Adaptive Communications and Signal Processing laboratory
Cornell University, Ithaca, NY 14853

# An Almost Surely Complete Subset of Planar Disks

**Ting He and Lang Tong**

Technical Report No. ACSP-TR-04-05-01

April 2005

## I. INTRODUCTION

Let $(X, \mathcal{F})$ be a given measurable space, and $\mathcal{A} \subseteq \mathcal{F}$ be a set of (possibly infinitely many) measurable sets. Assume $X \subseteq \mathbb{R}^2$,*i.e.*, the sample space is a planar region. In accordance with [1], define two probability spaces $(X, \mathcal{F}, P_1),\ (X, \mathcal{F}, P_2)$, and draw two collections of samples $S_1,\ S_2$ i.i.d. from $P_1,\ P_2$ respectively. Let $S = S_1 \cup S_2$. Our goal is to find a finite subset of $\mathcal{A}$ such that this subset gives the same set of intersections with $S$ as $\mathcal{A}$. This property is what we will define as “completeness”.

In this report, we let $\mathcal{A}$ be the set of planar disks. We propose a way to choose a finite subset of $\mathcal{A}$, and prove that it is almost surely complete under certain conditions.

## II. COMPLETE SUBSETS

We call a subset “complete” if it satisfies the following definition:

**Completeness** Given $\mathcal{A}$ being a collection of measurable subsets of space $X$, and $S \subset X$ be a set of points in $X$. Let $\mathcal{H}(S) \subset \mathcal{A}$ be a finite sub-collection of measurable sets which is a function of $S$. We call the collection $\mathcal{H}(S)$ *complete for* $S$ *with respect to* $\mathcal{A}$ if $\forall A \in \mathcal{A}$, there exists a $B \in \mathcal{H}(S)$ such that $S \cap A = S \cap B$.

Now let $\mathcal{A}$ be the collection of two dimensional disks. Hereafter all sets are closed unless otherwise stated.

For the set of sample points $S \subseteq X$, consider the finite sub-collection of $\mathcal{A}$ defined by

$$\mathcal{H}_{\mathrm{D}}(S) \triangleq \bigcup_{(s_i, s_j, s_k) \in \mathcal{T}} \mathcal{H}_{\mathrm{D}}(s_i, s_j, s_k) \tag{1}$$

where

$$\mathcal{T} \triangleq \{s_i, s_j, s_k \in S^3 : s_i, s_j, s_k \text{ are not collinear}\},$$

and

$$\mathcal{H}_{\mathrm{D}}(s_i, s_j, s_k) \triangleq \{D(s_i, s_j, s_k), D(s_i, s_j, s_k) \setminus \{s_i\},$$

$$D(s_i, s_j, s_k) \setminus \{s_j\}, \ldots, D(s_i, s_j, s_k) \setminus \{s_i, s_j, s_k\}\}$$

where $D(s_i, s_j, s_k)$ is the disk with $s_i$, $s_j$, and $s_k$ on its boundary, *i.e.*, $\mathcal{H}_{\mathrm{D}}(s_i, s_j, s_k)$ is $D(s_i, s_j, s_k)$ and all the $7$ variations for excluding some of the $3$ boundary points. See Figure 1.

We claim the following is true:

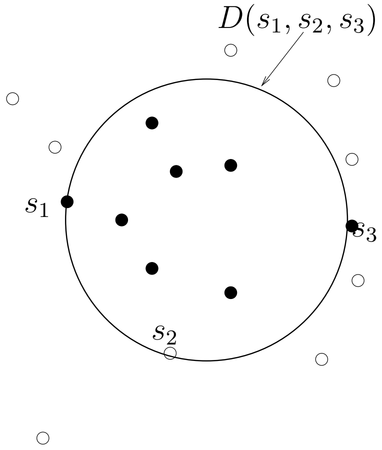

Fig. 1. Members of $\mathcal{H}_\text{D}$: $D(s_1, s_2, s_3) \in \mathcal{H}_\text{D}$; $D(s_1, s_2, s_3) \setminus \{s_2\} \in \mathcal{H}_\text{D}$

*Claim 0.1:* Let $S = (s_1, s_2, \dots, s_n)$, $n \geq 3$, where $s_i = (x_i, y_i)$. If $S$ is such that no three points in $S$ are collinear, and no four points in $S$ are concyclic, then for any disk $A$, there is a disk $A'$ s.t. $A'$ has three points on the boundary, and

$$A' \cap S = (A \cap S) \cup \{\text{at most three more boundary points}\}.$$

Proof of the claim: See Appendix.

Using the claim, we can derive the main result of this report:

*Proposition 2.1:* Let $\mathcal{A}$ be the collection of two dimensional disks. For $S_1$ and $S_2$ drawn from $P_1$ and $P_2$, if $P_1$ and $P_2$ are such that any set with Lebesgue measure $0$ has probability $0$[1], then the finite collection $\mathcal{H}_\text{D}(S_1 \cup S_2)$ in (1) is complete with respect to $\mathcal{A}$ a.s.(almost surely).

