CENTRAL SIMILARITY PROXIMITY MAPS IN DELAUNAY TESSELLATIONS

September 22, 2003

Elvan Ceyhan*, Johns Hopkins University, Carey E. Priebe, Johns Hopkins University
*Department of Mathematical Sciences, Johns Hopkins University,
Baltimore, MD 21218 (ceyhan@mts.jhu.edu)

Abstract:

Proximity maps and (di)graphs are used in many areas for various purposes. We introduce a new type of proximity map (and the associated (di)graph) and investigate its properties in Delaunay tessellations, and we propose an illustrative application to support estimation.

1 Introduction

Proximity maps (and the associated graphs) are used in disciplines where shape and structure are important. Examples include computer vision (dot patterns), image analysis, pattern recognition (prototype selection), geography and cartography, visual perception, biology, etc. In this paper, we define a new type of proximity map and investigate its properties. We illustrate our proximity map with an application to support estimation.

Let $(\Omega, \mathcal{M})$ be a measurable space. The *proximity map* $N(\cdot)$ is given by $N : \Omega \to \wp(\Omega)$, where $\wp(\cdot)$ is the power set functional, and the *proximity region* of $x \in \Omega$, denoted $N(x)$ is the image of $x \in \Omega$ under $N(\cdot)$. The points in $N(x)$ are thought of as being "closer" to $x \in \Omega$ than are the points in $\Omega \setminus N(x)$. Proximity maps are the building blocks of the *proximity graphs* introduced by Toussaint in [10]; an extensive survey is available in [11].

If $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ is a set of Ω-valued random variables then $N(X_i)$ are random sets, and the associated digraph is a *data random proximity digraph*. If X_i are independent identically distributed then so are the random sets $N(X_i)$. The *proximity digraph* D has the vertex set $\mathcal{V} = \mathcal{X}_n$ and the arc set $\mathcal{A}$ is defined by $X_iX_j \in \mathcal{A}$ iff $X_j \in N(X_i)$.

The *Delaunay tessellation* in $\mathbb{R}^d$ constructed by using a data set, B, is the dual of the *Voronoi diagram* constructed by using the same data set. This tessellation partitions the convex hull $\mathcal{C}_H(B)$ of B. In $\mathbb{R}^2$, the tessellation is a *triangulation* that yields triangles T_j, $j = 1, \ldots, J$ (see, e.g. [3]) provided that no more than three points are cocircular. See Figure 1 (top) for an example with $n = 200$ and $|\mathcal{Y}| = 20$, where X and Y are $\overset{iid}{\sim} \mathcal{U}((0,1) \times (0,1))$ and the Delaunay triangulation is based on $\mathcal{Y}$.

The data random proximity digraphs are introduced by Priebe et al. [5] in $\mathbb{R}^1$ where the Delaunay tessellation yields intervals. Suppose, $\mathcal{Y} = \{y_1, y_2\}$ with $y_1 < y_2$ and $X_i \overset{iid}{\sim} \mathcal{U}(y_1, y_2)$, then X_i can be transformed into iid $\mathcal{U}(0,1)$ random variables, that is, we can assume the support of $\mathcal{X}_n$ to be $\mathcal{S} = (0,1)$. The proximity map considered in Priebe et al. [5] is $N(x) := B(x, r(x)) = (x - r(x), x + r(x))$, where $r(x) = \min(x, 1-x)$. We note some of the appealing properties of this proximity region:

P1 $x \in N(x)$ for all $x \in \mathcal{S}$.

P2 x is at the *center* of $N(x)$.

P3 $N(x)$ and $\mathcal{S}$ are of the *same type*; they are both intervals.

P4 $N(x)$ mimics the shape of $\mathcal{S}$; i.e., it is *similar* to $\mathcal{S}$.

P5 $N(X)$ is a proper subset of the $\mathcal{S}$ a.s.

