ON THE QUESTION OF OBTAINING OPTIMAL PARTITIONS Of POINT SETS IN E^d WITH HYPERPLANE CUTS

By

Paul Lemke

IMA Preprint Series # 452

October 1988

ON THE QUESTION OF OBTAINING OPTIMAL PARTITIONS OF POINT SETS IN E^d WITH HYPERPLANE CUTS

PAUL LEMKE*

Abstract. Define the *average dissimilarity* of two non-empty multisets A, $B \subset E^d$ to be the average Euclidean distance between all $|A||B|$ pairs of points with one point in A, and one point in B. I show that the partition of a multiset $S \subset E^d$ into non-empty multisets of maximum dissimilarity may be accomplished with a hyperplane cut when $d = 1$, which leads to an efficient algorithm to find this partition. I show, however, that a hyperplane cut does not always produce the maximum dissimilarity partition when $d \geq 2$, although perhaps it always comes close.

Introduction. The problem of partitioning a finite set (or, more generally, a *multiset*, a set which may have repeated members) of points in E^d into two (or more) subsets which are as distinct as possible is a problem that arises in pattern recognition and classification theory, where the points each represent a list of d pieces of numerical data about some object in a population of such objects. The mathematical nature of the problem depends upon just what we mean by "distinct", which is in turn determined by the nature of the patterns or classes that we wish to identify. One of the most widely used methods to identify clusters of points in a set is the "bottom up" method, whereby each point is considered to be a cluster, and the two clusters which are nearest to each other (by some metric) are joined to make a larger cluster located at the center of mass of the two, with this process being continued until we are left with just 2 clusters. This method has the advantage of efficient computability, but does not always produce two subsets of maximum distinctness as we might have chosen to define "distinctness". I investigate here an issue bearing on the difficulty of computing the partition of maximum distinctness, where the measure of distinctness is simply the average Euclidean distance between points in opposite subsets.

The problem may be stated mathematically as follows: For any two finite non-empty multisets $A, B \subset E^d$ define the *average dissimilarity* $d(A, B)$ of A and B by:

$$d(A,B) = \frac{\sum_{a \in A} \sum_{b \in B} \rho(a,b)}{|A||B|}, \tag{1}$$

where $\rho(x, y)$ denotes the standard Euclidean distance between x and y.

Problem 1. Given a finite multiset $S \subset E^d$, find non-empty multisets A and B which maximize $d(A, B)$ subject to

$$A \cup B = S, \tag{2}$$

$$\text{and} \quad A, B \neq \phi. \tag{3}$$

*Institute for Mathematics and its Applications, University of Minnesota, 514 Vincent Hall, 206 Church Street S.E., Minneapolis, Minnesota 55455

If for some multiset S a pair of multisets A and B satisfy (2) and (3) we say that the unordered pair $\{A, B\}$ is a ***partition*** of S. If A and B also solve Problem 1 we say that $\{A, B\}$ is an ***optimal partition*** of S.

As one might suspect intuitively, solving problem 1 tends to produce a partition which is "polarized", with multisets A and B at opposite ends of the multi set S. In fact, there was some empirical evidence that the solution is always completely polarized, with the convex hull of A having no points in common with the convex hull of B ([1]), so that A and B are separable by a hyperplane in E^d. If this were true, it would not only give a neat partial characterization of optimal partitions in E^d, but would result in an efficient algorithm for finding them for small d, since the number of partitions of a multiset S generated by hyperplane cuts increases only polynomially with $|S|$ for fixed d. We therefore consider:

Conjecture 1. If $S \subset E^d$ is a finite multiset with nonzero diameter and $\{A, B\}$ is an optimal partition of S, then $hull(A) \cap hull(B) = \phi$, or, equivalently, there exists $v \in E^d$ and a scalar c such that

$$v \cdot x < c \quad \text{for all} \quad x \in A, \tag{4a}$$

$$\text{and} \quad v \cdot x > c \quad \text{for all} \quad x \in B. \tag{4b}$$

Problem 1 is easier to analyze if we add an additional constraint equating $\min(|A|, |B|)$ to some fixed value k (and hence equating $\max(|A|, |B|)$ to $|S| - k$), because then we need only maximize the numerator of the right-hand side of (1). We call a partition $\{A, B\}$ of S a *k-optimal* partition of S if it maximizes (1) subject to the constraint $\min(|A|, |B|) = k$, in addition to constraints (2) and (3). Clearly, all solutions to Problem 1 for a given S are k-optimal partitions of S for some value of k.