*Proof:*

Let $S = S_1 \cup S_2$. Since $P_1$,$P_2$ have $0$ probability on $0$-measure set and both a line and the boundary of a disk have Lebesgue measure $0$, we have that with probability $1$ (*i.e.,* almost surely) no three points in $S$ are collinear, and no four points in $S$ are concyclic.

Apply Claim 0.1:

---

[1] This is true if $P_1$,$P_2$ are absolutely continuous, *i.e.*, having pdf, because any measurable function has integration $0$ on a $0$-measure set.

Given sample $S = S_1 \cup S_2$ and a disk $A$, a.s. one can find a disk $A'$ such that $A'$ has three points on the boundary and the same intersection with $S$ as $A$ except on the boundary points. Now define $A'' \subseteq A'$ such that for any $s \in S$ on the boundary of $A'$, $s \in A'' \iff s \in A$. Then $A'' \cap S = A \cap S$. Since $A'$ has 3 points of $S$ on its boundary, $A'' \in \mathcal{H}_{\text{D}}(S)$.

■

## APPENDIX

### A. *Proof of Claim 0.1*

*Proof:*

Let $S = (s_1, s_2, \ldots, s_n),\ n \geq 3$, where $s_i = (x_i, y_i)$.

Define $D(a, b, r)$ to be the disk centered at $(a, b)$ with radius $r$ ($r \geq 0$), *i.e.,*

$$D(a, b, r) = \{(x, y) :\ (x - a)^2 + (y - b)^2 \leq r^2\}.$$

Given disk $D(a_0, b_0, r_0)$, let

$$D(a_0, b_0, r_0) \cap S = \{s_1, \ldots, s_k\}.$$

To avoid the trivial case, suppose $k \geq 2$. Define

$$\begin{aligned} \mathcal{P} \;=\; & \{(a, b, r) : \\ & (x_i - a)^2 + (y_i - b)^2 \leq r^2,\ \ i = 1, \ldots, k \qquad (2) \\ & (x_i - a)^2 + (y_i - b)^2 \geq r^2,\ \ i = k + 1, \ldots, n\big\}. \qquad (3) \end{aligned}$$

Any disk $D(a, b, r),\ (a, b, r) \in \mathcal{P}$ shatters the same set of points as $D(a_0, b_0, r_0)$, except that it may contain additional points on the boundary. To prove the claim, it is equivalent to show that there exists $(a, b, r) \in \mathcal{P}$ s.t. three of the inequalities (2,3) attain equality (*i.e.,* there are three tight constraints).

We know $(a_0, b_0, r_0) \in \mathcal{P}$. Take

$$r_1 = \max_{i=1,\ldots,k} \sqrt{(x_i - a_0)^2 + (y_i - b_0)^2}.$$

Then $(a_0, b_0, r_1) \in \mathcal{P}$ and has at least one tight constraint among the first $k$ constraints. Suppose it is for constraint $i = 1$.

Assume $a_0 \geq x_1$. Define function

$$a(r,b) = x_1 + \sqrt{r^2 - (y_1 - b)^2}.$$

Consider an optimization problem

$$\begin{aligned} &\min r \\ \text{s.t.} \quad &(a(r,b), b, r) \in \mathcal{P} \\ &r \geq 0 \end{aligned}$$

The optimization is feasible (since $(b_0, r_1)$ is a feasible solution) and bounded, so the optimal solution exisis. Let it be $(\tilde{b}, \tilde{r})$, $\tilde{r} > 0$ (because $k \geq 2$).

Note that constraint $i = 1$ always holds, since

$$(x_1 - a(r,b))^2 + (y_1 - b)^2 = r^2.$$

Hence the optimal solution must attain equality on at least one of $i = 2, \ldots, n$. Otherwise, by continuity of the function $(x - a(r,b))^2 + (y - b)^2 - r^2$ with respect to $r$, we know $\exists \epsilon > 0$ s.t.

$$(a(\tilde{r} - \epsilon, \tilde{b}), \tilde{b}, \tilde{r} - \epsilon) \in \mathcal{P}, \text{ and } \tilde{r} - \epsilon \geq 0,$$

which contradicts with that $\tilde{r}$ is optimal.

If $(\tilde{b}, \tilde{r})$ attains equality on more than one of $i = 2, \ldots, n$, then we are done. $(a(\tilde{r}, \tilde{b}), \tilde{b}, \tilde{r}) \in \mathcal{P}$ has three tight constraints.

Otherwise, suppose $(\tilde{b}, \tilde{r})$ attains equality only on $j$, $j \in \{2, \ldots, n\}$.

