Recovering lines with fixed linear probes

Extended Abstract

Mark de Berg Jit Bose David Bremner* William Evans
Lata Narayanan

1 Introduction

Suppose the only access we have to an arrangement of n input lines is to "probe" the arrangement with horizontal lines. A probe returns the set of *probe points* which are the intersections of the probe's horizontal line (the *probe line*) with all input lines. We assume that none of the input lines is horizontal, so a probe line intersects every input line. This does not imply that the number of probe points on a probe line is n. It is possible that two or more input lines may intersect the probe line at the same point. Our goal is to reconstruct the set of input lines using a small number of probes.

Aoki, Imai, Imai, and Rappaport [1] observed that if one is allowed to place the probe lines after seeing the results of previous probes, then the number of probes required is at most three. The first two probes serve to define a set of at most n^2 *candidate* lines: the set of lines through one probe point on the first probe line and one probe point on the second probe line. Placing the third probe line so that it does not intersect an intersection of the arrangement of candidates serves to distinguish the input lines from the other candidates.

This paper addresses the problem of *fixed* probes; in our setting the locations of the probes must be chosen before the input is examined. We show that for each natural number n there is a set of $n+1$ probe lines that will serve to determine any arrangement of n input lines and that $n+1$ are sometimes necessary. We further obtain asymptotically tight upper and lower bounds on the maximum number of input lines that are compatible with k probes each of which reports at most n probe points. A line is *compatible* with a set of probes if every intersection of the line with a probe line occurs at a probe point. We give an algorithm to reconstruct an arrangement of n lines that runs in $O(n^2 \log^2 n)$. A randomized version runs in time $O(n^2 \log n)$.

2 Intersection Probes

Let $L \sqcap H = \{\, l \cap h \mid l \in l, h \in H \,\}$. If L is a set of (hidden) non-horizontal lines and H is a

*bremner@math.washington.edu

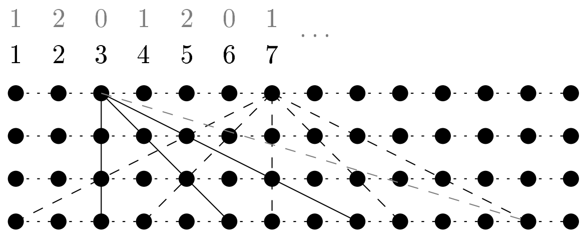

Figure 1: An (n, k) probe set compatible with $n^2/(k-1)$ lines.

set of (known) horizontal probe lines, we want to determine L given just $L \sqcap H$. A natural question is how many probe lines are necessary to determine an arrangement in this way.

Proposition 1 *Let L_1 and L_2 be sets of n non-horizontal lines, and H a set of $n+1$ horizontal lines. If $H \sqcap L_1 = H \sqcap L_2$ then $L_1 = L_2$.*

Proof. If $L_1 \neq L_2$ then chose $\ell \in L_2 \setminus L_1$. There must be exactly $n+1$ intersections $s_1, s_2, \dots s_{n+1}$ between ℓ and H. Since $|L_1| = n$, there must be some s_i that is not intersected by any line in L_1. □

Let S be a set of points in $\mathbb{R}^2$. Let $\mathcal{H}(S)$ be the set of horizontal lines that contain a point in S. We say that a set L of non-horizontal lines is *compatible* with S if $S = \mathcal{H}(S) \sqcap L$. We define the *height* of S as $|\mathcal{H}(S)|$. We define the *width* of S as $\max_{l \in \mathcal{H}(S)} |l \sqcap S|$. If S has width n and height k, we call S an (n, k) *probe set.* Let $f_k(n)$ be the maximum cardinality of a set of non-horizontal lines compatible with an (n, k) probe set. For example, $f_2(n) = n^2$, and Proposition 1 implies that $f_{n+1}(n) = n$. We next prove bounds on $f_k(n)$ for other values of k. Consideration of a rectangular grid of probe points (Figure 1) yields the following.