If $k = 1$, then (1) reduces to

$$\frac{\sum_{y \in S} \rho(x, y)}{|S| - 1} \quad . \tag{5}$$

where x is the sole member of the cardinality 1 member of $\{A, B\}$. Since $\rho(x, y)$ is a convex function of x for fixed y and strictly convex at $x = y$, (5) is a convex function of x and is strictly convex at each of the points of S. It follows that (5) attains its maximum in $hull(S)$ at an extreme point of $hull(S)$ (which will be a point of S), but that (5) does not attain a maximum in $hull(S)$ at a point of S which is not an extreme point of $hull(S)$. Since every extreme point of a polytope P can be separated by a hyperplane from all but an arbitrarily small chunk of P by a hyperplane, the existence of v and c satisfying (4a) and (4b) follows, provided that the point x at which the maximum is attained is not a multiple point of S. If it is, one can still use a tangent hyperplane through x to satisfy a weaker version of Conjecture 1 where (4a) and (4b) are non-strict inequalities. Thus a couple of slightly weaker versions of Conjecture 1 are true for the more constrained problem for $k = 1$.

For $2 \le k \le \frac{1}{2}|S|$, however, we note that if we define $S \subset E^1$ by:

$$S = \{0\} \cup \{1\}^n \cup \{2\}^{k-1},$$

where $n \ge k$, then it is straightforward to verify that $\{A, B\}$, where

$$A = \{0\} \cup \{2\}^{k-1},$$
$$\text{and} \qquad B = \{1\}^n,$$

is a k-optimal partition of S (in fact, the *unique* k-optimal partition of S if $n \ge k+1$). Thus Conjecture 1 is false for some k-optimal partitions for any values of $|S| \ge 4$ and $2 \le k \le \frac{1}{2}|S|$. Since this example automatically generalizes to all $d \ge 2$, we may consider the status of Conjecture 1 settled for the k-optimization version of Problem 1. For the remainder of this paper we therefore consider only the original version of Problem 1. (It may be noted that the optimal partition $\{A, B\}$ of $S = \{0\} \cup \{1\}^n \cup \{2\}^{k-1}$ is then given by $A = \{0\}$ and $B = \{1\}^n \cup \{2\}^{k-1}$, which does not violate Conjecture 1.)

LEMMA 1. *If f and g are functions mapping $\{n, n+1, \ldots, n+m\}$ into the positive reals, with f convex and g strictly concave, then $\frac{f(k)}{g(k)}$ does not attain a local maximum on $\{n+1, n+2, \ldots, n+m-1\}$.*

Proof. Assume to the contrary, that $\frac{f(k)}{g(k)} \ge \frac{f(k-1)}{g(k-1)}$ and $\frac{f(k)}{g(k)} \ge \frac{f(k+1)}{g(k+1)}$ for some $k \in \{n+1, n+2, \ldots, n+m-1\}$. Then from the first inequality we derive:

$$f(k)g(k-1) \ge f(k-1)g(k); \tag{6}$$

and from the second we derive:

$$f(k)g(k+1) \ge f(k+1)g(k). \tag{7}$$

Adding (6) to (7) and then dividing by 2 gives

$$f(k)(\tfrac{1}{2}\, g(k-1) + \tfrac{1}{2}\, g(k+1)) \ge (\tfrac{1}{2}\, f(k-1) + \tfrac{1}{2}\, f(k+1))g(k).$$

But from the positivity of f and the strict concavity of g we have:

$$f(k)g(k) > f(k)(\tfrac{1}{2}\, g(k-1) + \tfrac{1}{2}\, g(k+1)),$$

and from the positivity of g and the convexity of f we have

$$(\tfrac{1}{2}\, f(k-1) + \tfrac{1}{2}\, f(k+1))g(k) \ge f(k)g(k),$$

which, together with the previous inequalities gives $f(k)g(k) > f(k)g(k)$, a contradiction. Therefore the maximum of $\frac{f(k)}{g(k)}$ in $\{n, n+1, \ldots, n+m\}$ is attained at either $k = n$ or $k = n+m$. □

THEOREM 2. *If $\{A, B\}$ is an optimal partition of $S \subset E^d$ and for some positive integer n and some $x \in E^d$ $\{X\}^n \subset S$ and $|S| \geq n+1$, then either $\{X\}^n \subset A$ or $\{X\}^n \subset B$.*