Note that **P4** and **P3** are equivalent when $d = 1$, but for $d > 1$ only **P4** $\Rightarrow$ **P3**. Two natural extensions of $N(x) = B(x, r(x))$ in $\mathbb{R}^1$ to $\mathbb{R}^d$ with

This work was partially supported by Office of Naval Research Grant N00014-01-1-0011 and Defense Advanced Research Projects Agency Grant F49620-01-1-0395.

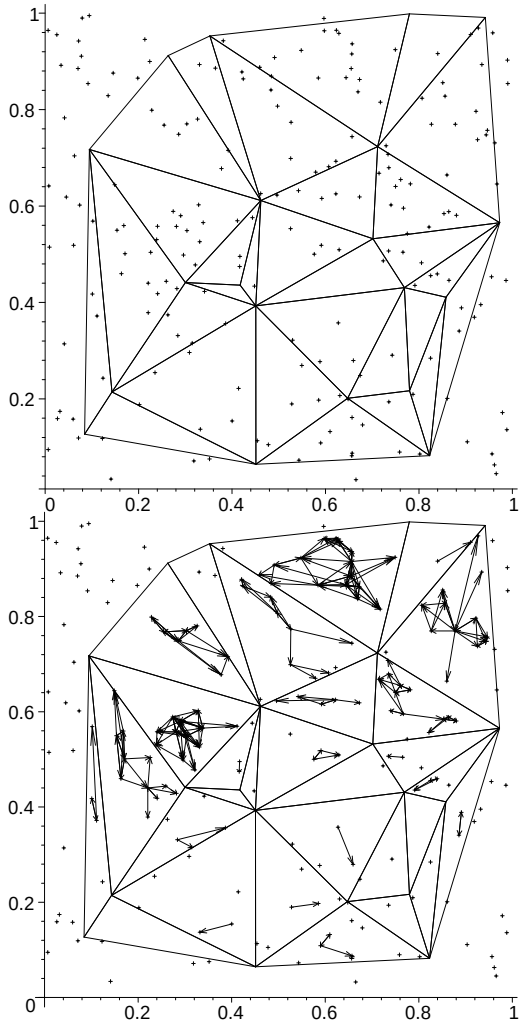

Figure 1: A realization of 200 X points and 20 Y points and the corresponding Delaunay triangulation based on $\mathcal{Y}$ (top); and the associated digraph (bottom).

$d > 1$ are the spherical proximity region $N_S(x) = B(x, r(x))$ where $r(x) = \min_{y \in \mathcal{Y}} d(x, y)$ and the arcslice proximity region $N_{AS}(x) = B(r, r(x)) \cap \mathcal{T}(x)$ where $\mathcal{T}(x)$ is the Delaunay tessellation cell that contains x. However, $N_S(\cdot)$ satisfies only **P1** and **P2**, and $N_{AS}(\cdot)$ satisfies only **P1** and **P5**. Below, we introduce a new type of proximity map that satisfies all of the properties **P1 – P5**.

2 Central Similarity Proximity Maps

Let $\Omega = \mathbb{R}^2$ and let $\mathcal{Y} = \{y_1, y_2, y_3\} \subset \mathbb{R}^2$ be three non-collinear points. Let $T(\mathcal{Y})$ be the triangle formed by these three points and e_j be the edge opposite vertex y_j for $j = 1, 2, 3$. For $x \in T(\mathcal{Y})$, the *central similarity proximity region* $N_{CS}(x)$ is defined to be the triangle similar to $T(\mathcal{Y})$ and adjacent to an edge e of $T(\mathcal{Y})$ with its centroid at x. To choose the adjacent edge in $T(\mathcal{Y})$, we use *centroidal edge regions* that are obtained by joining the centroid (center of mass) C_M and vertices with straight lines. So the similarity ratio is $d(x, e)/d(C_M, e)$. See Figure 2 (top) for $N_{CS}(x)$ with $e = e_3$. Furthermore, a quick investigation shows that $N_{CS}(\cdot)$ satisfies all the properties **P1**-**P5**. The digraph associated with the entire $\mathcal{X}_n$ and $\mathcal{Y}$ of Figure 1 (top) using N_{CS} is depicted in Figure 1 (bottom). If the adjacent edge were picked as $\text{argmin}_{e \in \{e_1, e_2, e_3\}} d(x, e)$, then we would have been implicitly using *incenter edge regions*.