Proposition 2 $\qquad f_k(n) \geq n^2/(k-1) \qquad 2 \leq k \leq n$

Proof. Let S be a set of probe points arranged in a regular $n \times k$ grid as in Figure 1. Let L be the set of lines that contain the ith probe point in the bottom row (row 1) and the jth probe point in the top row (row k), for all $1 \leq i, j \leq n$ such that $i \equiv j \mod (k-1)$. Each line in L is compatible with S since the line containing the ith probe point in row 1 and the jth probe point in row k contains the $(i + m(j-i)/(k-1))$th probe point in row $m+1$. Let L_i be the subset of L intersecting point i on the top row.

$$|L_i| = \begin{cases} \lfloor n/(k-1) \rfloor + 1 & 1 \leq i \leq n \mod (k-1) \\ \lfloor n/(k-1) \rfloor & \text{otherwise} \end{cases}$$

$$\sum |L_i| = n \left\lfloor \frac{n}{k-1} \right\rfloor + (n \bmod (k-1)) \left\lceil \frac{n}{k-1} \right\rceil$$

$$
\begin{aligned}
&= \frac{n^2}{k-1} + (n \bmod (k-1)) \left(\left\lceil \frac{n}{k-1} \right\rceil - \frac{n}{k-1} \right) \\
&\geq \frac{n^2}{k-1}
\end{aligned}
$$
□

Let $\mathcal{A}(L)$ denote the arrangement induced by a set of lines L. Let $\deg(v)$ be the number of lines that intersect at a vertex v in an arrangement. We will make use in the sequel of the following bound.

Proposition 3 [2] *Let L be a set of N lines and V be any set of M vertices of $\mathcal{A}(L)$.*

$$\sum_{v \in V} \deg(v) \in O(N^{2/3} M^{2/3} + N + M)$$

Our first application of Proposition 3 will be find an asymptotically matching upper bound for the lower bound of Proposition 2.

Theorem 4 $f_k(n) \in O(n^2/k) \qquad 2 \leq k \leq n$

Proof. Consider a set L of N non-horizontal lines and a set H of k horizontal probe lines with at most $n \leq N$ probe points on each; there are $N + k \leq 2N$ lines in total and at most nk points. For n sufficiently large, it follows from Proposition 3 that the sum of the degrees of these points (in $\mathcal{A}(L \cup H)$) is at most

$$
\begin{aligned}
\sum \deg(v) &\leq c_0 \left((2N)^{2/3} (nk)^{2/3} + 2N + nk \right) \\
&\leq c_1 (N^{2/3} (nk)^{2/3} + N + nk)
\end{aligned}
$$

for some constants c_0, c_1. If all lines are to be compatible with all probes then we must have

$$
\begin{aligned}
Nk \leq \sum \deg(v) &\leq c \left(N^{2/3} (nk)^{2/3} + N + nk \right) \\
N(k - c) &\leq c N^{2/3} (nk)^{2/3} + cnk \\
N^{1/3} &\leq \frac{c(nk)^{2/3}}{k - c} + \frac{cnk}{(k - c) N^{2/3}}
\end{aligned}
$$

We know that $N \geq n$, so

$$N \in O(n^2/k + n)$$
□

We now present algorithms to compute L from $L \sqcap H$ for sufficiently large H. In order that the number of input lines need not be known in advance, we assume that the probe points are presented grouped by line. If this is not the case, then an additional $O(nk)$ preprocessing needs to be done for an (n, k) probe set.

Theorem 5 *Let L be a set of n non-horizontal lines and $H = \{ h_1, \ldots, h_k \}$ be a set of horizontal lines such that $n < k$. Given $L \sqcap H$ grouped by probe line, we can compute L in time $O(n^2 \log^2 n)$.*

Proof. Form the set of n^2 candidate lines compatible with h_1 and h_2. Incrementally add probes and "weed out" impossible candidates. The ith weeding out step takes $O(N_i \log n)$ where N_i is the number of candidate lines that survived the first $i-1$ weeding out steps. By Theorem 4, $N_i \in O(n^2/i)$. By Proposition 1, we can stop when the number of probes processed is greater than the number of remaining candidates. Thus the time required is at most $c \sum N_i \log n = cn^2 \log n \sum 1/i = O(n^2 \log^2 n)$. □

Theorem 6 *Let L be a set of n non-horizontal lines and $H = \{ h_1, \ldots, h_k \}$ be a set of horizontal lines such that $n < k$. Given $H \sqcap L$ grouped by probe line, we can compute L in expected time $O(n^2 \log n)$.*

Proof. Initially form the set L_0 of $N_0 \leq n^2$ candidate lines obtained by joining every probe point on h_1 with every probe point on h_2. Let $H_0 = H \setminus \{ h_1, h_2 \}$.

At each step $i \geq 1$, we have a set L_{i-1} of N_{i-1} candidate lines and a set H_{i-1} of at least $n - i$ unprocessed probe lines. Choose a probe line h from the set H_{i-1} at random. Let $H_i = H_{i-1} \setminus \{h\}$. Let L_i be the candidate lines from L_{i-1} that are compatible with h. Forming L_i, that is testing each of the N_{i-1} candidate lines in L_{i-1} to see if it is eliminated, requires $O(N_{i-1} \log n)$ time. By Proposition 1, we know we can stop when $N_i < i$.

