Mobile Sensor Networks:
Bounds on Capacity and Complexity of Realizability

YIZHEN CHEN

Abstract

In a restricted combinatorial mobile sensor network (RCMSN), there are n sensors that continuously receive and store information from outside. Every two sensors communicate exactly once, and at an event when two sensors communicate, they receive and store additionally all information the other has stored. We study the capacity of information diffusion in mobile sensor networks, which was proposed by C. Gu, I. Downes, O. Gnawali, and L. Guibas. They collected all information received by two sensors between a communication event and the previous communication events for each of them into one information packet, and considered the number of sensors a packet eventually reaches. Then they defined the capacity of an RCMSN to be the ratio of the average number of sensors the packets reach and the total number of sensors. While they have studied the expected capacity of an RCMSN (when the order of communications is random), we found the RCMSNs with maximum and minimum capacities. We also studied similar problems for several related mobile sensor network constructions, such as ones generated from intersections of lines, as well as complexity results concerning when a mobile sensor network can be generated in such ways.

Contents

§1. Introduction

We consider n *sensors* that continuously receive and and store information from outside. At a *communication event* of two sensors, they receive additionally all information the other has stored. In this paper we study *combinatorial mobile sensor networks* (CMSN), in which the communication events and their order are to be specified. In order to evaluate the capacity of information diffusions in CMSN, the authors of paper [1] collect into an *information packet* all information received by a sensor between two successive communication events, and merge the information packets from the two sensors at a communication event, as afterwards the information in the packet always goes together. Thus, if there are k communcation events in total, then k information packets are generated. In an ideal situation, all sensors receive all packets eventually. So the authors of [1] define the *capacity* as the ratio of the actual number of packet deliveries and kn, or more rigorously as the following.

Definition ((R)CMSN)**.** A *restricted combinatorial mobile sensor network* (RCMSN) of n sensors, numbered by $1, 2, \ldots, n$, is a sequence of $\binom{n}{2}$ distinct *packets*, which are two-element sets $\{x, y\}$ where $1 \le x < y \le n$. If we denote this sequence by $a_1, \ldots, a_{n(n-1)/2}$, then we say that a packet a_{k_1} *can reach sensor* x ($1 \le x \le n$) if there exist $1 < k_2 < \cdots < k_m \le \binom{n}{2}$ such that $k_2 > k_1 \ge 1$, $a_{k_j} \cap a_{k_{j-1}} \ne \emptyset$ for each $2 \le j \le m$, and $x \in a_{k_m}$. We say a_{k_1} reaches x *in* $m-1$ *hops* if m is the minimum number that makes this condition hold. A *combinatorial mobile sensor network* (CMSN) is like an RCMSN, except that the length L of the sequence is arbitrary and the packets may be the same.

Definition (Capacity of (R)CMSN)**.** If the length of an RCMSN or CMSN of n sensors is L (we clearly have $L = \binom{n}{2}$ for an RCMSN), we define the *capacity* of the RCMSN or CMSN to be

$$\frac{1}{nL} \sum_{k=1}^{L} \#\{1 \le x \le n \mid a_k \text{ can reach } x\}.$$

To provide a comparison with RCMSN, we define the *absolute capacity* of a CMSN to be

$$\frac{1}{n\binom{n}{2}}\sum_{k=1}^{L}\#\{1 \le x \le n \mid a_k \text{ can reach } x\}.$$

We also study a specific type of RCMSN generated by the intersections of n lines, also considered in detailed by [1], defined as the following.

Definition (GMSN)**.** A *geometric mobile sensor network* (GMSN) of n sensors is an RCMSN of n sensors such that there exist n lines on an Euclidean plane where each line has a finite and distinct slope, no two intersections have the same x-coordinate, and the RCMSN is given by sorting $\{\{i,j\} \mid 1 \le i < j \le n\}$ according to the x-coordinate of the intersection between line i and line j in ascending order.

A variant of the GMSN is restricted GMSN (RGMSN), in which the number of slopes is limited:

Definition (RGMSN)**.** A *restricted geometric mobile sensor network* (RGMSN) of n sensors and k slopes is a CMSN of n sensors such that there exist n lines on an Euclidean plane with at most k distinct slopes, all of them finite, where no two intersections have the same x-coordinate, and the CMSN is given by sorting $\{\{i,j\} \mid \text{line } i \text{ and line } j \text{ intersect}\}$ according to the x-coordinate of the intersection between line i and line j in ascending order.

We note that the geometric mobile sensor network can be considered as the one-dimensional model where all sensors move with constant velocity, and they communicate whenever they meet. In restricted GMSNs, the velocities can only be selected from a few possible values.

For random RCMSNs, the authors of [1] claimed that when the size of the network approaches infinity, the expected capacity approaches one with high probability. However, their proof contains an error, which has been pointed out in Geneson's paper [2]. We propose a conjecture which, if proven, would fill the gap. Nevertheless, Geneson [2] gave the expected capacity of a random CMSN. Also, we found the maximum and minimum capacities of RCMSNs, which coincide with the maximum and minimum capacity of geometric MSNs. These results also have been shown before by Geneson [2], but we give simpler proofs in this paper (§2).

For random geometric MSNs, the authors of [1] found that the expected capacity does not approach one when the network size approaches infinity; instead, it is between $2/3$ and $5/6$. In this paper, we show in §2 that the expected capacity of a random GMSN when the size approaches infinity is exactly $5/6$.

Following the suggestion by the author of [2], in this paper we also study the restricted geometric MSNs. We give the maximum absolute capacity of RGMSNs with any number of slopes in §3, as well as an asymptotic formula for the maximum capacity of RGMSNs with any number of slopes. For the special case of RGMSNs with only two slopes, the capacity depends only on the size. With only three slopes, we give the exact expressions of the expected and maximum capacity (§4). With only four slopes, we also give the exact expression of the maximum capacity (§5). The remaining exact expressions (for expected capacity with ≥ 4 slopes, or maximum capacity with ≥ 5 slopes) are still left open.

Finally, we consider a new question about whether a given RCMSN or CMSN can be "realized" as a GMSN or RGMSN with a given number of slopes. We show in §6 that deciding if a given RCMSN can be realized as a GMSN is NP-Hard, using Shor's result [3] on the NP-Hardness of stretchability of pseudolines. However, the decision problem of whether a given CMSN can be realized as an RGMSN with s slopes is easier, and we give an algorithm (§6) to compute it in polynomial time with respect to the size of the network, although the complexity is exponential in variable s (thus it is still polynomial when s is bounded independent of network size).

§2. Geometric and restricted combinatorial MSNs

We first consider the minimum capacity of RCMSNs. We use the idea in [1].

Theorem 1 (RCMSN/GMSN Min. Capacity [2])**.** *The minimum possible capacity of an RCMSN or GMSN with n sensors is $2(n+1)/(3n)$.*

Proof. Clearly, any pair $\{i,j\}$ can reach both i and j, resulting in $2\binom{n}{2}$ deliveries. Any ordering of three pairs whose union contains only three distinct elements can always be expressed as $\{i,j\},\{j,k\},\{k,i\}$ in this order for distinct i,j,k. Here $\{i,j\}$ can reach k and $\{j,k\}$ can reach i, resulting in $2\binom{n}{3}$ deliveries. Therefore the minimum capacity is at least

$$\frac{2\binom{n}{2}+2\binom{n}{3}}{n\binom{n}{2}} = \frac{2(n+1)}{3n}.$$

Consider the lines $l_k : kx+(n+1-k)y = k(n+1-k)$ for $1 \le k \le n$. It can be verified that if $\max\{i,j\} < k$, then $\{i,j\}$ cannot reach line k, so the capacity of this GMSN is at most

$$\frac{1}{n\binom{n}{2}}\sum_{k=1}^{n} k(k-1) = \frac{2(n+1)}{3n}.$$

Therefore the exact minimum capacity is $\frac{2(n+1)}{3n}$. □

Now we consider the maximum capacity of RCMSNs. The idea is to consider those clearly unreachable sensors, and to construct a GMSN with those the only unreachable sensors.

Theorem 2 (RCMSN/GMSN Max. Capacity [2])**.** *The maximum possible capacity of an RCMSN or GMSN with n sensors is $2(n+1)/(3n)$.*

Proof. Let the sequence of communications be $\{a_1,b_1\},\{a_2,b_2\},\ldots,\{a_m,b_m\}$ where $m=\binom{n}{2}$. If a graph is formed with vertices $\{1,\ldots,n\}$ and undirected edges $(a_k,b_k),(a_{k+1},b_{k+1}),\ldots,(a_m,b_m)$, then the connected component containing (a_k,b_k) has at most $m-k+1$ edges and thus at most $m-k+2$ vertices. Therefore, a_k,b_k can reach at most $m-k+2$ nodes, and the total number of deliveries of a GMSN is no more than

$$\sum_{k=1}^{m-n+2} n + \sum_{k=m-n+3}^{m} (m-k+2) = n(m-n+2) + \frac{n^2-n-2}{2} = \frac{(n-1)(n^2-n+2)}{2}.$$

So the maximum capacity is at most

$$\frac{(n-1)(n^2-n+2)/2}{n^2(n-1)/2} = 1-1/n+2/n^2.$$

Now we place $n-2$ nonvertical and pairwise nonparallel lines randomly, put a line with a slope less than the minimum slope of the previous $n-2$ lines to the right of all previous intersections, and then put a line with a slope greater than the maximum of the previous $n-1$ lines to the right of all previous intersections. Then every intersection of the first $n-1$ lines obviously can reach all lines, and the last $n-1$ intersections can reach, from left to right, $n, n-1, \ldots, 2$ lines. This GMSN reaches our upper bound. □

In this proof, the construction used is called a *collector-distributor* construction [1, 2]. Similar constructions will be used in the next section.

What about the expected capacity of RCMSNs? It is claimed in [1] that the expected capacity of an RCMSN with n sensors is $1-O(\log^2 n/n)$. However, their proof is incorrect [2]. They partitioned the $\binom{n}{2}$ communications into groups of size $\lceil n\log n\rceil$, found that the probability that all n sensors appear in one given group is more than $1-1/n$, and concluded that with high probability, all groups contain all n sensors. But this is not true as the events of a sensor appearing in the groups are not independent. In order for their argument to hold, the following conjecture must be proven.

Open Question. If all pairs (i, j) where $1 \le i < j \le n$ are partitioned into groups of size $\lceil n \log n \rceil$, an expected proportion $1 - O(\log n/n)$ of the groups contain all numbers $1, \dots, n$.

In [1] it is also showed that the expected capacity of a random GMSN is no more than 5/6 when the network size approaches infinity, considering some lines with high or low slopes that clearly cannot be reached by intersections in certain regions. Here, we improve the result by directly computing the expected capacity using Wolfram Mathematica.

Theorem 3 (Expectation for GMSN)**.** *The expected capacity of a GMSN with n sensors is $5/6$ when $n \to \infty$.*

Proof. To randomly form a GMSN with n sensors, we independently choose n pair (a_k, b_k) from the uniform distribution in $[a^-, a^+] \times [b^-, b^+]$, then form the GMSN using the lines $l_k : y = a_k x + b_k$. By applying scaling, shear transformation, and translation, we may assume that $a^- = b^- = 0$ and $a^+ = b^+ = 1$ without changing the probability of each possible GMSN. The process of randomly choosing n lines can also be considered as the following process: first randomly choose two lines $l : y = ax + b$ and $l' : y = a'x + b'$, then randomly choose a set A of εn lines ($0 < \varepsilon < 1$), then randomly choose a set B of $(1 - \varepsilon)n - 2$ lines. We want to count the number of triples $(\{u, v\}, w)$ of lines, among all $n\binom{n}{2}$ such triples, such that $u \cap v$ can reach w. Theorem 3.4 in [1] shows that exactly $2\binom{n}{2} + 2\binom{n}{3}$ triples satisfy that $u \cap v$ can reach w in at most one hop. Therefore, we need only count the expected number of triples where $u \cap v$ can and only can reach w in exactly two hops, according to Theorem 3.8 of the same paper.