A *prototype* or *representative set* B_P of a finite data set, B, is a set that depends on B in such a way that B_P and B yield (nearly) the same result after a particular procedure, say classification or support estimation. If $B_P \subseteq B$, then $|B_P| \leq |B|$. When strict inequality holds, we say that the associated procedure is *edited* or *condensed*. However, $B_P \subseteq B$ is not required, see for example *learning vector quantization* in [2]. A reasonable choice for the prototype set is a *minimum dominating set*, which is often more "central" than arbitrary sets of the same size in $T(\mathcal{Y})$.

In a digraph D a vertex $v \in \mathcal{V}$ *dominates* itself and all vertices of the form $\{u : vu \in \mathcal{A}\}$. A *dominating set* S_D for the digraph D is a subset of $\mathcal{V}$ such that each vertex $v \in \mathcal{V}$ is dominated by a vertex in S_D. A *minimum dominating set* S_D^* is a dominating set of minimum cardinality, and the *domination number* $\gamma(D)$ is defined as $\gamma(D) := |S_D^*|$ [1]. If a minimum dominating set is of size 1, we call it a *dominating point*.

For $N_{CS}(\cdot)$ and $\mathcal{X}_n$, the associated digraph and the domination number are denoted as D_{CS}, and γ_n, respectively. In particular, if $\gamma_n = 1$, the prototype set of size one, namely one of the dominating points in $T(\mathcal{Y})$. Next we define a subset of $T(\mathcal{Y})$ associated with $\gamma_n = 1$ case.

The Γ_1-*region* associated with a proximity map $N(\cdot)$ for a set $Q \in \Omega$ and is defined to be $\Gamma_1(Q, N) := \{x \in \Omega : Q \subseteq N(x)\}$; i.e., the Γ_1-region is the subset of Ω whose intersection with Q yields the set of dominating point(s). For $x \in \Omega$, we denote $\Gamma_1(\{x\}, N)$ as $\Gamma_1(x)$. If $X \in \mathcal{X}_n$, then $\Gamma_1(X)$ and $\Gamma_1(\mathcal{X}_n)$ are random sets.

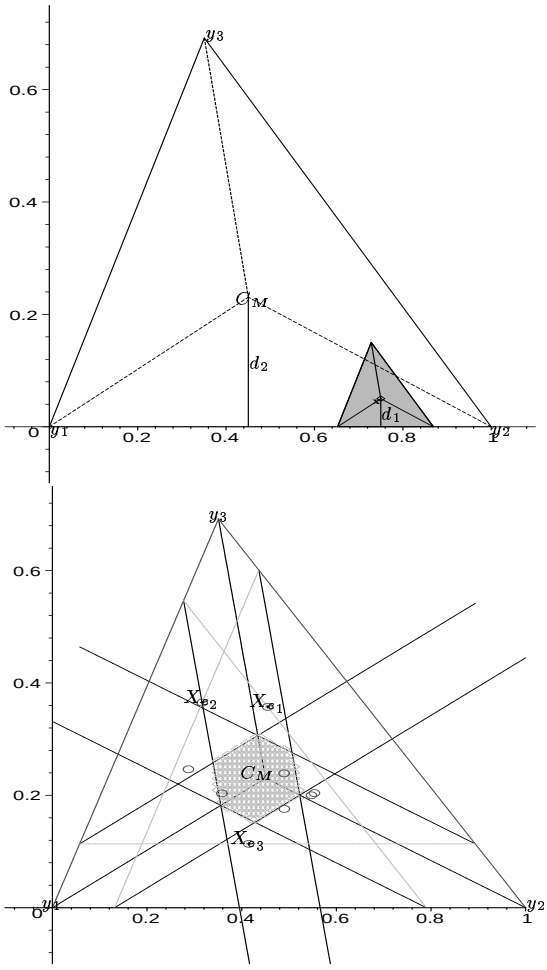

Figure 2: $N_{CS}(x)$ for an $x \in R(\overline{AB})$ (top); Γ_1 region for $n = 9$ for $N_{CS}(\cdot)$ (bottom).