We can divide the running of this incremental algorithm into two stages according to whether or not the following is satisfied:

$$(1) \qquad i \leq \frac{n}{2} \qquad \text{and} \qquad N_i > c_1 n$$

for some constant $c_1$to be chosen later. We will bound the time taken in stage 1, when (1) holds, below. In stage 2, when (1) is not satisfied, we know from Theorem 4 that $|L_i| \in O(n)$. Both stages together take time at most

$$c_2 \sum_{i=0}^{n/2} N_i \log n + O(n^2 \log n).$$

where the N_i are random variables. We will bound the expected running time of the algorithm by obtaining a bound on $\mathbf{E} N_i$.

Lemma 1 *For $0 \leq i \leq n/2$, $\mathbf{E} N_i \leq c_3 n^{1+(2/3)^i}$*

Proof. The proof is by induction on i. The lemma holds for $i = 0$. Consider the change in the number of candidate lines at step $i \geq 1$. Let H^*_{i-1} be the set of $n - i$ probe lines chosen after step $i - 1$. Let V be the set probe points on the probe lines in H^*_{i-1} ($|V| \leq n(n-i)$). We know (Proposition 3) that

$$(2) \qquad \sum_{v \in V} \deg(v) \leq c((N_{i-1} + n - i)^{2/3}(n(n-i))^{2/3} + (N_{i-1} + n - i) + n(n-i))$$

Let D be the right-hand side of (2). Let D_i denote the sum of the degrees of the probe points on the ith probe line h_i processed. Since all orderings of the $n-i$ probe lines to be chosen randomly are equally likely, $\mathbf{E}D_i = \mathbf{E}D_j$, for all $2 \le i \le j \le n$. It follows that $\mathbf{E}D_i \le D/(n-i)$. Thus the expected number N_i of candidate lines that survive the test at step i is at most $D/(n-i)$, or since $N_{i-1} > c_1 n$,

$$\mathbf{E}N_i \le c_4 \mathbf{E}\left(N_{i-1}^{2/3}\frac{n^{2/3}}{(n-i)^{1/3}} + \frac{N_{i-1}}{n-i} + n\right)$$

where $c_4 = c(1+1/c_1)^{2/3}$. Since $f(x) = x^{2/3}$ is a concave function, $\mathbf{E}[N_{i-1}^{2/3}] \le (\mathbf{E}N_{i-1})^{2/3}$ and by induction,

$$\mathbf{E}N_i \le c_4\left(c_3^{2/3}n^{(2/3)^i}\frac{n^{4/3}}{(n-i)^{1/3}} + \frac{n^{1+(2/3)^{i-1}}}{n-i} + n\right)$$

Since $n-i \ge n/2$,

$$\begin{aligned}\mathbf{E}N_i &\le c_4\left(c_3^{2/3}2^{1/3}n^{1+(2/3)^i} + 2n^{(2/3)^{i-1}} + n\right)\\ &\le c_4(c_3^{2/3}2^{1/3}+3)n^{1+(2/3)^i}\end{aligned}$$

which is at most $c_3 n^{1+(2/3)^i}$ for $c_3 > \max\{\sqrt{27/2}, 16c_4^3\}$. □

The expected running time $T(n)$ of the algorithm is then bounded as follows

$$T(n) \le c_2 \sum_{i=0}^{n/2} c_3 n^{1+(2/3)^i}\log n + O(n^2\log n)$$

We claim that

$$T(n) \in O(n^2 \log n) \tag{3}$$

To see that (3) holds, note that for $x \ge 8$, $x/2 \ge x^{2/3}$. Thus

$$\sum_{i=0}^{n/2} n^{(2/3)^i} \le 4n + \sum_{i=0}^{n/2} 2^{-i}n \le 6n$$

which suffices to prove the claim, and the theorem. □

Acknowledgments

This research came out of a workshop on Computational Tomography organized by Godfried Toussaint. The authors are grateful to Godfried Toussaint for organizing the workshop, and to the Bellairs Research Institute of McGill University for hosting it.

References

[1] Yasukazu Aoki, Hiroshi Imai, Keiko Imai, and David Rappaport. Probing a set of hyperplanes by lines and related problems. In *Proc. 3rd Workshop Algorithms Data Struct.*, volume 709 of *Lecture Notes Comput. Sci.*, pages 72–82. Springer-Verlag, 1993.

[2] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl. Combinatorial complexity bounds for arrangements of curves and spheres. *Discrete Comput. Geom.*, 5:99–160, 1990.