cut Principle is that the Resistance R of a graph is monotone, as adding new edges can only decrease or not affect the resistance R. On the other hand, it is also well-known that, in general, cover time is not a monotone property of graphs. As a simple demonstrative example we can take the line of n nodes which has cover time of $O(n^2)$, and by adding edges we can create the lollipop graph which is known to have cover time of $O(n^3)$, and if we keep adding edges we will get the complete graph which has optimal cover time, $O(n \log n)$ [10]. One can wonder if this is still the case if the graph G already has cover time of $O(n \log n)$. In other words, can we create, by adding more edges, a graph G' which has cover time of $\omega(n \log n)$?

Lemma 13 *Cover time of $O(n \log n)$ is not a monotone property of graphs.*

Proof: The proof will be by counter example and by the lower bound for cover time given in Theorem 1 . Let G be the 3D grid of n nodes. It is known that G has cover time of $C_G = O(n \log n)$. We construct a graph G' be adding $O(n^2)$ edges to G is such a way that the resistance of the graph will not change: Let u_0 be the node at $(0,0,0)$ and u_n the node at $(\sqrt{n}, \sqrt{n}, \sqrt{n})$. Make all the points at L_1 distance at most $\sqrt{n}$ from u_0 a clique. The number of nodes in this clique is $\approx n/2$, and so the number of edges in this clique is $\approx n^2/8$, making the total number of edges in G' $m = \Theta(n^2)$. Since the minimum degree in G' is the same as in G, namely degree of 3 at u_n, the resistance of $G' \geq \frac{1}{3}$, and by Theorem 1 we get $C_{G'} = \Omega(n^2)$. □

7.3 Proof of Lemma 3

Proof: Let $n = cB \log B$ and note that when $n \to \infty$ then $B \to \infty$. The upper bound (maximum) of $\Theta(\frac{n}{B})$ on the number of balls in any bin is given in [27]. The following is proof of the lower bound: Fix a bin, say the first bin, and let X_1 denote the size of the first bin. Consider $Pr[X_1 = \frac{\log B}{c_1}]$ for some constant c_1:[11]

$$\begin{aligned} Pr[X_1 = \frac{\log B}{c_1}] &= \binom{cB \log B}{\log B/c_1} (\frac{1}{B})^{\frac{\log B}{c_1}} (1 - \frac{1}{B})^{cB \log B - \frac{\log B}{c_1}} \\ &\leq (ec_1 c)^{\frac{\log B}{c_1}} e^{\frac{\frac{\log B}{c_1} - cB \log B}{B}} \\ &= B^{\frac{1}{c_1}} (c_1 c)^{\frac{\log B}{c_1}} B^{\frac{1}{Bc_1} - c} \\ &= B^{\frac{(\log c_1 + \log c)}{c_1} + \frac{1}{c_1} + \frac{1}{Bc_1} - c} \end{aligned}$$

Since we want for this probability to be $\frac{1}{B^{1+\epsilon'}}$ for $\epsilon' > 0$ we need

$$c - (\frac{(\log c_1 + \log c)}{c_1} + \frac{1}{c_1} + \frac{1}{Bc_1}) > 1$$

Let $c = 1 + \epsilon$ where $\epsilon > 0$ can be arbitrary small constant and so we need c_1 s.t

$$1 + \epsilon - (\frac{\log c_1 + \log(1+\epsilon)}{c_1} + \frac{1}{c_1} + \frac{1}{Bc_1}) > 1 \tag{3}$$

[11] by using $(1 - \frac{1}{n})^r \leq e^{-r/n}$, $\binom{n}{k} \leq (\frac{ne}{k})^k$ and $c_1 = e^{log(c_1)}$.