Proof. There must exist some integer k, with $0 \leq k \leq n$, and multisets $\overline{A}$ and $\overline{B}$ such that

$$A = \overline{A} \cup \{x\}^k$$
$$\text{and} \qquad B = \overline{B} \cup \{x\}^{n-k}.$$

If $\overline{A}$ is empty, it is easy to see that $d(A, B)$ with A and B defined as above is maximized when $k = n$. Similarly, if $\overline{B}$ is empty, we have $k = 0$. Otherwise, we note that the denominator of the right side of (1), $(|\overline{A}| + k)(|\overline{B}| + n - k)$, is a positive, strictly concave function of k, while the numerator, which may be written as:

$$\sum_{a \in \overline{A}} \sum_{b \in \overline{B}} \rho(a, b) + k \sum_{b \in \overline{B}} \rho(x, b) + (n-k) \sum_{a \in \overline{A}} \rho(a, x) + k(n-k)\rho(x, x),$$

is convex in k, since the last term is 0. Lemma 1 therefore applies to show that $d(A, B)$ cannot be maximal unless $k = 0$ or $k = n$. □

COROLLARY. *If $\{A, B\}$ is an optimal partition of $S \subset E^1$, then there exists an integer n and $x_1, x_2, \ldots, x_{n+1} \in E^1$ such that*

i) $x_1 < x_2 < \cdots < x_{n+1}$;

ii) *If the open interval (x_k, x_{k+1}) is denoted by I_k for $k = 1, 2, \ldots, n$, then* $S \subset \bigcup_{k=1}^{n} I_k$;

iii) *For $k = 1, 2, \ldots, n$, either $A \cap I_k$ or $B \cap I_k$ is empty, but not both;*

iv) *For $k = 1, 2, \ldots, n-1$, $A \cap (I_k \cup I_{k+1}) \neq \phi$ and $B \cap (I_k \cup I_{k+1}) \neq \phi$.*

Definition. For any multiset $S \subset E^1$ define $\underline{M(S)}$ by:

$$M(S) = \left\{ \frac{1}{|S|} \sum_{x \in S} x \right\}^{|S|}.$$

Thus $M(S)$ is a multiset of cardinality $|S|$ with all its points located at the center of mass of the points of S.

THEOREM 3. *If $\{A, B\}$ is a partition of $S \subset E^1$ and C is a submultiset of A such that $hull(C) \cap B = \phi$, and if A' is defined by $A' = (A - C) \cup M(C)$, then*

$$\sum_{a \in A'} \sum_{b \in B} \rho(a, b) = \sum_{a \in A} \sum_{b \in B} \rho(a, b).$$

Proof. If $b \in B$, then $b \notin hull(C)$, so that for fixed b the numbers $\{b-x\}$, $x \in C$, must all have the same sign. Therefore

$$\sum_{x\in C}\rho(b,x)=\sum_{x\in C}|b-x| = \left|\sum_{x\in C}(b-x)\right| = \left|\sum_{x\in C}b-\sum_{x\in C}x\right| = \left|\sum_{x\in M(C)}b-\sum_{x\in M(C)}x\right|$$

$$=\sum_{x\in M(C)}|b-x| = \sum_{x\in M(C)}\rho(b,x), \quad \text{so that } \sum_{x\in M(C)}\sum_{b\in B}\rho(x,b)=\sum_{x\in C}\sum_{b\in B}\rho(x,b),$$

and the conclusion follows. □

Of course, an equivalent result holds if "A" and "B" are interchanged in the above theorem.

THEOREM 4. *Let $\{A,B\}$ be an optimal partition of $S \subset E^1$ and let the intervals $I_1,\ldots,I_n$ be as in the Corollary to Theorem 2. For $k=1,2,\ldots,n$, define:*

$$S_k = S\cap I_k, \quad k=1,2,\ldots,n,$$

$$\overline{S}=\bigcup_{k=1}^{n} M(S_k),$$

$$\overline{A}=\bigcup_{k} M(S_k), \quad \textit{the union being taken over all}$$

k for which $I_k \cap A$ is non-empty,

$$\textit{and} \quad \overline{B}=\bigcup_{k} M(S_k), \; \textit{the union being taken over all}$$

k for which $I_k \cap B$ is non-empty.

Then $\{\overline{A},\overline{B}\}$ is an optimal partition of $\overline{S}$.