In set A, it is expected that

$$2 \int_0^1 \int_0^1 \int_a^1 \int_0^1 \int_{a'}^1 \max\left\{0, \min\left\{1, b + (b' - b) \cdot \frac{a - a''}{a - a'}\right\}\right\} \mathrm{d}a'' \, \mathrm{d}b' \, \mathrm{d}a' \, \mathrm{d}b \, \mathrm{d}a = \frac{1}{6}$$

of the lines (in the form $y = a''x + b''$) can be reached by $l \cap l'$ and has slope higher than that of both l and l'. Among these lines, the line l^* with the maximum slope is expected to have slope, when $n \to \infty$, equal to the upper limit of its range, which is one when $l \cap l'$ is not in the first quadrant, and $\min\{1, \mathrm{slope}(O(l \cap l'))\}$ when it is. In set B, the following expression counts the proportion of lines (i) with slope higher than that of both l and l', (ii) cannot be reached by $l \cap l'$ in at most one hop, but (iii) can be reached by $l \cap l'$ in two hops via l^*:

$$2\left(\int_0^1 \int_0^1 \int_a^1 \left(\int_b^1 \int_{a'}^1 + \int_0^b \int_{a'}^{\min\{1, \frac{a'b - ab'}{b - b'}\}}\right) \max\left\{0, \min\left\{1, 1 - b - (b' - b) \cdot \frac{a - a''}{a - a'}\right\}\right\} \mathrm{d}a'' \, \mathrm{d}b' \, \mathrm{d}a' \, \mathrm{d}b \, \mathrm{d}a\right) = \frac{1}{12}.$$

We can do a similar reasoning for lines with slope lower than that of both l and l'. When $\varepsilon \to 0$, we found that for every randomly chosen two lines l and l' among the n lines to be chosen, it is expected that $(n-2)/6$ triples $(\{l, l'\}, l'')$ require exactly two hops. By linearity of expectation, in total it is expected that $\binom{n}{2}(n - 2)/6$ triples require exactly two hops, so the expected capacity is

$$\frac{2\binom{n}{2} + 2\binom{n}{3} + \binom{n}{2}(n - 2)/6}{n\binom{n}{2}} = \frac{5}{6} + o(1).$$

□

Although we know the capacity of GMSNs with size approaching infinity, how concentrated they are is still left open.

Open Question. What is the variance of the capacity of a GMSN with n sensors when $n \to \infty$?

§3. General results for restricted geometric MSNs

Next, we consider the bounds on the capacity of RGMSNs. Clearly, the minimum capacity of any RGMSN is zero because we can always make all lines parallel. So we only find the maximum capacities. When limited to only two slopes, the geometric form of any RGMSN is a grid, so the capacity can be directly computed and is also easy, as shown by the following theorem.

Theorem 4 (2-Slope RGMSN Max. Capacity)**.** *The capacity for any RGMSN of n sensors and 2 slopes is $(n+1)/(2n)$. The expected and maximum absolute capacities of an RGMSN of n sensors and 2 slopes are $(n+2)/(4n-4)$ and $(n+2)/(4n)$, respectively.*

Proof. Suppose there are m lines with the larger slope and $n-m$ lines with the smaller slope. Without loss of generality, suppose all slopes are positive. We order the lines with the larger slope by their x-intercepts from the largest to smallest and label them $a_1, a_2, \ldots, a_m$, and order and label the lines with the smaller slope in the same way as $b_1, b_2, \ldots, b_{n-m}$. Clearly, the intersection $a_i \cap b_j$ can reach lines $a_1, \ldots, a_i$ and $b_1, \ldots, b_j$ but no other lines. Therefore, the capacity of the RGMSN is

$$\begin{aligned}\frac{1}{nm(n-m)} \sum_{i=1}^{m} \sum_{j=1}^{n-m} (i+j) &= \frac{1}{nm(n-m)} \cdot \left((n-m)\left(\sum_{j=1}^{m} j \right) + m\left(\sum_{j=1}^{n-m} j \right) \right) \\ &= \frac{1}{2} + \frac{1}{n}.\end{aligned}$$

The absolute capacity of the RGMSN is

$$\frac{m(n-m)}{\binom{n}{2}} \left(\frac{1}{2} + \frac{1}{n} \right) = \frac{m(n-m)(n+2)}{n^2(n-1)},$$

which clearly has maximum $(n+2)/(4n-4)$ and expected value

$$2^{-n} \sum_{m=0}^{n} \binom{n}{m} \frac{m(n-m)(n+2)}{n^2(n-1)} = 2^{-n} \cdot \frac{n+2}{n} \sum_{m=1}^{n-1} \binom{n-2}{m-1} = \frac{n+2}{4n}.$$

□

When there are three or more slopes, we can use a strategy similar to that in Theorem 2: trying to find as many unreachable lines as possible, and finding an example that attains the upper bound. While Theorem 2 only considers the last $n-1$ intersections, we need to consider the lines with the largest and smallest slopes when the number of slopes is limited.

We first deal with the easier absolute capacity, where the maximum is asymptotically less than one:

Theorem 5 (RGMSN Max. Abs. Capacity)**.** *The maximum absolute capacity of an RGMSN of n sensors and $s \geq 3$ slopes is asymptotically*

$$\frac{(135s^3 - 945s^2 + 2232s - 1796) + 4(9s^2 - 42s + 52)^{3/2}}{243(s-2)^3} + o(1).$$

Proof. We consider the intersections that lie on the a lines with the largest slope k and the b lines with the least slope k'. For a line whose slope is not equal to k or k', its any intersection with a line with slope k or k' cannot reach all lines with slopes k or k' to the left of it because of the maximality of k and k' (for if it reaches a line to the left, it can only do so via a line with higher absolute value of slope; but there are no such lines). Thus, for each of these $n-a-b$ lines, the intersections mentioned above cannot reach at least $1 + 2 + \cdots + (a+b-1)$ lines in total. For intersections of the lines with slopes k or k', the total number of lines they cannot reach is at least $(0+1+\cdots+(b-1)) + (1+2+\cdots+b) + \cdots + ((a-1)+a+\cdots+(a+b-2))$.

Suppose the slopes other than k and k' are $k_1, \ldots, k_{s-2}$, and the corresponding numbers of lines are $c_1, \ldots, c_{s-2}$, respectively. Then the total number of information deliveries is at most

$$n\left(ab + (a+b)\sum_{j=1}^{s-2} c_j + \sum_{i \neq j} c_i c_j \right) - (n-a-b) \sum_{j=1}^{a+b-1} j - \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} (i+j).$$

This is maximized when all the c_j's are equal and $a = b$; after that, we have only one free variable in the sum:

$$f(a) = n\left(a^2 + 2a(n-2a) + \frac{(s-2)(s-3)}{2} \left(\frac{n-2a}{s-2} \right)^2 \right) - (n-2a)a(2a-1) - a^2(a-1).$$

From this it is easy to find by differentiation that the maximum sum divided by $n\binom{n}{2}$ is equal to the formula given in the theorem statement when $n \to \infty$. It is not difficult to compute that the collector-distributor construction attains this upper bound—the formulas are almost entirely the same. $\square$

For the maximum capacity of an RGMSN, we need to consider both the largest/smallest slopes and the last $n-1$ intersections. The proof is essentially similar to that in Theorem 5.

Theorem 6 (RGMSN Max. Capacity)**.** *The maximum capacity of an RGMSN of n sensors and $s \geq 3$ slopes is asymptotically*

$$\begin{cases} 1 - \frac{s-2}{s-3}\frac{1}{n} + \Theta(n^{-2}) & s \geq 4 \\ 1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \Theta(n^{-3/2}) & s = 3. \end{cases}$$

Proof. Consider the a lines with the largest slope and the b lines with the smallest slope. They form a grid. (Note: the word *grid* below will simply mean these lines in the first two slopes.) Suppose, without loss of generality, that the largest slope is 1 and the smallest slope is -1. The intersections of lines in the grid cannot reach at least $\sum_{i=0}^{a-1}\sum_{j=0}^{b-1}(i+j)$ lines. The grid can also be seen as a set of $a+b$ curves, where the kth curve consists of those points on the gridlines such that there are $k-1$ gridlines on their right. Each of these curves intersects with all the remaining $n-a-b$ lines with slope $\neq \pm 1$; these intersections cannot reach at least $(n-a-b)(1+2+\cdots+(a+b-1))$ gridlines. Also, in the arrangement, the kth rightmost intersection obviously can reach at most $k+1$ lines (by induction), so the $n-1$ rightmost intersections cannot reach at least $1+2+\cdots+(n-2)$ lines. But we may have counted some intersections' unreachable lines twice. At worst, we overestimated it by $(n-1)(a+b)$ unreachable lines. Therefore, the maximum capacity with only s slopes is at most

$$1 - \frac{(1+2+\cdots+(n-2)) + (n-a-b)(1+2+\cdots+(a+b-1)) + \sum_{i=0}^{a-1}\sum_{j=0}^{b-1}(i+j) - (n-1)(a+b)}{n(ab+(a+b)(n-a-b)+\frac{(s-2)(s-3)}{2}(\frac{n-a-b}{s-2})^2)}.$$

To give an upper bound to this formula, we need only compute the minimum value of

$$\begin{aligned}
&\frac{(1+2+\cdots+(n-2)) + (n-a-b)(1+2+\cdots+(a+b-1)) + \sum_{i=0}^{a-1}\sum_{j=0}^{b-1}(i+j) - (n-1)(a+b)}{ab+(a+b)(n-a-b)+\frac{(s-2)(s-3)}{2}(\frac{n-a-b}{s-2})^2} \\
&= \frac{\frac{1}{2}(n-1)(n-2) + \frac{1}{2}(n-a-b)(a+b)(a+b-1) + \frac{1}{2}ab(a+b-2) - (n-1)(a+b)}{ab+(a+b)(n-a-b)+\frac{s-3}{2(s-2)}(n-a-b)^2} \\
&= \frac{\frac{1}{2}(n-a-b-1)(n-2) - \frac{1}{2}(a+b)^2 - \frac{1}{4}(1-\frac{1}{s-2})(a+b-2)(n-a-b)^2}{ab+(a+b)(n-a-b)+\frac{s-3}{2(s-2)}(n-a-b)^2} + \frac{a+b-2}{2} =: \kappa
\end{aligned}$$

When $s \geq 4$, we have

$$\begin{aligned}
\kappa &\leq \frac{\frac{1}{2}(n-a-b-1)(n-2) - \frac{1}{2}(a+b)^2 - \frac{1}{8}(a+b-2)(n-a-b)^2}{ab+(a+b)(n-a-b)+\frac{s-3}{2(s-2)}(n-a-b)^2} + \frac{a+b-2}{2} \\
&= \frac{\frac{1}{8}(n-a-b)(4(n-2)-(a+b-2)(n-a-b)) - \frac{1}{2}(n-2+(a+b)^2)}{ab+(a+b)(n-a-b)+\frac{s-3}{2(s-2)}(n-a-b)^2} + \frac{a+b-2}{2},
\end{aligned}$$

which is minimized either when $5 < a < n-6$ and $b = 1$ or when $a = b$ and $a + b \notin (6, n-5)$. In each case, the formula then has only one variable and an upper bound of the capacity is $1 - \frac{s-2}{s-3}\frac{1}{n} + \Theta(n^{-2})$ when $s \geq 4$. Again, it is not difficult to compute that the collector-distributor construction attains this upper bound—the formulas are almost entirely the same.