We present first a "geometry invariance" result which will simplify our subsequent analysis by allowing us to consider the special case of the equilateral triangle.

Theorem 2.1. *: Let $\mathcal{Y} = \{y_1, y_2, y_3\} \subset \mathbb{R}^2$ be three non-collinear points. For $i = 1, \cdots, n$, let $X_i \overset{iid}{\sim} F = \mathcal{U}(T(\mathcal{Y}))$, the uniform distribution on the triangle $T(\mathcal{Y})$. Then the distribution of γ_n and of $A(N_{CS}(X))/A(T(\mathcal{Y}))$, $A(\Gamma_1(X))/A(T(\mathcal{Y}))$ and $A(\Gamma_1(\mathcal{X}_n))/A(T(\mathcal{Y}))$ are independent of $\mathcal{Y}$, and hence the geometry of $T(\mathcal{Y})$.*

Based on Theorem 2.1 and our uniformity assumption, we may assume that $T(\mathcal{Y})$ is a standard equilateral triangle with $\mathcal{Y} = \{(0,0), (1,0), (1/2, \sqrt{3}/2)\}$ henceforth.

3 Γ_1-Regions

A *median line* in a triangle is the line joining a vertex to the midpoint of the opposite edge. The *edge extremum* for edge e_j, denoted X_{e_j}, is the point closest to e_j; i.e., $X_{e_j} := \text{argmin}_{X \in \mathcal{X}_n} d(X, e_j)$ for $j = 1, 2, 3$. We use median lines and edge extrema to describe $\Gamma_1(\mathcal{X}_n)$. For each edge, drawing a line at X_{e_j} parallel to e_j yields a new triangle with y_j. The median lines at the vertices (other than y_j) of this new triangle determine the Γ_1-region, $\Gamma_1(\mathcal{X}_n)$. See Figure 2 (bottom).

Let $\mathcal{F}$ be the class of all distributions on triangle $T(\mathcal{Y})$, and $\mathcal{F}_C$ be the class of distributions with continuous density on $T(\mathcal{Y})$. And let $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a random set iid F on $T(\mathcal{Y})$. Following are some results regarding the Γ_1-regions for N_{CS}.

Proposition 3.1. *Let $\mathcal{X}_n$ be from an $F \in \mathcal{F}_C$. Then $\Gamma_1(\mathcal{X}_n)$ is a convex hexagon a.s..*

Lemma 3.2. *Let $\mathcal{X}_n$ be from an $F \in \mathcal{F}$. Then $\Gamma_1(\mathcal{X}_n) = \cap_{j=1}^3 \Gamma_1(X_{e_j})$. But in general for any proximity map $\Gamma_1(\mathcal{X}_n) = \cap_{j=1}^n \Gamma_1(X_j)$.*

Theorem 3.3. *For a sequence of random variables $X_1, X_2, X_3, \ldots \overset{iid}{\sim} F \in \mathcal{F}$, let $\mathcal{X}(n) := \mathcal{X}(n-1) \cup \{X_n\}$, and $\mathcal{X}(0) := \emptyset$. Then $\Gamma_1(\mathcal{X}(n))$ is non-increasing in n, in the sense that $\Gamma_1(\mathcal{X}(n)) \supseteq \Gamma_1(\mathcal{X}(n+1))$ a.s.*

Corollary 3.4. *For $\mathcal{X}(n)$ defined as above for the sequence from an $F \in \mathcal{F}_C$, $\Gamma_1(\mathcal{X}(n)) \downarrow \{C_M\}$ as $n \to \infty$ a.s., in the sense that $\Gamma_1(\mathcal{X}(n)) \supseteq \Gamma_1(\mathcal{X}(n+1))$ and $A(\Gamma_1(\mathcal{X}(n))) \downarrow 0$ a.s.*

Note however that $\Gamma_1(\mathcal{X}_n)$ is neither non-increasing nor strictly decreasing, because we might have $\Gamma_1(\mathcal{X}_n) \subset \Gamma_1(\mathcal{X}_m)$ for some $m > n$. Nevertheless, Proposition 3.5 and Theorem 3.6 hold.