Now we write out the parametric equations for $a$, $b$, $r$ for disks with $s_1$, $s_j$ on the boundary:

$$r(t) = \sqrt{(\frac{s_1 s_j}{2})^2 + t^2} \tag{4}$$

$$a(t) = -\frac{(y_j - y_1)}{s_1 s_j} t + \frac{x_1 + x_j}{2} \tag{5}$$

$$b(t) = \frac{(x_j - x_1)}{s_1 s_j} t + \frac{y_1 + y_j}{2} \tag{6}$$

where $s_1 s_j = \sqrt{(x_1 - x_j)^2 + (y_1 - y_j)^2}$. See Fig. 2.

$\{D(a(t), b(t), r(t))\}_{t \in \mathbb{R}}$ is the set of disks with $s_1$, $s_j$ on the boundary.

Define

$$H_+ \triangleq \{(x,y) : \ (y - y_1)(x_j - x_1) \geq (x - x_1)(y_j - y_1)\},$$

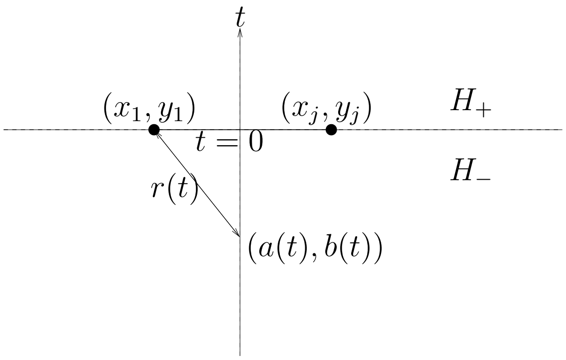

Fig. 2. Explanation of $a(t)$, $b(t)$, $r(t)$

and

$$H_- \triangleq \{(x,y) :\ (y-y_1)(x_j-x_1) \le (x-x_1)(y_j-y_1)\}.$$

Note that as $t \uparrow +\infty$,

$$D(a(t), b(t), r(t)) \to H_+,$$

and as $t \downarrow -\infty$,

$$D(a(t), b(t), r(t)) \to H_-.$$

Define function

$$f(x,y,t) = (x-a(t))^2 + (y-b(t))^2 - r(t)^2.$$

If $(x,y) \in \text{int}(H_+)$ (interior of $H_+$), then

$$\liminf_{t \to -\infty} f(x,y,t) > 0,$$

and

$$\limsup_{t \to +\infty} f(x,y,t) < 0.$$

Let $\tilde{t}$ be s.t. $\tilde{r} = r(\tilde{t})$. Then $(a(\tilde{t}), b(\tilde{t}), r(\tilde{t})) \in \mathcal{P}$. Since $n \ge 3$, there is at least one more point $s_i,\ i \in \{2, \ldots, n\},\ i \ne j$. Now we try to find the third tight constraint by varying $t$, starting from $t = \tilde{t}$.

If $s_i \in \text{int}(H_+)$ and $i \le k$, we can find $m$, the first constraint violated (*i.e.*, $f(x_m, y_m, t)$ changes sign) as $t \downarrow -\infty$ from $t = \tilde{t}$. By the continuity of $f(x,y,t)$ with respect to $t$, we have $\exists t_0 \le \tilde{t}$ s.t.

$$f(x_m, y_m, t_0) = 0,$$

and

$$\text{sign}(f(x_l, y_l, t_0)) = \text{sign}(f(x_l, y_l, \tilde{t})),$$

for $l = 2, \dots, n,\ l \neq j,\ l \neq m$. Here sign$(\cdot)$ is the sign function. In other words,

$$(a(t_0), b(t_0), r(t_0)) \in \mathcal{P},$$

and has three tight constraints, *i.e.*, $1,\ j,\ m$.

Such an $m$ exists because on one hand, $(a(\tilde{t}), b(\tilde{t}), r(\tilde{t})) \in \mathcal{P}$ and $i \leq k$ imply

$$f(x_i, y_i, \tilde{t}) < 0;$$

on the other hand, $s_i \in$ int$(H_+)$ implies

$$\liminf_{t \to -\infty} f(x_i, y_i, t) > 0.$$

At least constraint $i$ will be violated for $t$ small enough.

If $s_i \in$ int$(H_+)$ and $i > k$, we can find $t_0$ in the same way as $t \uparrow +\infty$ (from $\tilde{t}$). Similar arguments hold for $s_i \in$ int$(H_-)$. (These cover all the cases since no three points of $S$ are collinear.)

Therefore, we can always find $(a, b, r) \in \mathcal{P}$ with three tight constraints. This completes the proof.

■

## REFERENCES

[1] T. He, S. Ben-David, and L. Tong, “Nonparametric Change Detection and Estimation in Large Scale Sensor Networks.” submitted to IEEE Trans. on Signal Processing, December 2004.