When $s = 3$, we have

$$\kappa = \frac{(n-a-b-1)(n-2)-(a+b)^2}{2(ab+(a+b)(n-a-b))} + \frac{a+b-2}{2},$$

which is minimized either

$$\text{when } a = b < \frac{\sqrt{5n^2 - 16n + 12} - (n-2)}{4} \text{ or when } b = 1 \text{ and } a > \frac{\sqrt{5n^2 - 16n + 12} - n}{2}.$$

In each case, the formula then has only one variable and an upper bound of the capacity is $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \Theta(n^{-3/2})$. However, the collector-distributor construction does not attain the bound here. That is because unless the previous cases, neither a nor b is $O(1)$ when the capacity formula is maximized. Instead, we use the following construction. Let $1 \le a \le \lfloor (n-1)/2 \rfloor$ be an integer whose value will be determined later, and $b = n - 2a$. In $\mathbb{R}^2$ we draw the b lines $y = k$ ($0 \le k \le b-1$), the $a-1$ lines $x = k$ ($1 \le k \le a-1$), the a lines $x + 2y = a + 2(b+k) - 3$ ($1 \le k \le a$), and the line $x = 3a + 2b$, and rotate all lines by $\pi/4$ clockwise about the origin (Fig. 1). Then we get an RGMSN with 3 slopes whose capacity is

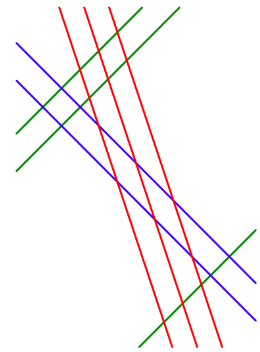

Figure 1. An RGMSN of 3 slopes. There are $a = 3$ red lines, $a = 3$ green lines, and $b = 2$ blue lines here.

$$\frac{1}{n(a^2+2ab)}\left(\sum_{k=2}^{b+1} k + (b+1)\sum_{k=b+2}^{a+b+1} k + (a-1)\sum_{k=b+2}^{a+b+1} k + a\sum_{k=1}^{a-1} k + b\sum_{k=a+b+2}^{2a+b} k\right)$$
$$= \frac{a+b+1}{n} - \frac{n-1}{2n}\frac{b}{a^2+2ab} =: f(a,b).$$

This can attain the upper bound (computed with basic differential calculus) $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \Theta(n^{-3/2})$ we found when a is allowed to be non-integers. When a must be an integer, the deviation from the upper bound is at most

$$f(a,b) - f(a + 1/2, b-1) = \frac{1}{2n\sqrt{n}} + \Theta(n^{-2}),$$

so this construction attains the upper bound. $\square$

§4. Three-slope geometric MSNs

Theorem 6 only gives the capacity of near-optimal constructions, not the optimal construction. We have found a method to produce optimal constructions for RGMSNs of no more than four slopes, although it does not generalize to higher numbers of slopes. We begin with RGMSNs with at most three slopes. The method is to move lines with the third slope in the "grid" formed by existing lines in two slopes.

Remark. In the proofs below, a *grid* is a collection of lines with exactly two different slopes. A *corner* of a grid is an intersection from which there exists a ray that intersects no other lines in the grid. A *path* is a line that intersects the grid. The intersection where a path *enters* or *leaves* a grid is one next to a corner, and the former one can reach the latter one but not vice versa.

Theorem 7 (3-Slope RGMSN Max. Capacity)**.** *The maximum capacity of an RGMSN of n sensors with only three possible slopes allowed is equal to*

$$\max\left\{\frac{(4p-1)n^2 - (10p^2 - 6p - 1)n + (6p^3 - 6p^2 - 2p)}{4pn^2 - 6p^2 n}, \frac{(4q+1)n^2 - (10q^2 + 4q - 1)n + (6q^3 + 3q^2 - 3q - 2)}{(4q+2)n^2 - (6q^2 + 6q + 2)n}\right\},$$

where $q = \lfloor \sqrt{n}/2 \rfloor$ and either $p = q$ or $p = q + 1$. This is asymptotically $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + r_n n^{-3/2}$, where $\liminf_{n\to\infty} r_n = \frac{15}{32}$ *and* $\limsup_{n\to\infty} r_n = \frac{19}{32}$.

Proof. Suppose there are a lines in slope k_1, b lines in slope k_2, and c lines in slope k_3, where $a + b + c = n$. Without loss of generality, suppose the rightmost intersection lies on lines with slopes k_1 and k_2. All lines with those two slopes form an a-by-b grid. Now we investigate where we should put the remaining c lines in between so as to maximize the capacity.

Because the intercepts of the lines in the grid are not determined, the new lines can be put anywhere as long as it does not contradict the relations of slopes. Thus, any new line placed into the grid must start outside a corner adjacent to the corner of the rightmost intersection and end in the opposite corner; it must go through a sequence of adjacent cells without going in the opposite direction of how it went before. Equivalently, it intersects any existing line exactly once. For example, the drawing below (Fig. 2) shows two valid paths and two invalid paths. All diagrams in this proof will uniquely determine an arrangement because all four corners of the grid are assigned different roles: the corner of the rightmost intersection, the corner opposite to it, the starting corner, and the ending corner.

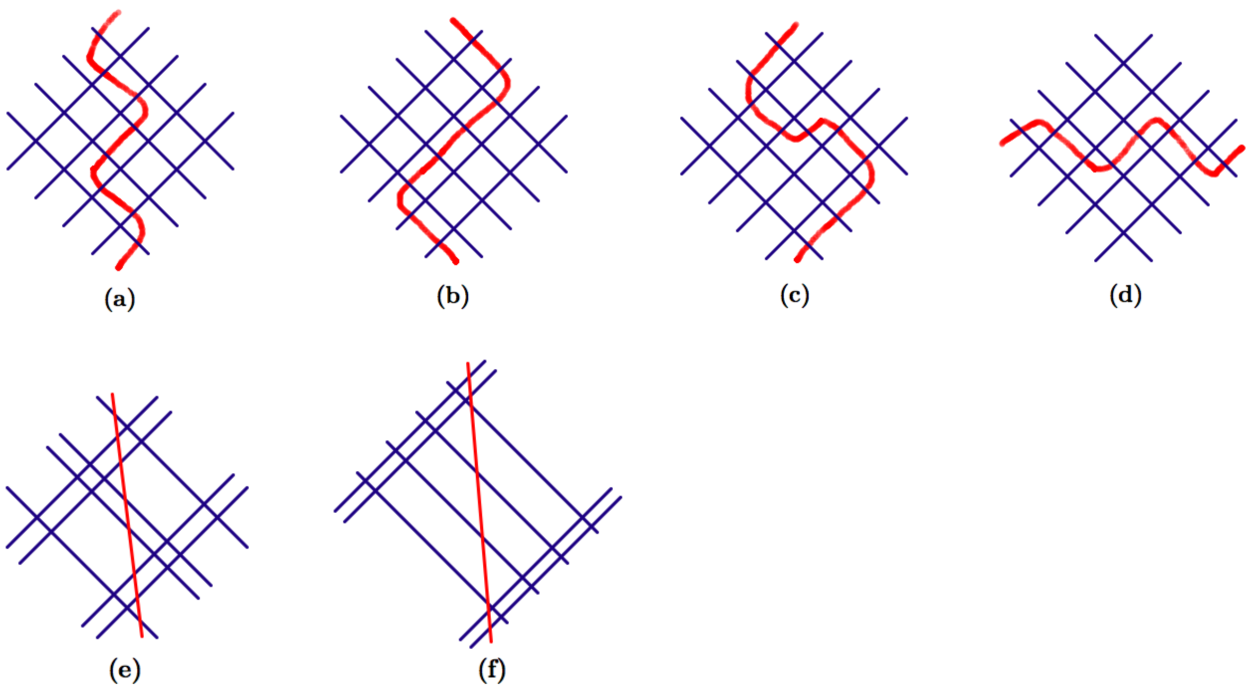

Figure 2. **(a)** and **(b)** are valid paths. Path **(c)** is invalid because it intersects an existing line three times. Path **(d)** is invalid because it contradicts the given rightmost intersection, even if it can be pulled straight. Diagrams **(e)** and **(f)** show the realizations of **(a)** and **(b)** in straight lines, respectively.

We first find the best arrangement when $c = 1$. In all diagrams below in this proof, the new lines added have slope $k_3 < 0$ when pulled straight. Also, $k_3 < k_2 < 0 < k_1$. (Note that the only two possible cases are this and $k_2 < 0 < k_1 < k_3$, since otherwise the rightmost intersection would not come from lines with slopes k_1 and k_2.) In addition, we will use the following setting: the nearest line the starting corner can reach without adding new lines is labeled s_a; the next line is then labeled s_{a-1}, and so on. Without loss of generality, those lines are assumed to have slope k_1. From the ending corner we similarly label the lines it can reach $t_b, t_{b-1}, \dots, t_1$. Then, without adding new lines, $s_j \cap t_i$ can reach exactly $i + j$ lines. The new line is labeled l. Lastly, each intersection is labeled the number of lines it can reach.

Step 1. *Where should the path leave the grid?*

We have included a diagram ((a) and (b) in Fig. 3) to help understand the description below. Suppose we initially have a path that leaves the grid between t_i and t_{i+1} (that is, it intersects t_i, s_1, and t_{i+1} in this order, as shown in (a)). Now we modify the path and make it leave the grid between t_{i-1} and t_i, assuming that the path is still valid and the positions of other intersections relative to the grid are not changed. Then only the labels on $s_1 \cap t_i$ and $t_i \cap l$ change. Other intersections are not affected because either they cannot reach l and are thus unrelated to the change, or they can still travel through l to the lines they previously can reach. Note that although $l \cap s_1$ changes its relative position to the grid, its label does not change as it can still reach l and all t_j, but not other lines. The following table shows the changes in the two intersections' labels:

Intersection	Original	Reason	New	Reason
$s_1 \cap t_i$	$i+1$	it cannot reach l	$b+2$	it can reach all t_j through l and s_1
$t_i \cap l$	$b+2$	it can reach all t_j through l and $s_1 \cap t_i$, plus s_1	$b-i+2$	it can reach $t_i, \ldots, t_b$ through l

This table tells us that whether the new network has a larger number of information deliveries depends on whether $i > (b+1)/2$.

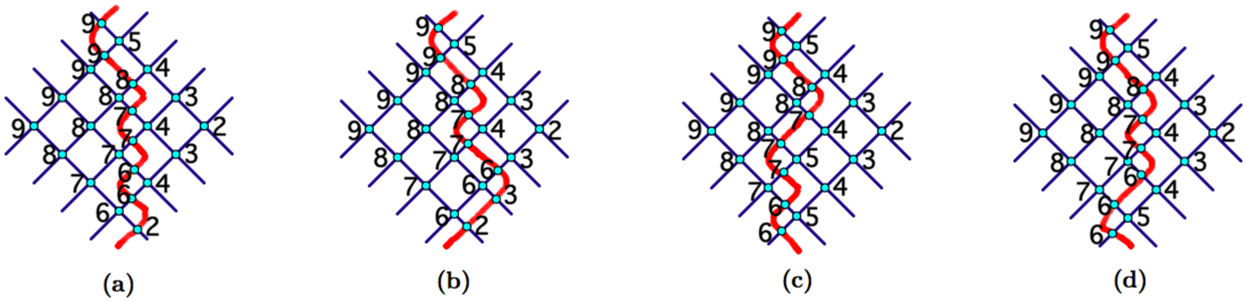

Figure 3. In **(a)** and **(b)**, the place that the path leaves the grid changes; in **(c)** and **(d)**, the relative position of the path and one intersection changes. When pulled straight, the new lines all have negative slope. The number of lines that can be reached from each intersection is labeled.

Step 2. *Where should the path cross the grid?*

Diagrams (c) and (d) in Fig. 3 are concerned here. Suppose we initially have a path that intersects t_i and s_j consecutively in this order ($j > 1$). Now we switch their order while leaving other intersections unchanged. Then, for reasons similar to those in the previous paragraph, only the labels on $s_j \cap t_i$ and $t_i \cap l$ change. All other intersections are unrelated. The following table shows the changes in the two intersections' labels:

Intersection	Original	Reason	New	Reason
$s_j \cap t_i$	$i+j$	it cannot reach l	$b+j+1$	it can reach all t_i through l and s_j, and also $s_1, \ldots, s_j$
$t_i \cap l$	$b+j+1$	it can reach all t_i through l and $s_j \cap t_i$, and also $s_1, \ldots, s_j$	$b+j$	it can reach all t_i through l and $s_{j-1} \cap t_i$, and also $s_1, \ldots, s_{j-1}$

This table tells us that the new network always has a larger or equal number of information deliveries.

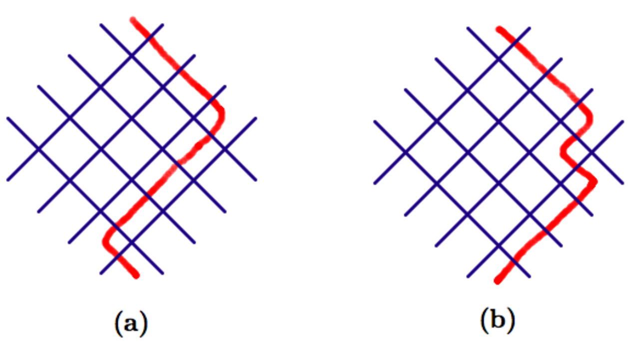

Figure 4. The best network and the second best network.