Proof. Since the points of $M(S_k)$ are contained in I_k and since the intervals $\{I_k\}$ are disjoint, Theorem 3 may be applied n times consecutively (letting $C=S_1,S_2,\ldots,S_n$) to show that $d(\overline{A},\overline{B})=d(A,B)$.

Also, if $\{C,D\}$ is an optimal partition of $\overline{S}$, then we can define multisets $\overline{C}$ and $\overline{D}$ by:

$$\overline{C}=\bigcup_{k} S_k, \text{ the union being taken over all } k \text{ for which } M(S_k)\cap C \text{ is non-empty},$$

$$\text{and } \overline{D}=\bigcup_{k} S_k, \text{ the union being taken over all } k \text{ for which } M(S_k)\cap D \text{ is non-empty}.$$

From Theorem 2 we know that each of the submultisets $M(S_k)$ of $\overline{S}$ contains points from exactly one of the partition sets C and D, so that $\overline{C}\cap\overline{D}=\phi$ and $\overline{C}\cup\overline{D}=S$. We can then apply Theorem 3 again to show that $d(\overline{C},\overline{D})=d(C,D)$. But since $\{\overline{C},\overline{D}\}$ is a partition of S and $\{A,B\}$ is an optimal partition of S, we have $d(\overline{C},\overline{D})\le d(A,B)$ and therefore $d(C,D)\le d(\overline{A},\overline{B})$. But since $\{C,D\}$ is optimal by definition, $\{\overline{A},\overline{B}\}$ must also be optimal. □

COROLLARY. *If there is a multiset $S \subset E^1$ and a partition $\{A, B\}$ of S which constitute a counterexample to Conjecture 1, then we can assume without loss of generality that A, B and S have the form:*

$$\begin{aligned} S &= \bigcup_{k=1}^{n} \{y_k\}^{m_k}, \\ A &= \bigcup_{k \text{ odd}, \le n} \{y_k\}^{m_k}, \\ \text{and} \quad B &= \bigcup_{k \text{ even}, \le n} \{y_k\}^{m_k}, \end{aligned}$$

where $n, m_1, m_2, \ldots, m_n$ are positive integers and $y_1, y_2, \ldots, y_n$ are real numbers such that $y_1 < y_2 < \cdots < y_n$.

The proof of this corollary uses Theorem 4, the symmetry between A and B, and the easily observed fact that $hull(\overline{A}) \cap hull(\overline{B}) \neq \phi$ if $hull(A) \cap hull(B) \neq \phi$.

THEOREM 5. *If S, A, and B are of the form given in the corollary to Theorem 4, and $n \geq 4$, then $\{A, B\}$ is not an optimal partition of S.*

Proof. Define:

$$\begin{aligned} e &= \min(m_2, m_3) \\ A_1 &= A \cup \{y_2\} - \{y_1\} \\ B_1 &= B \cup \{y_1\} - \{y_2\} \\ A_2 &= A \cup \{y_2\}^e - \{y_3\}^e \\ B_2 &= B \cup \{y_3\}^e - \{y_2\}^e \\ A_3 &= A \cup \{y_4\} - \{y_3\} \\ B_3 &= B \cup \{y_3\} - \{y_4\}. \end{aligned}$$

Then $\{A_1, B_1\}, \{A_2, B_2\}$ and $\{A_3, B_3\}$ are each partitions of S. Noting that $|A||B| = |A_1||B_2| = |A_2||B_2| = |A_3||B_3|$, we can calculate:

$$|A||B|(d(A_1, B_1) - d(A, B)) = (y_2 - y_1)(2 - m_1 - m_2 + m_3 - \sum_{i=4}^{n}(-1)^i\, m_i) \tag{8}$$

$$|A||B|(d(A_2, B_2) - d(A, B)) = e(y_3 - y_2)(2e + m_1 - m_2 - m_3 + \sum_{i=4}^{n}(-1)^i\, m_i) \tag{9}$$

$$|A||B|(d(A_3, B_3) - d(A, B)) = (y_4 - y_3)(2 - m_1 + m_2 - m_3 - \sum_{i=4}^{n}(-1)^i\, m_i). \tag{10}$$