Proposition 3.5. *For positive integers $m > n$, let $\mathcal{X}_m$ and $\mathcal{X}_n$ be from an $F \in \mathcal{F}$. Then $A(\Gamma_1(\mathcal{X}_m)) \leq^{ST} A(\Gamma_1(\mathcal{X}_n))$.*

Theorem 3.6. *Let $\{\mathcal{X}_n\}_{n=1}^{\infty}$ be a sequence of data sets from $F \in \mathcal{F}$. Then $\Gamma_1(\mathcal{X}_n) \to \{C_M\}$ as $n \to \infty$ a.s.*

Proposition 3.7. *Let $\mathcal{X}_n$ be a data set of size n from an $F \in \mathcal{F}_C$, then $P(\gamma_n = n) > 0$ for all $n < \infty$.*

Proposition 3.8. *The expected area of the Γ_1-region, $E[A(\Gamma_1(\mathcal{X}_n))] \to 0$ as $n \to \infty$.*

Proposition 3.9. *The domination number γ_n is between 3 and 6 with probability 1 as $n \to \infty$, i.e. $P(3 \le \gamma_n \le 6) \to 1$ as $n \to \infty$.*

For an $F \in \mathcal{F}_C$, edge extrema are distinct with probability 1 as $n \to \infty$, so we use only the asymptotically accurate joint density of the edge extrema in propositions 3.8 and 3.9.

4 Support Estimation

Estimating the support of a distribution, especially in high dimensions, has received considerable attention in the statistical literature. For example, in [8], Scholkopf et al. give a brief survey and propose a new method to estimate a region that has a high probability, say θ. If a region of minimum volume with some regularity conditions has probability $\theta = 1$, this region yields an estimate for the support of the density. We propose the use of $N_{CS}(\cdot)$ in estimating the support of $\mathcal{X}$ using $\mathcal{Y}$, or vice versa where we implicitly assume that $\min(n, |\mathcal{Y}|) \ge d + 1$.

Let f_X and f_Y be the densities and $\mathcal{S}_X$ and $\mathcal{S}_Y$ the supports of X and Y, respectively. If $d(\mathcal{S}_X, \mathcal{S}_Y) := \inf\{d(x,y) : (x,y) \in \mathcal{S}_X \times \mathcal{S}_Y\} = \delta > 0$, we say $\mathcal{S}_X$ and $\mathcal{S}_Y$ are *δ-separable*. Also let $\widehat{S}_n := [\cup_{x \in \mathcal{X}_n \cap \mathcal{C}_H(\mathcal{Y})} N_{CS}(x)] \cup [\cup_{x \in \mathcal{X}_n \setminus \mathcal{C}_H(\mathcal{Y})} (V_C(x) \setminus \mathcal{C}_H(Y))]$, where $V_C(x)$ is the Voronoi cell of x constructed by using all X and Y points shuffled together and for $\theta \ge 0$, $S_\theta^X := \{x : f_X(x) \ge \theta\}$. Then we have

Proposition 4.1. *Suppose $\mathcal{S}_X$ and $\mathcal{S}_Y$ are δ-separable with $\delta > 0$. Then for fixed $|\mathcal{Y}|$, $S_\theta^X \subset \widehat{S}_n$ a.s. for all $\theta \ge 0$ for sufficiently large n and $P(X \notin \widehat{S}_n \text{ and } Y \in \widehat{S}_n) \to 0$ (Bayes optimal error for δ-separable densities) as $n \to \infty$.*