Given any valid path, one can always perform the two operations to transform it to one of the two paths shown in Fig. 4 above as follows: a path can be mapped to a string of `S` and `T` indicating at the current cell in the grid if the path crosses an s_j to another cell or a t_i to another cell. For example, the path in Fig. 2(a) is mapped to the string `TSSTSTST`. Now, the discussions above tell us that a substring `TS` can be replaced by `ST` without reducing capacity if there exists at least one `S` following this substring. Repeating this operation, any string can be transformed to the form `SS...STT...TSTT...T`. The discussions above also tell us that the last `S` may either be moved to the end or the beginning; one of the two choices does not reduce the capacity. Therefore any string is transformed to `SS...STT...TS` or `SS...STSTT...T` (without violating the rightmost intersection), which correspond to Fig. 4 (a) and (b) respectively.

Now we consider the case $c > 1$. For the rightmost line with slope k_3, we may apply the same discussions above to move it to one location in Fig. 4. Then, no matter how other lines with slope k_3 move, the total capacity does not change because each intersection on the left of l can reach l, and thus all the t_i's; it always can reach all the s_j's below it and no s_j's above it; and it always can reach all lines with slope k_3 on the right of it and none on the left. Moving a line with slope k_3 across an intersection only exchanges two intersections' labels.

It is easy to compute that the total number of information deliveries in Fig. 4(a) is

$$\sum_{k=2}^{b+1} k + (b+1)\sum_{k=b+2}^{b+c+1} k + (a-1)\sum_{k=b+2}^{b+c+1} k + c\sum_{k=1}^{a-1} k + b\sum_{k=b+c+2}^{b+c+a} k,$$

and the total number of information deliveries in Fig. 4(b) is

$$2 + c\sum_{k=1}^{b-1} k + (b-1)\sum_{k=1}^{c} k + (a+1)\sum_{k=b+2}^{b+c+1} k + c\sum_{k=1}^{a-1} k + (b-1)(b+c+1) + b\sum_{k=b+c+2}^{a+b+c} k.$$

The difference between the two sums is $(b-1)(b(c-1)+2) \geq 0$, so Fig. 4(a) has no less capacity than Fig. 4(b). The sum in Fig. 4(a) equals $((2b+a+c+1)(ab+bc+ca)+(ac+b-b^2))/2$, so the capacity to be maximized is

$$\frac{(2b+a+c+1)(ab+bc+ca)+(ac+b-b^2)}{2n(ab+bc+ca)} = \frac{2b+a+c+2}{2n} - \frac{n-1}{2n}\frac{b}{ab+bc+ca}.$$

If a, b, and c may not be integers, this is maximized when $a = c = \sqrt{n}/2 + O(n^{-1/2})$, so when the variables must be integers, each of a and c must be equal to p or $p+1$; this leads to the formula given in the theorem. For the asymptotic formula, we observe that the capacity has a maximum value equal to $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \frac{19}{32n\sqrt{n}} + \Theta(n^{-2})$, which is achieved when n is an even perfect square and $a = c = \sqrt{n}/2$, giving the limsup value. Also, when $a = c = \sqrt{n}/2 - t$ is an integer, the capacity is $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \frac{19-64t^2}{32n\sqrt{n}} + \Theta(n^{-2})$, and when $a = \sqrt{n}/2 - t$ and $c = \sqrt{n}/2 + 1 - t$ are integers, the capacity is $1 - \frac{1}{\sqrt{n}} + \frac{9}{8n} + \frac{3-64t(t-1)}{32n\sqrt{n}} + \Theta(n^{-2})$. If $\frac{19-64t^2}{32} < \frac{15}{32}$ we must have $|t| > \frac{1}{4}$, and in this case $\frac{3-64|t|(|t|-1)}{32} > \frac{15}{32}$, so we always have $r_n \geq \frac{15}{32}$; this liminf is achieved by $n = 4k^2 + 2k + 1$ for positive integers k. $\square$

Regarding RGMSNs with at most three slopes, we have also found their expected capacity for three slopes and all intercepts chosen uniformly in some interval. The idea for the proof is similar to the idea in Theorem 3 (expected capacity for GMSN)—identifying all possible terms that contribute to the capacity of the RGMSN. We note that the integrals here are also computed using Wolfram Mathematica like in Theorem 3.

Theorem 8 (Expectation for Three-Slope RGMSN)**.** *The expected capacity of an RGMSN with at most three slopes and n sensors is $23/36$ when $n \to \infty$.*

Proof. To randomly form an RGMSN with at most three slopes and n sensors, we first choose the three slopes independently from the uniform distribution in $[a^-, a^+]$. Let them be $u < v < w$ (the probability that two of them are equal is zero). Then we choose the slopes of the n lines independently from the three slopes with equal probability, and choose the intercepts of the n lines independently from the uniform distribution in $[b^-, b^+]$. Like in Theorem 3, we assume $a^- = b^- = 0$ and $a^+ = b^+ = 1$ without loss of generality. For $a+b+c = n$, the probability that there are a lines with slope u, b lines with slope v, and c lines with slope w is $3^{-n}\binom{n}{a,b,c}$. They form $ab+bc+ca$ intersections.

No matter where the lines are, any intersection of a line with slope u and a line with slope w can reach all lines with slope v. For other intersections, because for two independently randomly chosen lines l and l' (with random slope), a third independently randomly chosen line with slope not lower than that of both

l and l' has equal probability to be reachable or not reachable from $l \cap l'$, it is expected that $a/2$ lines with slope u are reachable from any intersection of a line with slope v and a line with slope w. Similarly, it is expected that $c/2$ lines with slope w are reachable from any intersection of a line with slope u and a line with slope v.

An intersection cannot reach a line with slope different from both lines that form the intersection because there are only three different slopes. Because u and w are the minimum and maximum slopes, it is also not possible for an intersection of lines with slopes u and w to reach another line with slope u or w if it cannot be reached in one hop. For an intersection of lines with slopes v and w, if it can reach a line with slope w, then it can do so in one hop because w is the largest slope; for lines with slope v, it can reach all of them if it can reach the rightmost line with slope u—the one with the largest intercept. The probability that this happens is

$$\frac{\int_0^1 \int_u^1 \int_v^1 \int_0^1 \int_0^1 \min\left\{1, \max\left\{0, 1 - q - (p - q)\frac{v-u}{v-w}\right\}\right\} \cdot ap^{a-1}\,\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}w\,\mathrm{d}v\,\mathrm{d}u}{\int_0^1 \int_u^1 \int_v^1 \int_0^1 \int_0^1 ap^{a-1}\,\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}w\,\mathrm{d}v\,\mathrm{d}u} = \frac{3a+1}{4(a+1)}.$$

Otherwise, it can only reach those lines with slope v that it can reach in one hop. Similarly, an intersection of lines with slopes u and v can only reach lines with slope u that it can reach in one hop as u is the smallest slope. However, if it can reach the rightmost line with slope w, then it can reach all lines with slope v. The probability that this happens is

$$\frac{\int_0^1 \int_u^1 \int_v^1 \int_0^1 \int_0^1 \min\left\{1, \max\left\{0, q + (p - q)\frac{v-u}{v-w}\right\}\right\} \cdot c(1-p)^{c-1}\,\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}w\,\mathrm{d}v\,\mathrm{d}u}{\int_0^1 \int_u^1 \int_v^1 \int_0^1 \int_0^1 c(1-p)^{c-1}\,\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}w\,\mathrm{d}v\,\mathrm{d}u} = \frac{3c+1}{4(c+1)}.$$

Otherwise, it can only reach those lines with slope v that it can reach in one hop. We have discussed all cases that require two hops, so by Theorem 3.8 in [1], the expected capacity is

$$\begin{aligned}\kappa = \lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \Bigg(acb + \sum_{i=1}^{a}\sum_{k=1}^{c}(i+k) \\ + \frac{bca}{2} + \sum_{j=1}^{b}\sum_{k=1}^{c}\left(\frac{3a-1}{4(a+1)}(b+k) + \frac{a+5}{4(a+1)}(j+k)\right) \\ + \frac{abc}{2} + \sum_{i=1}^{a}\sum_{j=1}^{b}\left(\frac{3c-1}{4(c+1)}(i+b) + \frac{c+5}{4(c+1)}(i+j)\right)\Bigg).\end{aligned}$$

This is equal to

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \Bigg(2abc + \frac{ac(a+c+2)}{2} \\ &+ \frac{3a-1}{4(a+1)}b\left(bc + \frac{c(c+1)}{2}\right) + \frac{a+5}{4(a+1)}\frac{bc(b+c+2)}{2} + \frac{3c-1}{4(c+1)}b\left(ba + \frac{a(a+1)}{2}\right) + \frac{c+5}{4(c+1)}\frac{ba(b+a+2)}{2}\Bigg) \\ &= \lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \Bigg(2abc + \frac{ac(a+c)}{2} \\ &+ \frac{3}{4}b\left(bc + \frac{c^2}{2}\right) + \frac{1}{4}\frac{bc(b+c)}{2} + \frac{3}{4}b\left(ba + \frac{a^2}{2}\right) + \frac{1}{4}\frac{ba(b+a)}{2}\Bigg)\end{aligned}$$

because the expected value of $1/a$ is zero when $n \to \infty$. Then this is equal to

$$\lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \left(2abc + \frac{5}{8}(ab(a+b)+bc(b+c)+ca(c+a))\right)$$

because the sum is unchanged when we average over all permutations of $\{a,b,c\}$. Then this is equal to

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \left(\frac{1}{8}abc + \frac{5}{8}(a+b+c)(ab+bc+ca)\right) \\
&= \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} \frac{1}{3^n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{1}{n(ab+bc+ca)} \\
&\leq \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{a+b+c}{9n} = \frac{5}{8} + \frac{1}{8}\cdot\frac{1}{9} = \frac{23}{36}.
\end{aligned}$$

On the other hand, this is also equal to

$$\begin{aligned}
\kappa &= \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{abc}{n(ab+bc+ca)} \\
&\geq \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{abc}{n(a+b+c)^2/3} \\
&= \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n} \sum_{\substack{a+b+c=n \\ ab+bc+ca\neq 0}} \binom{n}{a,b,c} \frac{3abc}{n^3} \\
&= \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n} \sum_{\substack{a+b+c=n \\ a,b,c\geq 1}} \binom{n-3}{a-1,b-1,c-1} \frac{3n(n-1)(n-2)}{n^3} \\
&= \frac{5}{8} + \frac{1}{8} \lim_{n\to\infty} 3^{-n}\cdot 3^{n-3}\cdot \frac{3n(n-1)(n-2)}{n^3} = \frac{5}{8} + \frac{1}{8}\cdot\frac{1}{9} = \frac{23}{36}.
\end{aligned}$$

Therefore, $\kappa = 23/36$. $\square$

§5. Four-slope geometric MSNs

As promised before, we give the exact maximum capacity of an RGMSN with at most three slopes. The idea used is similar to that in Theorem 7, but there are more cases here.

Theorem 9. *The maximum capacity of an RGMSN of n sensors with only four possible slopes allowed is equal to*

$$\begin{cases} \dfrac{n^3+2n^2-2n-4}{n(n^2+4n-8)}, & n \text{ is odd}, \\ \dfrac{n^3+2n^2-3n-4}{n(n^2+4n-9)}, & n \text{ is even}. \end{cases}$$

Proof. Suppose there are a lines of slope k_1, b lines of slope k_2, c lines of slope k_3, and d lines of slope k_4, and the rightmost intersection lies on two lines of slopes k_1 and k_2. All lines of slopes k_1 and k_2 form a grid. Lines of slopes k_3 and k_4 can be considered directed paths (or "pseudolines" in the literature [3]) between two neighboring corners of the rightmost corner; the direction indicates whether the slope is positive. They also need to satisfy the following conditions: (i) each path must be valid as in Theorem 7's description; (ii) paths of the same slope cannot intersect, and two paths with slopes k_3 and k_4 intersect exactly once; (iii) the paths' starting points (at infinity) and ending points must be consistent with the slopes.

For the convenience of drawing diagrams, we suppose, without loss of generality, that the grid and all intersections are labeled as in Theorem 7. The paths will be labeled in different ways during the discussion below. We will divide into two cases: $k_3k_4 > 0$ and $k_3k_4 < 0$. In the first case the new paths are all in the same direction and start and end at the same corners of the grid. Thus we will use the words "starting corner," "ending corner," as well as "rightmost corner." In the second case the two neighboring corners of the rightmost corner are symmetric. In all diagrams below, if a line is an arrow, then its slope is positive if the arrow points upward and negative otherwise; if it is not an arrow, then its slope is indicated by the diagram.