If we multiply (8) by $(y_2 - y_1)^{-1}$ and add the result to the product of (9) and $e^{-1}(y_3 - y_2)^{-1}$, we get:

$$2e + 2 - 2m_2. \tag{11}$$

If we multiply (9) by $e^{-1}(y_3 - y_2)^{-1}$ and add the result to the product of $(y_4 - y_3)^{-1}$ and (10) we get:

$$(12) \qquad 2e + 2 - 2m_3.$$

Since $e = \min(m_2, m_3)$, one of (11) and (12) must be positive, and since each of these is a non–negative linear combination of (8), (9) and (10), one of these quantities must be positive, thereby proving that $\{A, B\}$ is non-optimal. □

THEOREM 6. *If S, A and B are of the form given in the corollary to Theorem 4, and $n = 3$, then $\{A, B\}$ is not an optimal partition of S.*

Proof. If $y_2 - y_1 \geq y_3 - y_2$, then since $d(A, B)$ represents the average distance between points in A and points in B, and all those distances are either $y_2 - y_1$ or $y_3 - y_2$, we must have $d(A, B) \leq y_2 - y_1$. But clearly the partition $\{C, D\}$, where

$$C = \{y_1\}^{m_1},$$
$$\text{and} \quad D = \{y_2\}^{m_2} \cup \{y_3\}^{m_3},$$

has $d(C, D) > y_2 - y_1$, since all interset distances are greater than or equal to $y_2 - y_1$, with some greater. Therefore $\{A, B\}$ is non-optimal if $y_2 - y_1 \geq y_3 - y_2$. Since the case $y_3 - y_2 \geq y_2 - y_1$ is symmetrically equivalent, $\{A, B\}$ is always non-optimal. □

COROLLARY. *Conjecture 1 is true for $d = 1$.*

Proof. Since Theorems 5 and 6 leave only the case of the corollary to Theorem 4 where $n = 2$, and since this case does not give a counterexample to Conjecture 1, we conclude that Conjecture 1 is true for $d = 1$.

COMMENT. It would seem that there should be a shorter proof for the $d = 1$ case of Conjecture 1 than Lemma 1–Theorem 6.

Algorithm 1. The above result leads to an easy algorithm to find the optimal partitions of multisets in E^1. If the points are $x_1, x_2, \ldots, x_n$, with $x_1 \leq x_2 \leq \cdots \leq x_n$, then we only have to find the value of k between 1 and $n - 1$ inclusive which maximizes:

$$d(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}) = \frac{\sum_{i=k+1}^{n} \sum_{j=1}^{k} (x_i - x_j)}{k(n-k)}$$

Since $\displaystyle\sum_{i=k+1}^{n} \sum_{j=1}^{k} (x_i - x_j) = k \sum_{i=k+1}^{n} x_i - (n-k) \sum_{j=1}^{k} x_j = ks_n - ns_k$, where

$$s_k = \sum_{i=1}^{k} x_i \quad \text{for} \quad k = 1, 2, \ldots, n-1,$$

we need only compute all values of s_k, and then compute and compare all values of:

$$\frac{ks_n - ns_k}{k(n-k)},$$

which can be done in $O(n)$ arithmetic operations (assuming that points were given originally in numerical order).

Comparison to Bottom Up Result. The partition produced by the bottom up method is not always the partition of maximum dissimilarity, and, in fact, may produce a partition whose dissimilarity is smaller than the maximum by an arbitrarily large factor. An example in E^1 is:

$$S = \bigcup_{i=0}^{\ell} \{i + \epsilon\tbinom{i}{2}\}^{(n^i)},$$

where ℓ is a given integer, ϵ is a small positive number, and n is a large integer. As $n \to \infty$, the average dissimilarity of the partition given by the bottom up method approaches $1 + (\ell - 1)\epsilon$, while the value for the optimal one approaches $\ell + \binom{\ell}{2}\epsilon$, so that as $\epsilon \to 0$, the ratio approaches ℓ. I do not know whether there are examples where the number of points increases only as a polynomial function of the ratio.

Like the partition obtained by the bottom–up method, however, the optimal partition of a set in E^1 does not separate the two closest points. The proof of this involves a fair amount of algebraic calculations, and I will omit it. I shall mention that this result is not true in E^2. An example is:

$$S = \{(-\sqrt{3}-1, 1), (-\sqrt{3}-1, -1), (-1+\epsilon, 0), (1-\epsilon, 0), (\sqrt{3}+1, 1), (\sqrt{3}+1, -1)\},$$

where ϵ is a sufficiently small positive number. The two closest points, $(-1+\epsilon, 0)$ and $(1-\epsilon, 0)$, are separated in the optimal partition, which places the first 3 points of S (in the order that they are listed above) in one set, and the remaining points in the other.