Recall the spherical proximity region $N_S(x) = B(x, r(x))$ and arc-slice proximity region $N_{AS}(x) = B(x, r(x)) \cap T(\mathcal{Y})$. Let $\widehat{S}_n(N_S) := \cup_{x \in \mathcal{X}_n} N_S(x)$ and $\widehat{S}_n(N_{AS}) := [\cup_{x \in \mathcal{X}_n \cap \mathcal{C}_H(\mathcal{Y})} N_{AS}(x)] \cup [\cup_{x \in \mathcal{X}_n \setminus \mathcal{C}_H(\mathcal{Y})} (V_C(x) \setminus \mathcal{C}_H(Y))]$, and $\widehat{S}_n(N_{CS})$ be defined as above. Also let $\Pi(N) := P(X \notin \widehat{S}_n(N) \text{ and } Y \in \widehat{S}_n(N))$, then

Proposition 4.2. *For $\mathcal{X}_n$ and $\mathcal{Y}$ as above with $|\mathcal{Y}|$ fixed, $\Pi(N_{CS}) \le \Pi(N_{AS}) \le \Pi(N_S)$ for sufficiently large n and $\Pi(N_S) \to 0$ as $n \to \infty$.*

For a realization of $\widehat{S}_n$ restricted to $\mathcal{C}_H(\mathcal{Y})$ with the same $\mathcal{X}_n$ and $\mathcal{Y}$ points of 1 (which are not separable), see Figure 3 (top). For a realization of $\widehat{S}_n$ restricted to $\mathcal{C}_H(\mathcal{Y})$ with separable $\mathcal{X}_n$ and $\mathcal{Y}$ with $\delta = .01$, see Figure 3 (bottom). For the latter, 200 X points are generated from $\mathcal{U}(P_1)$ and 50 Y points from $\mathcal{U}(P_2)$, where P_1 is the (convex) polygon with the vertices $(0,0), (\frac{1}{4}, 0), (\frac{3}{4}, \frac{1}{2}), (\frac{1}{4}, 1), (0,1)$ and P_2 is the polygon with vertices $(\frac{1}{4}+\delta, 0), (\frac{3}{4}+\delta, \frac{1}{2}), (\frac{1}{4}+\delta, 1), (1,1), (1,0)$.

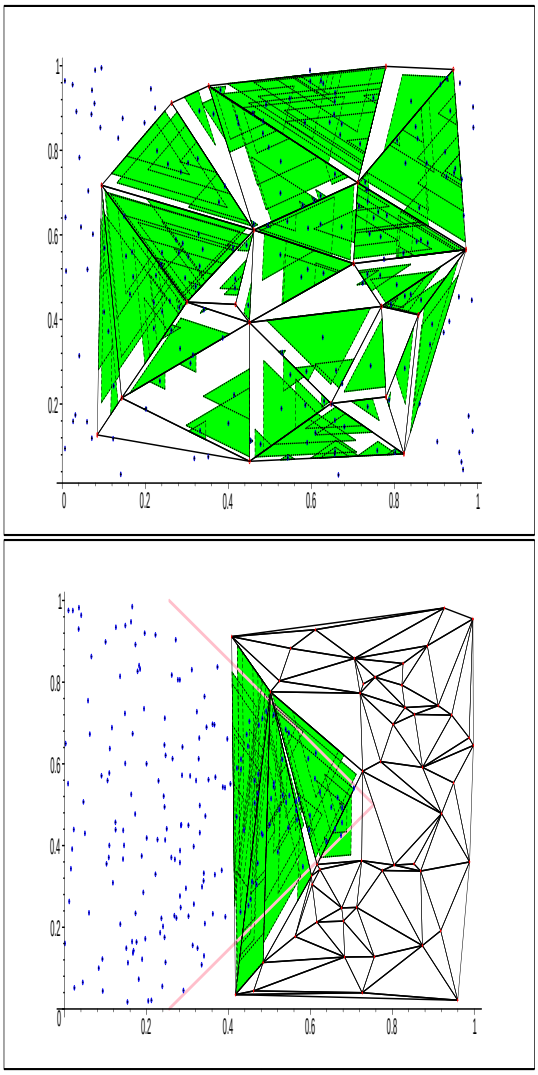

Figure 3: $\widehat{S}_n \cap \mathcal{C}_H(\mathcal{Y})$ for the data sets in Figure 1 (top); the separable data sets (bottom).