Case One

Step 1. *Moving the intersections.*

In this paragraph we discuss the effect of moving an intersection across an edge (Fig. 5) while leaving all other parts of the diagram unchanged. The paths l_1 and l_2 referred below are labeled in the diagram. There are two cases: from (a) to (b), the intersection moves across t_i; from (c) to (d), the intersection moves across s_j ($j > 1$). Only the labels listed in the table below are involved in the change.

(a) to (b)	Change	Reason	(c) to (d)	Change	Reason
$l_1 \cap l_2$	0	It reaches t_i via s_j in (b)	$l_1 \cap l_2$	-1	It cannot reach s_j in (d)
$l_1 \cap t_i$	0	It reaches l_2 via l_1 in (b)	$l_1 \cap s_j$	0	It reaches l_2 via l_1 in (d)
$l_2 \cap t_i$	$+1$	It cannot reach l_1 in (a)	$l_2 \cap s_j$	$+1$	It cannot reach l_1 in (c)

In (c), if $j = 1$, then the label on $l_1 \cap l_2$ is reduced by at least two (s_j and t_{i-1}). Therefore, the intersection of two paths should never exit the grid, but otherwise it can be freely moved in the k_2 direction, or be moved toward the ending corner in the k_1 direction, in order for the capacity to increase.

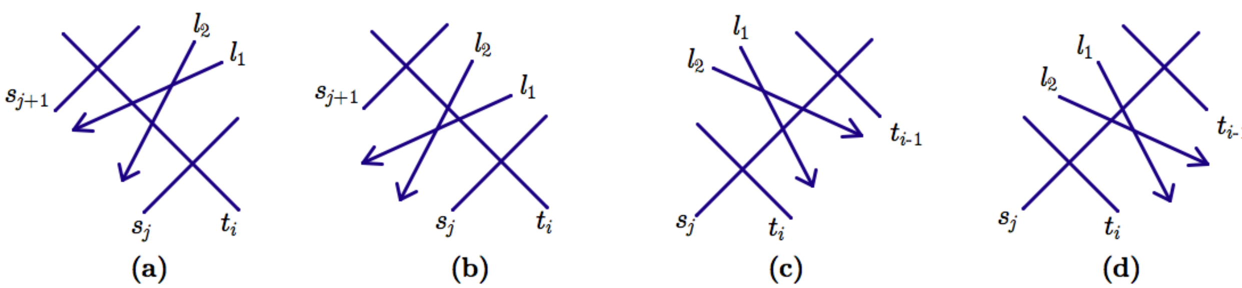

Figure 5. An intersection is moved (from **(a)** to **(b)** and from **(c)** to **(d)**) across an edge in the grid. Note that the intersection moved is of two paths whose directions indicate their slopes, whereas in the grid the slope is drawn.

Step 2. *Moving a non-rightmost path across an intersection.*

In this paragraph we discuss the case shown in Fig. 6. The path l_1 is "non-rightmost" because there is another path l_2 to the right of it. If l_1 and l_2 have the same slope, then the capacity does not change no matter how l_1 moves on the left of l_2 according to the discussion in Theorem 7. When they intersect, there are two cases: from (a) to (b), l_1 crosses a cell when it has already intersected with l_2; from (c) to (d), l_1 crosses a cell when it has not intersected with l_2. Only the labels listed in the table below are involved in the change.

(a) to (b)	Change	Reason	(c) to (d)	Change	Reason
$t_i \cap s_j$	$+1$	It cannot reach l_1 in (a)	$t_i \cap s_j$	0	It reaches l_1 via l_2 in (c)
$l_1 \cap t_i$	-1	It cannot reach s_j in (b)	$l_1 \cap t_i$	-1	It cannot reach s_j in (d)
$l_1 \cap s_j$	0	It reaches t_i via l_1 in (b)	$l_1 \cap s_j$	0	It reaches t_i via l_1 in (d)

Therefore, a non-rightmost path can be moved freely across an intersection if it has already intersected with a path on its right. Before the intersection with that path, it should move farther from the rightmost corner in order to increase the capacity. How a rightmost path should move across intersections in the grid or exit the grid has been discussed in Theorem 7.

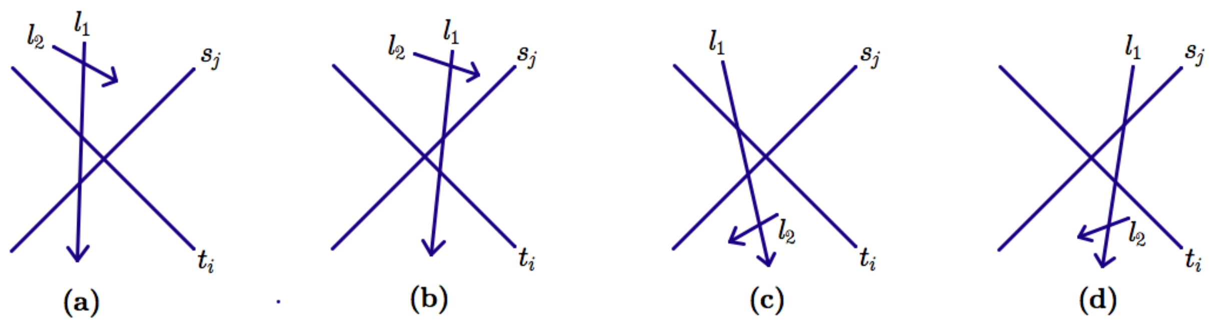

Figure 6. A non-rightmost path l_1 is moved across an intersection in the grid (from **(a)** to **(b)** and from **(c)** to **(d)**). The diagram does not indicate where the path l_2 is; only where $l_1 \cap l_2$ is relative to the intersection $t_i \cap s_j$.

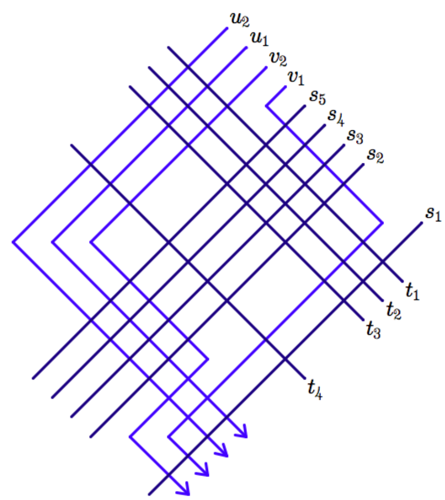

Figure 7. The best arrangement for Case One, where $a = 5$, $b = 4$, and $c = d = 2$.

Now our tools are enough to deal with Case One. All the added paths (with slopes k_3 and k_4 form another grid with its own rightmost, starting, and ending corners. We can first move the ending corner to the last row (between s_1 and s_2) of the k_1-k_2 grid using our transformations above that do not reduce the capacity. This is possible because we can move together the parts of the two paths that are after their intersection, and then move the intersection along the route. Then we repeat the operation for the closest corner (to the ending corner), and so on, until all intersections of the paths are in the cell closest to the ending corner of the k_1-k_2 grid. After that, the paths can be moved in the grid without considering the intersections using our transformations to their best locations shown in Fig. 7.

We can compute that the number of deliveries in Fig. 7 is

$$\sum_{k=2}^{b+c+d+1} k + c \sum_{k=b+c+2}^{b+c+d+1} k + b \sum_{k=b+c+2}^{a+b+c+1} +$$

$$d \sum_{k=b+c+2}^{a+b+c} k + (a-1) \sum_{k=1}^{d} k + c \sum_{k=b+c+d+2}^{a+b+c+d} k + b \sum_{k=a+b+c+2}^{a+b+c+d} k + bc(a+b+c+d).$$

Case Two

In this case, k_3 and k_4 have different signs. All lines (shown as paths) of slopes k_3 and k_4 form another grid. The lines l, l' on which the rightmost intersection of this k_3-k_4 grid lies on divide the plane into four regions. We will call the region on the left of both l and l' the *first region*, that between the two lines the *second region*, and that on the right of both lines the *third region*. We assume that the part of $l \cup l'$ that can reach both lines is in the first region, and the part of $l \cup l'$ that can reach only one of l and l' is in the second region. Fig. 8 shows the regions. Depending on the arrangement, the intersections in the k_1-k_2 grid might belong to any of the three regions (except the rightmost one, which must be in the third region); however, any intersection cannot reach an intersection in another region with smaller index, otherwise one of l and l' would intersect some line with slope k_1 or k_2 more than once.

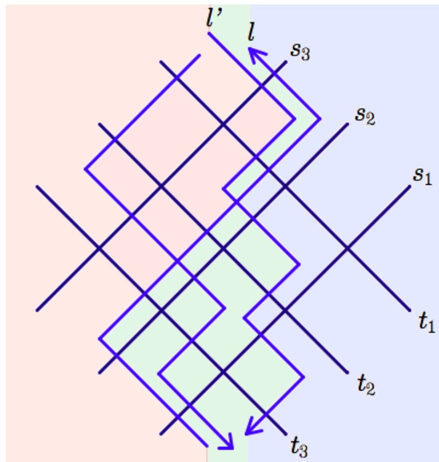

Figure 8. In this arrangement, the first regions are painted red, the second regions are painted green, and the third regions are painted blue. The lines should be included in the regions, but to be visible, they are not painted.

Because $k_3k_4 < 0$, the intersection $l \cap l'$ can reach any line with slope k_1 or k_2, and so does any intersection in the first region. The number of lines that an intersection in the first region can reach is then n minus the number of lines in the k_3-k_4 grid that the intersection cannot reach. This is a fixed number for any intersection in the k_3-k_4 grid in the first region, and at most n for other intersections. It is clear that by moving paths across intersections in the k_1-k_2 grid, we can produce an arrangement with no less capacity where any intersection in the k_1-k_2 grid in the first region can reach all lines with slopes k_3 and k_4.

Step 1. *Moving all lines with slopes k_3 and k_4 together.*

We have the following two operations: (i) move a line in the k_3-k_4 grid across an intersection in the k_1-k_2 grid; (ii) move an intersection in the k_3-k_4 grid across an edge in the k_1-k_2 grid. These operations are shown in Fig. 9 and Fig. 10.

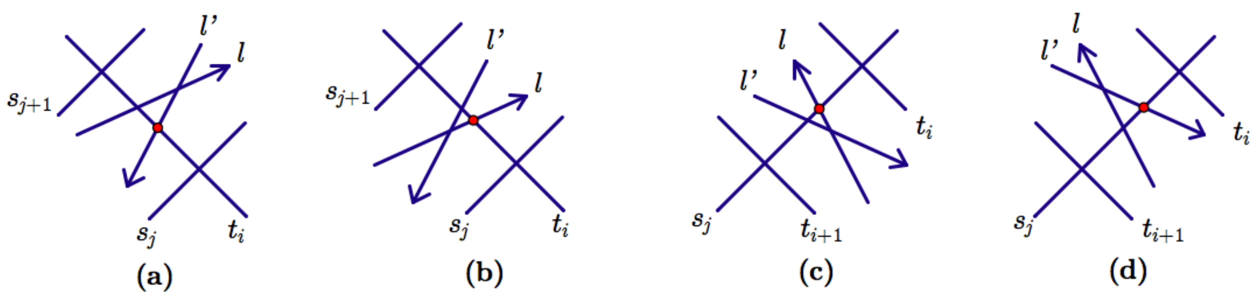

Figure 9. From **(a)** to **(b)**, the rightmost intersection $l \cap l'$ in the k_3-k_4 grid is moved below t_i; from **(c)** to **(d)** it is moved above s_j. Only the number of lines the marked point can reach changes.

Because $l \cap l'$ is the rightmost intersection in the k_3-k_4 grid, there cannot be any intersections in this grid in the second region. So we need only discuss operation (i) on $l \cap l'$ (Fig. 9), which is the only intersection on the boundary between the second region and the third region. (The discussion on the boundary between the first region and the second region is similar, except the same number of new lines in the k_3-k_4 grid can be reached in each arrangement which do not change the result). In the diagrams, only the intersection painted red may change its number of lines it can reach because the other two intersections involved in the change are both on either l or l', and we know that each of them can reach all lines in the k_1-k_2 grid and only the lines l and l' in the k_3-k_4 grid. The lines that the red point can reach are listed in the table below.