The Case $d = 2$. Conjecture 1 is false for $d \geq 2$ and I will illustrate a general (non-rigorous) method of constructing counterexamples.

If $\{A, B\}$ is an optimal partition of a multiset $S \subset E^2$ and p is some point $\in E^2$, we might expect that the optimal partition of $S \cup \{p\}$ might be either $\{A \cup \{p\}, B\}$ or $\{A, B \cup \{p\}\}$, depending on which side of the "equal dissimilarity curve", defined by $d(A \cup \{x\}, B) - d(A, B \cup \{x\}) = 0$, that p lies on. Since the Euclidean metric is nonlinear, we expect that the equal dissimilarity curve associated with most partitions $\{A, B\}$ will not be a straight line. There will therefore be colinear points p_1, p_2 and p_3, with p_2 between p_1 and p_3, such that p_2 is on one side of the curve and p_1 and p_3 are on the other. The optimal partition of $S \cup \{p_1\} \cup \{p_2\} \cup \{p_3\}$ may then have p_2 in one partition set and p_1 and p_3 in the other, in which case p_2 would be in the convex hulls of both partition sets. This would therefore give a counterexample to Conjecture 1 for $d = 2$.

There are, however, three obstacles to constructing a counterexample to Conjecture 1 in this manner:

i) The optimal partition of $S \cup \{p_1\} \cup \{p_2\} \cup \{p_3\}$ may divide up S differently than the optimal partition of S alone does.

ii) Even if the difficulty mentioned in i) does not occur, the optimal partition of $S \cup \{p_1\} \cup \{p_2\} \cup \{p_3\}$ may place p_i with A, for some $i \in \{1,2,3\}$, while the optimal partition of $S \cup \{p_i\}$ may place it with B, or vice versa. This may occur because, for example, the equal dissimilarity curve associated with the partition $\{A \cup \{p_1\}, B \cup \{p_2\}\}$ may be different from that associated with $\{A, B\}$, so that p_3 may be on different sides of the curves (relative to A and B). This may mean that the optimal partition of $S \cup \{p_1\} \cup \{p_2\} \cup \{p_3\}$ does not divide up $\{p_1, p_2, p_3\}$ in the desired way.

iii) The optimal partition of $S \cup \{p_1\} \cup \{p_2\} \cup \{p_3\}$ may put all the points of S in the same partition set.

Obstacles i) and ii) may be overcome by giving S and its optimal partition sets a sufficiently large number of points, keeping certain other things constant. This may be done by replacing S with S^n, since it follows from Theorem 2, and the fact that $d(A^n, B^n) = d(A, B)$ for any two multisets A and B, that if $\{A, B\}$ is an optimal partition of S, $\{A^n, B^n\}$ is an optimal partition of S^n. It is then not hard to show that obstacle i) will always be overcome if n is made sufficiently large, provided that the optimal partition $\{A, B\}$ of S is unique. Likewise, obstacle ii) will be overcome as $n \to \infty$ because the equal dissimilarity curve associated with a partition such as $\{A^n \cup \{p_1\}, B^n \cup \{p_2\}\}$ will differ from that associated with $\{A^n, B^n\}$ by $O(\frac{1}{n})$, as both will approach the "limiting equal dissimilarity curve" (LEDC) given by: $\sum_{y \in A} \rho(x,y) - \sum_{y \in B} \rho(x,y) = (|A| - |B|)d(A,B)$.

Obstacle iii) cannot be overcome just by making n large, but it will not be a problem for large n if $d(S, \{p_i\}) < d(A, B)$ for $i = 1, 2, 3$.

To construct a concrete counterexample we choose $S = \{0,0\}^2 \cup \{0,1\}$. (This is the simplest multiset whose optimal partition has a LEDC which is not a straight line.) The optimal partition $\{A, B\}$ of S is given by:

$$A = \{(0,0)\}^2$$
$$\text{and} \qquad B = \{(0,1)\},$$

with the LEDC being given by:

$$2\sqrt{x_1^2 + x_2^2} - \sqrt{x_1^2 + (1 - x_2)^2} = 1,$$
$$\text{or} \qquad 9x_2^4 + 12x_2^3 + (18x_1^2 - 12)x_2^2 + 12x_1^2 x_2 + 9x_1^4 - 16x_1^2 = 0.$$

Setting $x_1 = 0$ we find $x_2 = \frac{2}{3}$.