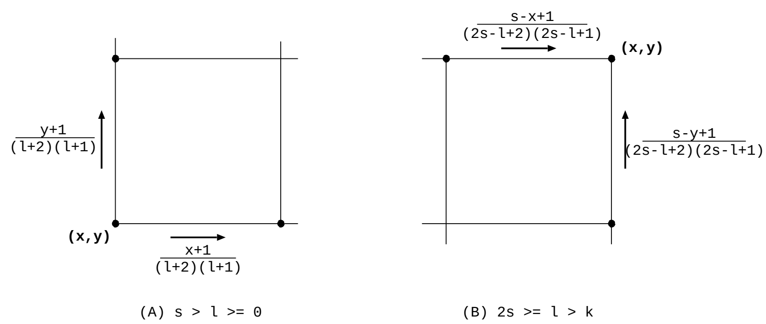

Figure 3: Unit flow for upper bound

The following c_1 will satisfy (3)

$$\frac{1}{c_1-1}+\frac{1}{B(c_1-1)}+\frac{\log c_1}{c_1-1}<\epsilon \tag{4}$$

So it is clear that for any constant $\epsilon > 0$ small as we want, there exists a constant c_1 that will satisfy (4), let that constant be c^*. Note easily that $Pr[X_1 = \frac{\log B}{c^*}] \geq Pr[X_1 = \frac{\log B}{c^*} - Q]$ for any $0 \leq Q \leq \frac{\log B}{c^*}$. Therefore, we have that for large enough B

$$\begin{aligned} Pr[X_1 \leq \frac{\log B}{c^*}] &\leq (\frac{\log B}{c^*})Pr[X_1 = \frac{\log B}{c^*}] \\ &\leq (\frac{\log B}{c^*})\frac{1}{B^{1+\epsilon'}} \end{aligned}$$

Finally to get the lower bound (minimum) for all bins, we use that the probability of the union of events is no more than their sum. Letting U denote the event that *some bin has less than* $\frac{\log B}{c^*}$ *balls*:

$$\begin{aligned} Pr[U] &\leq \sum_{i=1}^{B} Pr[X_i \leq \frac{logB}{c^*}] \\ &= \sum_{i=1}^{B} Pr[X_1 \leq \frac{\log B}{c^*}] \\ &= B\frac{\frac{\log B}{c^*}}{B^{1+\epsilon'}} \\ &= \frac{\frac{\log B}{c^*}}{B^{\epsilon'}} = o(1) \end{aligned}$$

Therefore, with high probability every bin has at least $\frac{\log B}{c^*} = \Theta(\frac{n}{B})$ balls. Now, clearly, choosing $n > cB\log B$ can only increase the probability that every bin has at least $\frac{\log B}{c^*} = \Theta(\frac{n}{B})$ balls. So, we are done. □

7.4 A Demonstrative Example: The 2D Grid

Let G be a $\sqrt{n} \times \sqrt{n}$ $2-$dimensional grid of n nodes[12], and for convenience denote $s = \sqrt{n}$. Let u_0 be the point at the origin (i.e $(0,0)$) and u_n the point at $(\sqrt{n}, \sqrt{n})$.

Claim 14 *The resistance between u_0 and u_n in G is $\Theta(\log n)$*

Proof: The proof is similar to that for the mesh [10] (6.3 , 8.7), and we state parts of it in a way that is generalizable to other proofs.

The upper bound is found using Thomson Principle and by constructing a unit flow c from u_o to u_n, where u_0 is the only source and u_n is the only sink. Let V_l denote the set of nodes at Manhattan distance (i.e. L_1 distance) l from the origin u_0. Note that $|V_l| = l+1$. And, let E_l denote the set of edges between V_l and V_{l+1}. Note that, for any l, the set V_l can be viewed as a contour of equi-potential nodes. So, we refer to V_l as *layer l*, and the flow that we construct can be seen as progressing uniformly through the E_l's.

The flow **out** of a point $(x,y) \in V_l$ (where $s > l \geq 0$) from u_0 is $\frac{x+1}{(l+2)(l+1)}$ to point $(x+1,y)$ and $\frac{y+1}{(l+2)(l+1)}$ to point $(x,y+1)$ (Fig. 3 (A)). For points $(x,y) \in V_l$ where $2s \geq l > s$ from u_0 the flow **incoming** from $(x-1,y)$ is $\frac{s-x+1}{(2s-l+2)(2s-l+1)}$ and $\frac{s-y+1}{(2s-l+2)(2s-l+1)}$ from the point $(x,y-1)$ (Fig. 3 (B)). Clearly, the flow out of u_0 is 1 and the flow into u_n is 1. Moreover, it is easy to check that the net flow at each point $(x,y) \neq u_0, u_n$ is 0.