Diagram	Lines it can reach	Number	Diagram	Lines it can reach	Number
(a)	$t_1, \ldots, t_b, s_1, \ldots, s_j, l'$	$b+j+1$	(b)	$t_1, \ldots, t_i, s_1, \ldots, s_a, l$	$a+i+1$
(c)	$t_1, \ldots, t_i, s_1, \ldots, s_a, l$	$a+i+1$	(d)	$t_1, \ldots, t_b, s_1, \ldots, s_j, l'$	$b+j+1$

Therefore, the intersection $l \cap l'$ can be moved either above t_i or below s_j. Because moving from Fig. 9(b) to Fig. 9(a) changes the number of deliveries by $(j-i)+(b-a)$, and moving from Fig. 9(d) to Fig. 9(c) changes the number of deliveries by $(i-j)+(a-b)$, one of the changes must be nonnegative (as they sum to zero). And this movement makes the capacity no less.

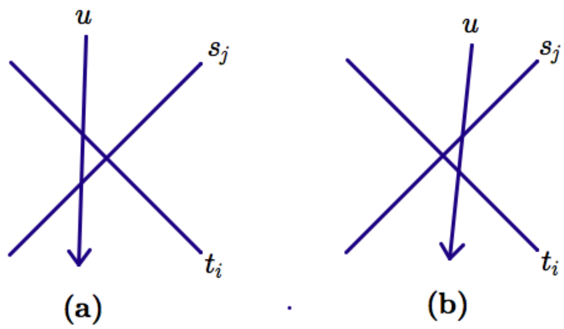

Figure 10. The line u of slope k_3 or k_4 is moved across an intersection $s_j \cap t_i$. We assume that $u \neq l$ and $u \neq l'$.

Operation (ii) is much simpler. On the boundary between the first region and the second region, it always increases the capacity because one intersection in the second region is moved into the first region, and intersections initially in the first region still remain in the first region. At other locations (the second region and the third region), only the lines in the k_3-k_4 grid whose directions are away from $l \cap l'$ can occur, so the discussion is essentially the same as Case One, which we omit. Now, we notice that every operation (i) can be immediately followed by an operation (ii) that puts one more intersection in the k_1-k_2 grid in a region with smaller index. Therefore, after a series of operations, the arrangement is transformed to an arrangement with no less capacity such that there are no intersections in the k_1-k_2 grid between lines in the k_3-k_4 grid. Note that in this step we do not alter where l or l' leaves the k_1-k_2 grid; it will be discussed in Step 3.

Step 2. *Moving the intersections.*

In this step, only operation (i) is needed to move the intersections in the k_3-k_4 grid. From right to left, we label the lines with slope k_3 by $u_1, u_2, \ldots, u_c$ and the lines with slope k_4 by $v_1, v_2, \ldots, v_d$. Suppose we want to move the intersection $u_k \cap v_l$ across either a t_i (from Fig. 11(a) to Fig. 11(b)) or a s_j (from Fig. 11(c) to Fig. 11(d)). Only the red point in the diagrams needs discussion because the other two points involved in the movement can directly reach $u_k \cap v_l$, and thus all lines in the k_1-k_2 grid. And clearly the movement cannot change the number of lines these points can reach in the k_3-k_4 grid. We have several cases.

Case i. $k \neq 1$ and $l \neq 1$. The red point can always reach an intersection in the k_3-k_4 grid and thus all lines in the k_1-k_2 grid. So the capacity does not change.

Case ii. $k = 1$ and $l \neq 1$. Moving from (a) to (b) and from (c) to (d) always results in an arrangement with no less capacity because the red points in (b) and (d) can reach $u_k \cap l_1$, and thus all lines in the k_1-k_2 grid.

Case iii. $k \neq 1$ and $l = 1$. This case is symmetric to Case ii, so similarly moving from (b) to (a) and from (d) to (c) always results in an arrangement with no less capacity.

Case iv. $k = l = 1$. This case has already been discussed in Case One.

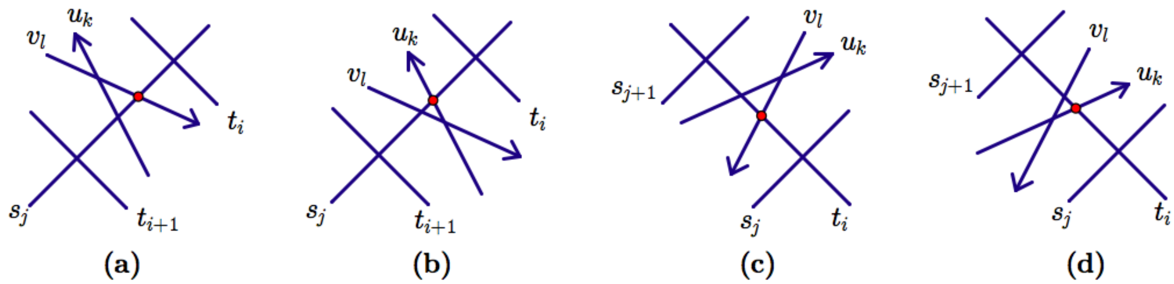

Figure 11. From **(a)** to **(b)**, the intersection $u_k \cap v_l$ in the k_3-k_4 grid is moved below s_j; from **(c)** to **(d)** it is moved below t_i. Only the number of lines the marked point can reach changes. Here the index l is not related to the rightmost line l mentioned before.

It is not hard to see that the operations above that result in no less capacity can move all intersections in the k_3-k_4 grid outside the k_1-k_2 grid, transforming the arrangement to the form in Fig. 12 (where the enclosed region in (a) should be topologically equivalent (for the locations of the lines) to the enclosed region in (b)).

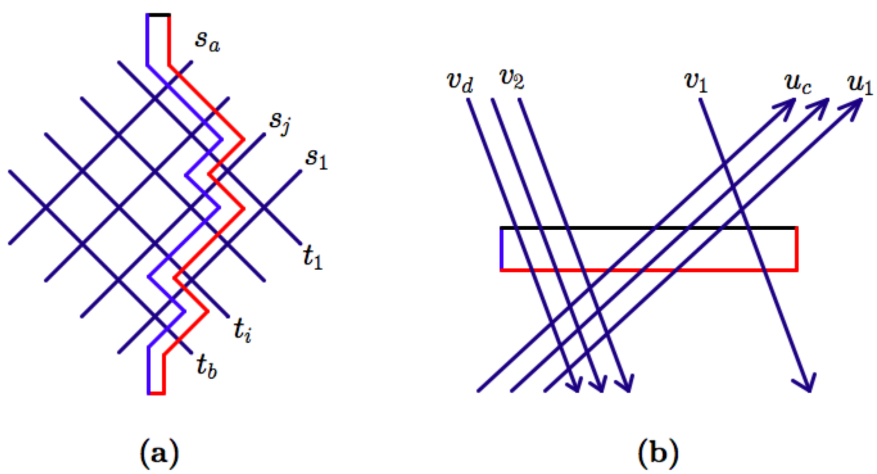

Figure 12. An arrangement can be obtained if the enclosed region in (b) is embedded as the enclosed region in (a). In this diagram, $a = b = d = 4$, $c = i = 3$, and $j = 2$.

Step 3. *Where l and l' leave the k_1-k_2 grid.*

Suppose lines l and l' intersect s_1 between t_i and t_{i+1}, and intersect t_1 between s_j and s_{j+1}. We directly compute the total number of deliveries.

Intersections	Total deliveries
$s_1 \cap t_1, \ldots, s_1 \cap t_i, s_2 \cap t_1, \ldots, s_j \cap t_1$	$\sum_{k=2}^{i+1} k + \sum_{k=3}^{j+1} k$
$v_1 \cap s_k, v_1 \cap t_k$	$\sum_{k=2}^{b-i+1} k + (b+2)i + \sum_{k=b+3}^{b+j+1} k + (b+j+1) + \sum_{k=b+j+2}^{a+b+1} k$
$v_1 \cap u_k$	$\sum_{k=a+b+2}^{a+b+c+1} k$
$u_k \cap s_l, u_k \cap t_l$	$(a+b)\sum_{k=a+b+2}^{a+b+c+1} k$
$u_k \cap v_2, \ldots, u_k \cap v_d$	$(d-1)\sum_{k=a+b+2}^{a+b+c+1} k + c\sum_{k=1}^{d-1} k$
$v_k \cap s_l, v_k \cap t_l$	$(a+b)\sum_{k=a+b+c+2}^{a+b+c+d} k$
Other $s_k \cap t_l$	$(a+b+c+d)(ab-i-j+1)$

In the formula for the total number of deliveries, the terms depending on i and j are

$$2(i^2 - (n-2)i) + (j^2 - (2n-5)j).$$

When this is maximized, we must have $i = j = 1$ due to the range of i and j. However, even when $i = j = 1$, the formula is smaller than that in Case One by $2(n-3) > 0$ when expanded, so Case One gives the maximum capacity, which is the maximum of

$$\frac{a+d+2(b+c)+1}{2n}+\frac{ad+bc(a+d-1)+(b+c)-(b+c)^2}{2n(ad+bc+(a+d)(b+c))}.$$

To maximize this, we observe that for given $a+d$ and $b+c$, ad must be minimized while bc must be maximized, so we have $d=1$ and $|b-c|\le 1$. In this case basic calculus methods (omitted) show that the capacity is decreasing with a increasing, so $a=1$ or $a=2$. If we split into two cases according to the parity of n, in each case we need only compare two numbers, and we found that in each case one number is always larger than the other. The larger number is thus the maximum capacity, which is

$$\begin{cases} \dfrac{n^3+2n^2-2n-4}{n(n^2+4n-8)}, & n \text{ is odd,} \\ \dfrac{n^3+2n^2-3n-4}{n(n^2+4n-9)}, & n \text{ is even.} \end{cases}$$

□

Unfortunately, the casework method above can only give the exact maximum capacity for at most four slopes. For five or more slopes, it fails to work.

Open Question. What is the maximum capacity of an RGMSN of n sensors with only s slopes allowed ($s\ge 5$)? We conjecture it to be

$$1-\frac{(n-1)(n-2)}{n((s-2)(s-3)k^2+2(s-3)kt+t(t-1)+4n-6)}$$

where $n=(s-2)k+t+2$, $0\le t\le s-3$, and k and t are integers.

We have also attempted to compute the expected capacity of an RGMSN of n sensors with at most four slopes by adapting the method in Theorem 8. However, the definite integrals we encountered are out of reach of our computational power. For example, let four lines be $l_1: y=a_1x+b_1$, $l_2: y=a_2x+b_2$, $l_3: y=a_3x+b_3$, and $l_4: y=a_4x+b_4$, and assume that $a_1>a_2>a_3>a_4$. The probability that l_4 can be reached by $l_1\cap l_2$ while b_4 being the largest of m lines, and the probability that both l_4 can be reached by $l_1\cap l_2$ while b_4 being the largest of m lines and l_3 cannot be reached by $l_1\cap l_2$ in one hop are the following integrals, respectively.

Open Question. We have

$$\int_0^1\int_0^1\int_0^1\int_0^{a_1}\int_0^{a_2}\int_{\substack{\max\{0,\min\{1,\\ b_1+(b_2-b_1)\frac{a_1-a_4}{a_1-a_2}\}\}}}^1 mb_4^{m-1}\,\mathrm{d}b_4\,\mathrm{d}a_4\,\mathrm{d}a_2\,\mathrm{d}b_2\,\mathrm{d}b_1\,\mathrm{d}a_1=\frac{3m+1}{24(m+1)},$$

$$\int_0^1\int_0^1\int_0^1\int_0^{a_1}\int_0^{a_2}\int_0^{\substack{\max\{0,\min\{1,\\ b_1+(b_2-b_1)\frac{a_1-a_3}{a_1-a_2}\}\}}}\int_0^{a_3}\int_{\substack{\max\{0,\min\{1,\\ b_1+(b_2-b_1)\frac{a_1-a_4}{a_1-a_2}\}\}}}^1 mb_4^{m-1}\,\mathrm{d}b_4\,\mathrm{d}a_4\,\mathrm{d}b_3\,\mathrm{d}a_3\,\mathrm{d}a_2\,\mathrm{d}b_2\,\mathrm{d}b_1\,\mathrm{d}a_1=\frac{13m^2+18m+2}{1728(m+1)(m+2)}.$$

(Note: we are unable to compute these in Mathematica.)

Besides these integrals, we also need the probability that $l_1\cap l_2$ can reach neither l_3 nor l_4 while b_3 being the largest of k lines and b_4 being the largest of m lines is the following.