Setting $x_1 = \pm\frac{1}{2}$ we find $x_2 \cong 0.6362$.

A convenient rational number between $\frac{2}{3}$ and 0.6362 is $\frac{13}{20}$, and we may then choose as our three colinear points $(-\frac{1}{2}, \frac{13}{20})$ and $(0, \frac{13}{20})$ and $(\frac{1}{2}, \frac{13}{20})$. To deal only with integer coordinates, we now scale the example up by a factor of 20, and for $n = 1, 2, \ldots$ define the multisets $\{T_n\}$ by:

$$T_n = \{(0,0)\}^{2n} \cup \{(0,20)\}^n \cup \{(-10,13)\} \cup \{(0,13)\} \cup \{(10,13)\}.$$

We expect that for n sufficiently large the optimal partition of T_n will be the pair of multisets

$$\begin{aligned} &\{(0,0)\}^{2n} \cup \{(0,13)\} \\ \text{and} \quad &\{(0,20)\}^n \cup \{(-10,13)\} \cup \{(10,13)\}, \end{aligned}$$

whose convex hulls intersect at $(0,13)$.

We now attempt to verify this and to find out how large n must be. Let $\{A, B\}$ be an optimal partition of T_n for some n. From Theorem 2 we know that $\{(0,0)\}^{2n}$ and $\{(0,20)\}^n$ are each contained entirely in one of the two sets A and B. If they are in the sample partition set (say A), then it is easily verified that the maximum possible average distance from any point $b \in B$ to the points of A occurs when $b = (10,13)$ and $A = \{(0,0)\}^2 \cup \{(0,20)\} \cup \{(-10,13)\}$, when it is $\frac{\sqrt{149}+2\sqrt{269}+20}{4} \cong 16.25$, so that $d(A,B) \leq 16.26$. However, if $\{(0,0)\}^{2n}$ and $\{(0,20)\}^n$ are in different partitions sets, then $d(A,B) \geq \frac{40n^2}{2n(n+3)}$, and $\frac{40n^2}{2n(n+3)} = 20 - \frac{60}{n+3} > 16.26$ when $n \geq 14$. Thus, for $n \geq 14$ the optimal partition of T_n puts $\{(0,0)\}^{2n}$ and $\{(0,20)\}^n$ in different partition sets (say A and B, respectively).

We now look at the eight possible ways to assign $(-10,13), (0,13)$ and $(10,13)$ to A and B, and give the corresponding value of $d(A,B)$.

i) $(-10,13), (0,13), (10,13) \in A$	$\frac{40n^2+(2\sqrt{149}+7)n}{n(2n+3)}$
ii) $(-10,13), (0,13) \in A,\ (10,13) \in B$	$\frac{40n^2+(\sqrt{149}+2\sqrt{269}+7)n+30}{(n+1)(2n+2)}$
iii) $(-10,13), (10,13) \in A,\ (0,13) \in B$	$\frac{40n^2+(2\sqrt{149}+26)n+20}{(n+1)(2n+2)}$
iv) $(0,13), (10,13) \in A,\ (-10,13) \in B$	$\frac{40n^2+(\sqrt{149}+2\sqrt{269}+7)n+30}{(n+1)(2n+2)}$
v) $(-10,13) \in A,\ (0,13), (10,13) \in B$	$\frac{40n^2+(\sqrt{149}+2\sqrt{269}+26)n+30}{(n+2)(2n+1)}$
vi) $(0,13) \in A,\ (-10,13), (10,13) \in B$	$\frac{40n^2+(4\sqrt{269}+7)n+20}{(n+2)(2n+1)}$
vii) $(10,13) \in A,\ (-10,13), (0,13) \in B$	$\frac{40n^2+(\sqrt{149}+2\sqrt{269}+26)n+30}{(n+2)(2n+1)}$
viii) $(-10,13), (0,13), (10,13) \in B$	$\frac{40n^2+(4\sqrt{269}+26)n}{2(n+3)n}$

Calculations show that assignment vi) gives a larger value of $d(A, B)$ than the other assignments for all $n \geq 33$. Thus T_{33} provides a counterexample to Conjecture 1, with optimal partition sets

$$\begin{aligned} A &= \{(0,0)\}^{66} \cup \{(0,13)\} \\ \text{and} \quad B &= \{(0,20)\}^{33} \cup \{(-10,13)\} \cup \{(10,13)\}. \end{aligned}$$

It may be noted that this counterexample automatically generalizes to all $d \geq 3$. Also, since the above partition is the unique optimal partition of T_{33}, we can move the point at $(0,13)$ slightly to put it in the interior of hull(B), so that Conjecture 1 is still false even if (4a) and (4b) are changed to non-strict inequalities. Uniqueness also gives us the latitude to disperse the points at (0,0) and (0,20) slightly to give a counterexample with no multiple points.