Now by Thomson Principle we have $R_{u_0u_n} \leq P(c) = \sum_{e \in E} r(e)c^2(e)$.

We note that when $l < s$, $|E_l| = 2(l+1)$, and $\forall e \in E_l, c(e) \leq \frac{1}{1+1}$. When $2s \geq l > s$, letting l_u be the distance from u_n (i.e $l_u = 2s - l$), $|E_{l-1}| = 2(l_u+1)$, and $\forall e \in E_{l-1}, c(e) \leq (1/(l_u+1))$. So,

$$
\begin{aligned}
R_{u_0u_n} \leq P(c) &= \sum_{e \in E} \rho(e)c^2(e) \\
&= \sum_{l=1}^{s-1} \sum_{e \in E_l} O(1/l^2) + \sum_{l=s+1}^{2s} \sum_{e \in E_{l-i}} O(1/l_u^2)) \\
&= \sum_{l=0}^{s-1} O(l)O(1/l^2) + \sum_{k=0}^{s-1} O(k)O(1/k^2)) \\
&= 2\sum_{l=0}^{s-1} O(1/l) \\
&= O(\log n)
\end{aligned}
$$

A more general way to view this is as follows: Note the following in line with the intuition on equi-potential contours: For any layer l and any node $v \in V_l$, the flow into v equals the flow out of v which equals $\frac{1}{|V_l|}$ (or $\frac{1}{|V_{l_u}|}$ if $s \leq l \leq 2s$). Moreover, the total flow into V_l equals the total flow out of V_l equals 1. Note further that, optimally, to minimize power, for any fixed layer l, flow would be distributed almost uniformly over the edges E_l. So, letting δ_l denote the minimum degree of any node $v \in V_l$ in E_l, we would have a flow c and $e \in E_l$ s.t

$$
|E_l| = \Theta(\delta_l |V_l|) \;\; \text{and} \;\; c(e) = O(1/|E_l|) \tag{5}
$$

[12] We consider the grid, not the torus. To the best of our knowledge, all previous work on meshes has only considered the torus.

For example, in the case of the $2D$ Grid, for any layer l, we get $\delta_l \geq 1$ by considering border edges. This satisfies 5 because $|E_l| \approx 2|V_l| = \Theta(l)$ and for $e \in E_l$ $c(e) = O(1/l)$. Now, letting l_{max} be the maximum layer in the construction, for any flow satisfying 5, we may write the equation for power as follows:

$$
\begin{aligned}
R_{u_0 u_n} \leq P(c) &= \sum_{e \in E} \rho(e) c^2(e) \\
&= \sum_{l=1}^{l_{max}} \sum_{e \in E_l} O(\frac{1}{|E_l|^2}) \\
&= O(\sum_{l=1}^{l_{max}} \frac{1}{|E_l|}) \\
&= O(\sum_{l=1}^{l_{max}} \frac{1}{|V_l|\delta_l})
\end{aligned}
\tag{6}
$$

In particular, for the case of the $2D$ Grid, this immediately yields:

$$R_{u_0 u_n} \leq P(c) = O(\sum_{l=1}^{l_{max}} \frac{1}{2l+2}) = O(\log n)$$

The lower bound is by the "Short/Cut" Principle. For any $l \geq 0$, shorting all vertices in V_l to each other we get the following:

$$
\begin{aligned}
R_{u_0 u_n} &\geq \frac{1}{2} + \frac{1}{4} + \dots \frac{1}{2(s-1)} + \frac{1}{2(s-1)} \dots + \frac{1}{2} \\
&= 2 \sum_{l=1}^{s-1} \frac{1}{2(s-1)} \\
&= \Omega(\log n)
\end{aligned}
$$

□

Noting that in this case $m \approx 2n$ and using (1) we get that $C_G = O(n \log^2 n)$