As an experimental application, we use a minefield data set from The Coastal Battlefield Reconnaissance and Analysis (COBRA) Program [9], [12]. The observations are detections of mines and minelike targets obtained from an unmanned aerial vehicle via a multispectral sensor.

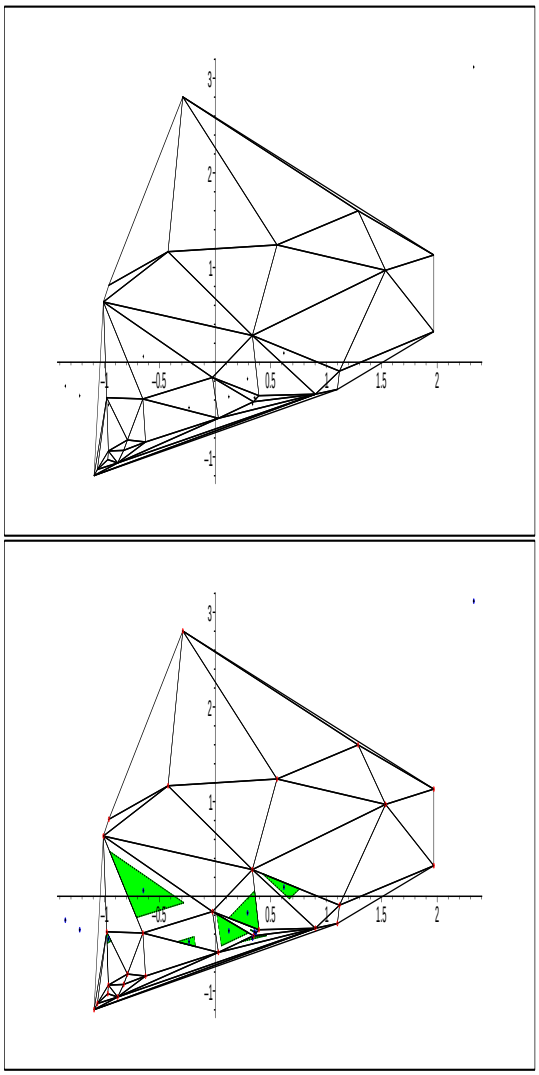

Figure 4: The Delaunay triangulation for the minefield data (top); the associated $\widehat{S}_n \cap \mathcal{C}_H(\mathcal{Y})$ (bottom).

We treat true mines as X points and false detections as Y points. In this data set, $n = 12$, $|\mathcal{Y}| = 27$. The original data set consists of 6-dimensional imagery, but Priebe et al. [7] and Olson et al. [4] demonstrate that dimensions 3 and 5 carry most of the relevant information for our purposes, so the data can be viewed in $\mathbb{R}^2$. See Figure 4 for the Delaunay triangulation and the corresponding support estimate restricted to $\mathcal{C}_H(\mathcal{Y})$, respectively.

As a final remark we note that, in the above procedure, we obtain $\widehat{S}_n^X$, and if we switch the roles of $\mathcal{X}_n$ and $\mathcal{Y}$ above, we obtain $\widehat{S}_n^Y$ also. These estimates can then be used to build discriminant regions for classification in a manner analogous to the procedure proposed in [6].

5 Extension to Higher Dimensions

Let $\Omega = \mathbb{R}^d$ with $d > 1$. The extensions of $N_S(\cdot)$ and $N_{AS}(\cdot)$ are straightforward and described above.