It may also be noted that the method given above for constructing counterexamples should work for any metric ρ on E^2 such that there exist x and y in E^2 such that the LEDC associated with $\{\{x\}, \{y\}^2\}$ has a bend in it near the point where it intersects the line segment that is the convex hull of x and y. This seems to be the case for any reasonable metric one can think of.

There is also a counterexample to Conjecture 1 with points at only 4 distinct locations. The multiset:

$$S = \{(0,0)\} \cup \{(0,11)\}^{12} \cup \{(-6,0)\} \cup \{(6,0)\}$$

has the unique optimal partition $\{A, B\}$, where

$$\begin{aligned} A &= \{(0,0)\} \cup \{(0,11)\}^{12}, \\ \text{and} \quad B &= \{(-6,0)\} \cup \{(6,0)\}. \end{aligned}$$

Of course, there are no such counterexamples with 3 or fewer locations. However, there are probably counterexamples with fewer *points* than the 15 in the above example. Perhaps one could find some of them with the aid of a computer with a program capable of rapidly finding optimal partitions of small sets of points in E^2. I would suggest trying a set containing three colinear points sandwiched in between two small sets of points whose precise locations can be adjusted until computation shows that the middle point of the three colinear points belongs in a different partition set than the other two. Using the fact that the sum of the lengths of the diagonals of a quadrilateral in E^2 is greater than half the sum of the lengths of the sides, it is easy to show that Conjecture 1 holds for sets of 4 points in E^2. The cardinality of the smallest set for which Conjecture 1 fails, however, is otherwise an open question as of this writing.

Another open question of potentially greater importance involves finding the maximum ratio of the average dissimilarity of the optimum partition of a set S to the maximum average dissimilarity of the partitions of S obtainable by a hyperplane split. Precisely stated, if for any multiset $S \subset E$ we define $f(S)$ to be $d(A,B)$, where $\{A,B\}$ is an optimal partition of S, and we define $g(S)$ to be the maximum value of $d(C,D)$ over all partitions $\{C,D\}$ of S for which $hull(C) \cap hull(D) = \phi$, then we wish to know the supremum of $\frac{f(S)}{g(S)}$ over all multisets S in E^d, for fixed d. From the counterexamples to Conjecture 1 given in this paper, we know that the supremum is greater than one for $d \geq 2$. By adjusting the parameters of certain counterexamples to Conjecture 1, we can find the supremum of the ratio over all possible values of the parameters, and this will give us a lower bound on the supremum of $\frac{f(S)}{g(S)}$. In particular, for $S = \{(-1,0)\} \cup \{(0,0)\} \cup \{(1,0)\} \cup \{(0,y)\}^m$, the maximum occurs at $m \approx 34.30873744$ and $y \approx 2.480072911$, for which $\frac{f(S)}{g(S)} \approx 1.006614386$ (the fact that m is not an integer is not important because we can use an arbitrarily close rational approximation $\frac{k}{n}$ for m and then substitute S^n for S, since $f(S^n) = f(S)$ and $g(S^n) = g(S)$). More parameters besides m and y could be added to the example if a computer were used to perform the optimization.

On the other hand, if $x, y \in S$ are such that $\rho(x,y) = \max_{x,y \in S} \rho(x,y)$, then clearly $f(S) \leq \rho(x,y)$. Also, since $\rho(x,z) + \rho(y,z) \geq \rho(x,y)$ for each $z \in S$, it follows that $d(\{x\}, S - \{x\}) + d(\{y\}, S - \{y\}) \geq \rho(x,y)$, so that $g(S) \geq \frac{1}{2}\, \rho(x,y)$. We therefore have $\frac{f(S)}{g(S)} \leq 2$ for all finite sets $S \subset E^d$.

The above results leave considerable uncertainty in the value of the maximum of $\frac{f(S)}{g(S)}$.

Acknowledgement. I would like to thank Dr. David Matula for discussions concerning the nature and significance of this problem.

REFERENCE

[1] DAVID MATULA, *Personal Communication*, October, 1987.