Open Question. Compute the integral

$$\int_0^1\int_0^1\int_0^1\int_0^{a_1}\int_0^{a_2}\int_0^{\substack{\max\{0,\min\{1,\\ b_1+(b_2-b_1)\frac{a_1-a_3}{a_1-a_2}\}\}}}\int_0^{a_3}\int_0^{\substack{\max\{0,\min\{1,\\ b_1+(b_2-b_1)\frac{a_1-a_4}{a_1-a_2}\}\}}} kmb_3^{k-1}b_4^{m-1}\,\mathrm{d}b_4\,\mathrm{d}a_4\,\mathrm{d}b_3\,\mathrm{d}a_3\,\mathrm{d}a_2\,\mathrm{d}b_2\,\mathrm{d}b_1\,\mathrm{d}a_1.$$

Below is a list of conjectured integral values, but we are unable to compute any of them in Mathematica.

k	Conjectured integral value
1	$\dfrac{23m^2+90m+70}{1728(m+1)(m+2)}$
2	$\dfrac{223m^3+1526m^2+3069m+1782}{17280(m+1)(m+2)(m+3)}$
3	$\dfrac{363m^4+3915m^3+14530m^2+21700m+10752}{28800(m+1)(m+2)(m+3)(m+4)}$
4	$\dfrac{1247m^5+19623m^4+115657m^3+316033m^2+394560m+176000}{100800(m+1)(m+2)(m+3)(m+4)(m+5)}$

The expected capacity of an RGMSN of n sensors with at most four slopes can be computed if all of these integrals are computed.

§6. Realizability of (R)CMSNs as (R)GMSNs

Although many RCMSNs can be generated from a geometric construction (GMSN), most cannot. Knuth proved in [4] that the number of GMSNs of n sensors is less than $3^{\binom{n+1}{2}}$, but clearly the number of RCMSNs of n sensors is $\binom{n}{2}!$. Therefore, we propose the following question: which RCMSNs are GMSNs? However, a result on pseudoline arrangements by P. Shor implies that there is unlikely an efficient algorithm:

Theorem 10 (GMSN Realizability is NP-Hard)**.** *Given an RCMSN, deciding whether it is generated from a GMSN is NP-Hard.*

(Note: below, a *pseudoline* means a simple curve that is not closed; and we do not allow two pseudolines, or one pseudoline and one line, to intersect more than once.)

Proof. Shor [3] showed that deciding the stretchability of a pseudoline arrangement is NP-Hard. Because all pseudoline arrangements in his paper contain only pseudolines that intersect any additional vertical line at most once, we shall consider only such pseudolines. We also assume that any additional vertical line contains at most one intersection of the pseudolines. Because of this, it is possible to draw two vertical lines such that all intersections of the pseudolines fall between them, and each pseudoline "starts" (at infinity) on the left of the two vertical lines and "ends" on the right of the two vertical lines.

Suppose the arrangement contains n pseudolines, each two of them intersect exactly once between the two vertical lines. We convert this arrangement to a finite sequence of numbers as follows. Let all the intersections ordered by their x-coordinate be $p_1, \ldots, p_{\binom{n}{2}}$. For each $1 \le k \le \binom{n}{2}$, we draw a vertical ray upward from p_k and let a_k be the total number of intersections this ray form with all the pseudolines. Then $a_k \in [n-1]$, and we have a sequence $\{a_k\}_{k=1}^{n(n-1)/2}$.

Let σ be any permutation of $\{1, \ldots, n\}$. We can convert the sequence $\{a_k\}$ to an RCMSN using the following algorithm:

1 $\tau \leftarrow \sigma$
2 **for** k **from** 1 **to** $\binom{n}{2}$
3 $\quad c_k \leftarrow \{\tau(a_k), \tau(a_k+1)\}$
4 $\quad (\tau(a_k), \tau(a_k+1)) \leftarrow (\tau(a_k+1), \tau(a_k))$
5 **return** $\{c_k\}_{k=1}^{n(n-1)/2}$

Here σ is the labels of the pseudolines at the vertical line on the left of all intersections (ordered by their y-coordinates), and τ stores the current labels of the pseudolines at an additional vertical line immediately before and after each intersection, such that the labels are consistent with σ (guaranteed by line 4). Now, the geometric realizabilities of all $n!$ RCMSNs that can be generated from the pseudoline arrangement are equivalent by a relabeling of lines using σ. Therefore, we need to prove that the pseudoline arrangement (if exists) is uniquely determined from an RCMSN up to a relabeling of lines and isomorphism (two arrangements are isomorphic if the graphs, where the vertices are the regions in the arrangement and the edges are adjacency of regions, are isomorphic).

The following algorithm converts an RCMSN $\{c_k = \{x_k, y_k\}\}_{k=1}^{n(n-1)/2}$ (where each $x_k, y_k \in [n]$) into a sequence $\{a_k\}_{k=1}^{n(n-1)/2}$ (where each $a_k \in [n-1]$) as described before, and also reports some clearly nonrealizable RCMSNs:

1 $\tau \leftarrow$ identity $\in S_n$
2 $T \leftarrow$ linked list of $[n]$ with no links
3 **for** k **from** 1 **to** $\binom{n}{2}$
4   **if** T.degree$(\tau(x_k)) = 2$ **or** T.degree$(\tau(y_k)) = 2$
5   **return** *non-realizable*
6   **else if not** T.linked$(\tau(x_k), \tau(y_k))$
7   T.link$(\tau(x_k), \tau(y_k))$
8         $(\tau(x_k), \tau(y_k)) \leftarrow (\tau(y_k), \tau(x_k))$
9 $\sigma(1) \leftarrow T$.head
10 **for** k **from** 2 **to** n
11     $\sigma(k) \leftarrow T$.next$(\sigma(k-1))$
12 $\tau \leftarrow \sigma$
13 **for** k **from** 1 **to** $\binom{n}{2}$
14   **if** $|\tau^{-1}(x_k) - \tau^{-1}(y_k)| \neq 1$
15   **return** *non-realizable*
16   **else**
17         $a_k \leftarrow \min\{\tau^{-1}(x_k), \tau^{-1}(y_k)\}$
18         $(\tau^{-1}(x_k), \tau^{-1}(y_k)) \leftarrow (\tau^{-1}(y_k), \tau^{-1}(x_k))$
19 **return** σ **and** $\{a_k\}_{k=1}^{n(n-1)/2}$

In lines 1–8 we first determine which two numbers in the given RCMSN might represent adjacent pseudolines on the left of both vertical lines we inserted in a GMSN. T is a list of consecutive lines, and τ represents that "the current kth pseudoline from the top is the $\tau(k)$-th pseudoline on the left of both vertical lines we inserted." Only adjacent pseudolines are allowed to intersect because otherwise the pseudolines between them cannot extend across the intersection. Thus, if we fail to build a whole list T, no such GMSN exists. On the other hand, if the list T is built, it represents the relabeling of lines discussed in the previous algorithm and is thus stored in σ (lines 9–11). In lines 12–18, the meaning of τ is the same as that in the previous algorithm, and we reverse the previous algorithm to find $\{a_k\}$. If we fail to reverse it, then at some point there must be two nonadjacent lines that are required to intersect, so clearly no such GMSN exists.

From this algorithm, we see that if an RCMSN is geometrically realizable, it corresponds to at most two pairs $(\sigma, \{a_k\})$ (because T can be read from both sides), and their corresponding pseudoline arrangements are isomorphic as they are mirror images of each other. We shall ignore σ because it does not change the pseudoline arrangement.

Now we need to prove that every sequence $\{a_k\}_{k=1}^{n(n-1)/2}$ where $a_k \in [n-1]$ describes at most one pseudoline arrangement. First, it describes one arrangement naturally: Consider all points (k, j) where $0 \le k \le \binom{n}{2}$ and $1 \le j \le n$ are integers. We draw a ray from each $(0, j)$ horizontally to the left and from each $(\binom{n}{2}, j)$ horizontally to the right. Then, for each k we connect $(k-1, a_k)$ to (k, a_k+1) by a segment, $(k-1, a_k+1)$ to (k, a_k) by a segment, and $(k-1, j)$ to (k, j) by a segment for all $j \notin \{a_k, a_k+1\}$. If this is not a pseudoline arrangement, then two curves must have intersected twice, and the RCMSN corresponds to no pseudoline arrangement. Otherwise, the RCMSN is generated by at least one pseudoline arrangement.

Given a pseudoline arrangement whose sequence is $\{a_k\}$, we need to prove that it is isomorphic to the pseudoline arrangement above. We give a region color 1 if it is adjacent to the region above all pseudolines. Then we give a region color r if it is adjacent to an already colored region with color $r-1$. By definition, the number of regions with color r is equal to the number of r's in $\{a_k\}$. So in our two arrangements the number of regions with each color is equal. Also, it is clear that in both arrangements, the leftmost and rightmost regions with color r must be adjacent to the leftmost and rightmost regions with color $r \pm 1$, respectively. If a region immediately on the left of intersection p_j and a region immediately on the right of intersection p_k have colors differing by one (which is the only case they might be adjacent, by definition), then they are adjacent if and only if $k < j$ because the pseudoline connecting the two intersections is either a shared edge or an edge that separates the two regions. There are no other possible cases of adjacent regions, so the graph of adjacent regions is determined by the sequences $\{a_k\}$. Therefore the two pseudoline arrangements we have are isomorphic.

Hence, if we have an algorithm for the geometric realizability of RCMSNs, then for every pseudoline arrangement we can convert it to an RCMSN in polynomial time and determine if it is generated from a GMSN; and the GMSN as a pseudoline arrangement must be isomorphic to the given one because they correspond to the same RCMSN, as discussed above. □

Although it is hard to determine if an RCMSN is generated from a GMSN, we may consider the similar problem of determining if a CMSN is generated from an RGMSN with limited number of slopes. For example, this problem is easy when the RGMSN has only one or two slopes, because the only possible shape of the network is a grid in this case. It turns out that for three or more slopes we can also solve this problem in polynomial time, as shown in the theorem below. We reduce the problem into linear programming and problems with a smaller number of slopes. However, our algorithm has complexity that is exponential on the number of slopes.

Theorem 11 (RGMSN Realizability is Polynomial)**.** *Given a CMSN and an integer $s \geq 3$, we can decide in polynomial time whether it is generated from an RGMSN with at most s slopes, and construct the GMSN when possible.*

Proof. First, we use the method in Theorem 10 to convert a CMSN into a pseudoline arrangement when possible. If this is not possible, then the CMSN is clearly not realizable. Now we construct a graph where the vertices represent the pseudolines, and two vertices are connected by an edge if the two pseudolines do not intersect. If any component in this graph is not a complete graph, then the pseudoline arrangement is clearly not stretchable since being parallel is transitive. If this graph has more than s connected components, then the CMSN is also not realizable with at most s slopes because we can find four pairwise intersecting pseudolines. If this graph has less than three connected components, then the CMSN is clearly realizable as a grid.

Base Case: $s = 3$.

We assume that the graph has three connected components A, B, and C. Also, we assume that all pseudolines in the arrangement we get are *monotone*, meaning that each pseudoline intersects any vertical line or horizontal line exactly once. We can easily modify Theorem 10's method so that it always gives monotone pseudolines. Now we label the pseudolines in A as $a_1, a_2, \ldots$ in their order from top to bottom when intersecting with a vertical line, and label the pseudolines in B as $b_1, b_2, \ldots$ similarly, but from bottom to top. Note that the direction "top to bottom" is arbitrary; it can be reversed. We assume that $\min\{|A|, |B|, |C|\} \geq 2$ since otherwise the arrangement is clearly stretchable. Also, for a horizontal line above all intersections in the arrangment, we assume without loss of generality that its leftmost intersection with the arrangment is with a line in B, and its rightmost intersection with the arrangement is with a line in A. These are merely relabeling of pseudolines.