Let $\mathcal{Y} = \{y_1, y_2, \ldots, y_{d+1}\} \subset \mathbb{R}^d$ be $d+1$ non-collinear points. Let $\mathfrak{S}(\mathcal{Y})$ be the simplex formed by these $d+1$ points and φ_j be the face opposite vertex y_j for $j = 1, 2, \ldots, d+1$. For $x \in \mathfrak{S}(\mathcal{Y})$, the *central similarity proximity region* $N_{CS}(x)$ is defined to be the simplex similar to $\mathfrak{S}(\mathcal{Y})$ and adjacent to a face φ of $\mathfrak{S}(\mathcal{Y})$ with its centroid at x. To choose the adjacent face in $\mathfrak{S}(\mathcal{Y})$, we use *centroidal edge regions* that are simplicies with the vertices of $\mathfrak{S}(\mathcal{Y})$ adjacent to face φ and C_M. So the similarity ratio is $d(x, \varphi)/d(C_M, \varphi)$. Γ_1-regions can be extended similarly. Moreover, the results in section 3 (except Proposition 3.9) and section 4 can be extended mutatis mutandis to $\mathbb{R}^d$ with $d > 2$. Letting $\gamma_n(d)$ be the domination number for $\mathbb{R}^d$ with $d > 1$, we have $\lim_{n\to\infty} P(\gamma_n(d) > 1) = 1$, also we believe that there exist $\kappa_1(d) > 1$ and $\kappa_2(d) < n$ such that $\lim_{n\to\infty} P(\kappa_1(d) \leq \gamma_n(d) \leq \kappa_2(d)) = 1$. We conjecture that $\kappa_1(d) = (d+1)$ and $\kappa_2(d) = d\,(d+1)$.

6 Acknowledgements

This work partially supported by Office of Naval Research Grant N00014-01-1-0011 and by Defense Advanced Research Projects Agency Grant F49620-01-1-0395. Minefield data was provided by NSWC Coastal Systems Station, Dahlgren Division, Panama City, Florida, with support from United States Marine Corps Amphibious Warfare Technology; special thanks to R.R. Muise.

References

[1] G. Chartrand and L. Lesniak. *Graphs & Digraphs.* Chapman & Hill, 1996.

[2] L. Devroye, L. Gyorfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition.* Springer Verlag, New York, 1996.

[3] A. Okabe, B. Boots, and Sugihara K. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams.* Wiley, 1992.

[4] T. Olson, J.S. Pang, and C.E. Priebe. A likelihood–mpec approach to target classification. *Mathematical Programming, Ser. A*, 96:1–31, 2003.

[5] C.E. Priebe, J.G. DeVinney, and D.J. Marchette. On the distribution of the domination number of random class catch cover digraphs. *Statistics and Probability Letters*, 55:239–246, 2001.

[6] C.E. Priebe, D.J. Marchette, J. DeVinney, and D. Socolinsky. Classification using class cover catch digraphs. *Journal of Classification*, 20(1):3–23, 2003.

[7] C.E. Priebe, J.S. Pang, and T. Olson. Optimizing sensor fusion for classification performance. In *Proceedings of the International Conference on Imaging Science, Systems, and Technology: CISST'99*, pages 397–403, 1999.

[8] B. Scholkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, and R.C. Williamson. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13:1443–1471, 2001.

[9] D.L. Smith. Detection technologies for mines and minelike targets. *SPIE : Detection Technologies for Mines and Minelike Targets*, 2496:404–408, 1995.

[10] G.T. Toussaint. The relative neighborhood graphs of a finite planar set. *Pattern Recognition*, 12:261–268, 1980.

[11] G.T. Toussaint and J.W. Jaromczyk. Relative neighborhood graphs and their relatives. *Proceedings of IEEE*, 80:1502–1517, 1992.

[12] N.H. Witherspoon, J.H. Holloway, K.S. Davis, R.W. Miller, and A.C. Dubey. The coastal battlefield reconnaissance and analysis (cobra) program for minefield detection. *SPIE : Detection Technologies for Mines and Minelike Targets*, 2496:500–508, 1995.