Our strategy is to suppose the arrangement is stretchable and try to construct it unless we find a contradiction. For an arrangement stretchable with at most three slopes, we first apply an affine transformation (which does not change the arrangement) to make all lines in A have slope -1 and all lines in B have slope 1. Now we try to make all lines in C vertical. We use the following procedure (Fig. 13):

```
(A', B') ← (A, B)
(l_A, r_A, l_B, r_B) ← (a_1, a_{|A|}, b_1, b_{|B|})
repeat
    (s_A, s_B) ← (|A'|, |B'|)
    for each a ∈ A'
        for each c ∈ C
            if a ∩ c is between l_B and r_B
                B' ← B' ∪ (the line through a ∩ c with slope 1)
    for each b ∈ B'
        for each c ∈ C
            if b ∩ c is between l_A and r_A
                A' ← A' ∪ (the line through b ∩ c with slope −1)
until s_A = |A'| and s_B = |B'|
```

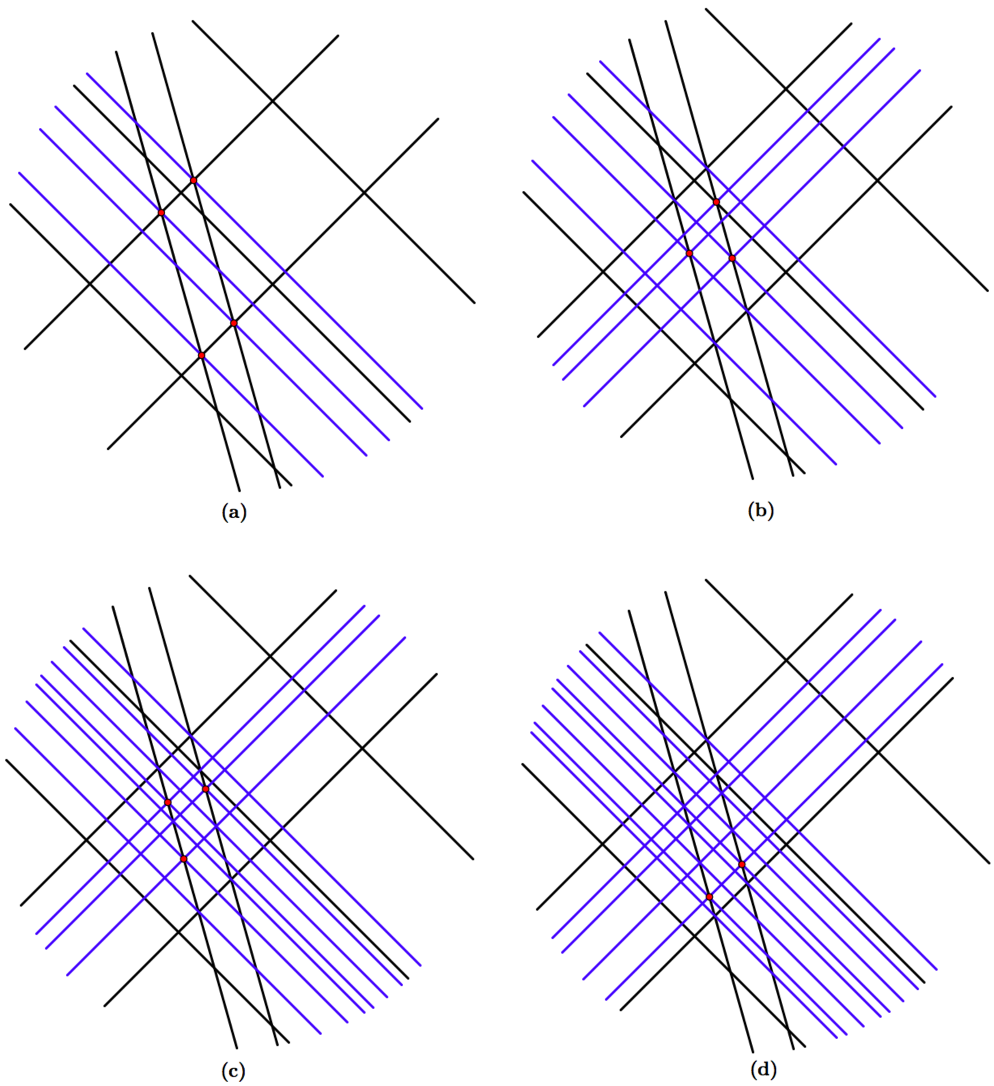

Figure 13. An example of the procedure. All additional lines are blue, and the intersections where additional lines are added are marked. (a) The arrangement at line 9 of the first cycle. (b) The arrangement at line 13 of the first cycle. (c) The arrangement at line 9 of the second cycle. (d) The arrangement at the end.

If this procedure terminates, then we can adjust the distances between the lines so that each cell in the grid formed by A' and B' is a square. Then, because there are no additional lines in satisfying the conditions in line 9 and line 13 in the algorithm, all intersections of lines in C and the grid are at intersections of gridlines; lines satisfying this condition can only be vertical.

Therefore we need to move the lines by some small distances before the procedure above such that the procedure would terminate. This is not difficult: we move all lines by sufficiently small distances such that they have rational intercepts; then we rotate all lines in C by a sufficiently small angle such that the angle between each line in C and each line in A is the arctangent of a rational number. After these operations, there are only finitely many lines that could possibly be generated in the procedure above, so the procedure must terminate.

Now we suppose that all lines in A have slope -1, all lines in B have slope 1, and all lines in C are vertical. Let x_k be the distance between lines a_k and a_{k+1}, and y_k be the distance between lines b_k and b_{k+1}. Consider the intersection $a_i \cap b_l$ and $a_j \cap b_k$, where $i < j$ and $k < l$. (In any other order of i, j, k, and l, the order of the two intersections' x-coordinates is determined.) If there is a pseudoline in C intersecting both ray $(a_i \cap b_l)(a_i \cap b_{l+1})$ and ray $(a_j \cap b_k)(a_j \cap b_{k-1})$ (Fig. 14), then in the stretched arrangement $a_i \cap b_l$ has a smaller x-coordinate than $a_j \cap b_k$, and we have $\sum_{t=i}^{j-1} x_t > \sum_{t=k}^{l-1} y_t$; if there is a pseudoline in C intersecting both ray $(a_i \cap b_l)(a_i \cap b_{l-1})$ and ray $(a_j \cap b_k)(a_j \cap b_{k+1})$, then in the stretched arrangement $a_i \cap b_l$ has a larger x-coordinate than $a_j \cap b_k$, and we have $\sum_{t=i}^{j-1} x_t < \sum_{t=k}^{l-1} y_t$. This is a linear programming problem, and we can determine its consistency in polynomial time. (Note: although we have strict inequalities here, we can scale the arrangement to convert any $p < q$ to the equivalent $p + 1 \leq q$.) If it is consistent, then the arrangement of lines with distances $\{x_k\}$ and $\{y_k\}$ and lines in C in appropriate places gives the GMSN when rotated a small angle. If it is inconsistent, then the pseudoline arrangement is not stretchable because we have found a contradiction.

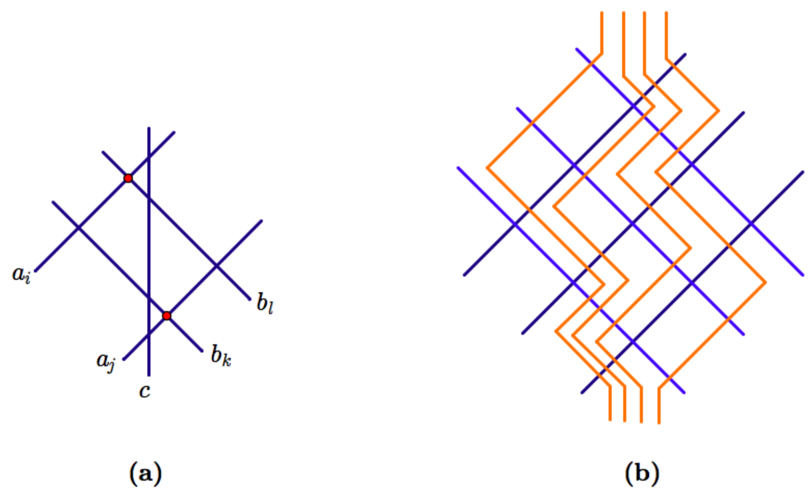

Figure 14. (a) The line $c \in C$ gives a constraint on the distances between lines. (b) An example of a nonstretchable pseudoline arrangement.

Second Base Case: $s = 4$.

From the $s = 3$ case, we see that every RGMSN with four slopes can be transformed into one with the four slopes equal to 0, 1, ∞, and another number $k < 0$ (hereby that will be called a RGMSN$_k$). So our question is that, given a CMSN, whether there exists k such that the network can be realized as a GMSN$_k$. Because we may reflect the whole network over the line $x + y = 0$, without loss of generality, the range of k can be further limited to $0 < -k < 1$.

If we know the number k, we can decide in polynomial time whether the CMSN is realizable as a RGMSN$_k$ using the inequalities that describe the order of intersections of one line and other lines on the line. This is a linear programming problem. For the problem with k unknown, our algorithm is to find $O(n^4)$ different potential values of k (if there are n sensors in the CMSN), and show that if the given CMSN is realizable, it can be realized with one of those specific k values. (The potential values of k depend on the CMSN given.) So we run linear programming at most $O(n^4)$ times to find out the realizability.

Now suppose k is unknown. Then we have a system of strict inequalities that become linear if k is regarded as constant, and all inequalities are linear for k. The set of values of k such that the linear program is feasible must be an open set because all inequalities are strict. Let $-k_0$ be the infimum of that set, then it is a boundary point. Clearly, with $k = k_0$ the linear program is not feasible, but we have a sequence of decreasing feasible values of k converging to k_0. When k follows this sequence, some of the strict inequalities converge to equalities, and with those equalities in place of strict inequalities, the linear program with $k = k_0$ becomes feasible. We need not care about those inequalities that still remain strict because they would still be satisfied when the variables shift by a small amount, restoring the strict inequalities. Regarding the resulting system of equations, we observe that k_0 has to be the unique solution; otherwise it would equal to a quotient of linear expressions of positive variables. The range of such a quotient is an open set, so we can find a smaller $-k_0$, contradicting the fact that it was taken as the infimum of the set of feasible k.

The only way that k_0 is the unique solution of the resulting system of equations is that we have two segments in the network such that the ratio of their lengths is a constant. From this fact, we can find all possible k_0 from the given CMSN in polynomial time. First, we remove all lines with the unknown slope.

Then k_0 must be the maximum slope of some extra line connecting two intersections in the resulting CMSN (which has only lines with known slope). There are $O(n^2)$ intersections, hence at most $O(n^4)$ possible extra lines; for each extra line the determination of k_0 is a linear fractional programming problem, which is known to have a polynomial algorithm.

However, we do not use these k_0 values to test the original system of strict inequalities, because being the infimum they would not satisfy strict inequalities. We get the k values by adding a sufficiently small number to the k_0 values such that the arrangements would not be altered because of the change. We may take that small number to be less than half the square of the smallest nonzero potential k_0 value, which is smaller than half of the difference between any two distinct potential k_0 values.

INDUCTIVE CASE: $s \geq 4$.

By induction, we see that every RGMSN with s slopes can be transformed into an RGMSN with the slopes 0, 1, -1, and one of polynomially many sets of $s-3$ other slopes. Using the same method from the previous case, for each set of those slopes, we can consider them as known slopes and compute in polynomial time polynomially many potential values of the remaining unknown slope. Then we have polynomially many sets of all s slopes to test using linear programming, thus giving a method with complexity polynomial in the network size, but exponential in s. $\square$

§7. Future work

Besides the conjectures listed above, there are many other questions about MSNs that we have not investigated. Although we have obtained the maximum capacities of RCMSNs, GMSNs, and RGMSNs, we have not found the expected capacities of RCMSNs and RGMSNs with more than three slopes. One can also consider a generalized form of GMSN which replaces the lines by polynomial curves, and compute its maximum and expected capacities, and try to determine whether a CMSN is realizable by polynomial curve-GMSNs. Another variant of the GMSNs is to allow three or more lines to intersect at one point, and consider questions similar to those mentioned above.

Acknowledgements

I thank Dr. Jesse Geneson from Department of Mathematics in Iowa State University for mentorship and proposal of this project. I thank Mr. Qiusheng Li from my high school for discussion about the problems. I thank Mr. Yongyi Chen from MIT and Prof. Imre Leader from Trinity College, Cambridge for suggestions about the paper. I thank Dr. Tanya Khovanova, Prof. Pavel Etingof, and Dr. Slava Gerovitch from the PRIMES-USA program in MIT Mathematics Department for the research opportunity.

References

[1] Gu, C., Downes, I., Gnawali, O., Guibas, L. (2018). On the Ability of Mobile Sensor Network to Diffuse Information. *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks* p. 37–47.

[2] Geneson, J. (2018). Sharp extremal bounds on information diffusion capacities of mobile sensor networks. Preprint. `osf.io/n46qv`

[3] Shor, P. W. (1991), "Stretchability of pseudolines is NP-hard", in Gritzmann, P.; Sturmfels, B. (eds.), *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 4, Providence, RI: American Mathematical Society, pp. 531–554.

[4] Knuth, D. E. (1992), *Axioms and Hulls*, Springer–Verlag Berlin.