Performance and Geometric Interpretation for Decision Fusion with Memory

Moshe Kam, *Senior Member, IEEE*, Chris Rorres, Wei Chang, and Xiaoxun Zhu

Abstract*—A binary distributed detection system comprises a bank of local decision makers (LDM's) and a central information processor (the data fusion center, DFC). All LDM's survey a common volume for a binary $\{H_0, H_1\}$ phenomenon. Each LDM forms a binary decision: it either accepts H_1 ("target-present") or H_0 ("target-absent"). The LDM is fully characterized by its performance probabilities (probability of false alarm and probability of detection). The decisions are transmitted to the DFC through noiseless communication channels. The DFC then optimally combines the local decisions to obtain a global decision ("target-present" or "target-absent") which minimizes a Bayesian objective function. Along with the local decisions, the DFC in our architecture remembers and uses its most recent decision in synthesizing each new decision. When operating in a stationary environment, our architecture converges to a steady-state decision in finite time with probability one, and its detection performance during convergence and in steady state is strictly determined. Once convergence is proven, we apply the results to the detection of signals with random phase and amplitude. We further provide a geometric interpretation for the behavior of the system: the unit square of the current (P_f, P_d) known to the DFC is partitioned into polygons, one of which defines a "stopping set" of values. If the current (P_f, P_d) falls into this region, there is no way to leave it, and hence there is no reason to continue testing.

NOMENCLATURE

$\beta(P_{fg}^{(t)}, P_{dg}^{(t)})$ — Global Bayes risk at time instant t for a 1-bit DFC memory system.

$\Delta\beta(\cdot,\cdot)$ — Difference in global probability of errors between two consecutive time instants $t+1$ and t: $\Delta\beta(P_{fg}^{(t)}, P_{dg}^{(t)}) = \beta(P_{fg}^{(t+1)}, P_{dg}^{(t+1)}) - \beta(P_{fg}^{(t)}, P_{dg}^{(t)})$.

$\delta_{\mathbf{i}}(\cdot,\cdot)$ — Function defined by $\delta_{\mathbf{i}}(x,y) = P_d^i(1-P_d)^{n-i}y - \tau P_f^i(1-P_f)^{n-i}x$.

$\gamma_{\mathbf{i}}(\cdot,\cdot)$ — Function defined by $\gamma_{\mathbf{i}}(x,y) = P_d^i(1-P_d)^{n-i}(1-y) - \tau P_f^i(1-P_f)^{n-i}(1-x)$.

$\Re_S$ — Stopping-set of a 1-bit DFC memory distributed detection system.

τ — Threshold of the DFC.

$[0,1]$ — Closed subset of the real between 0 and 1.

$\mathbf{D}(\cdot,\cdot)$ — Function defined by

$$\mathbf{D}(x,y) = \sum_{i=0}^{n} \binom{n}{i} P_d^i(1-P_d)^{n-i}[(1-y)\,\mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y)) + y\mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y))]$$

DFC — Data fusion center.

$\mathbf{F}(\cdot,\cdot)$ — Function defined by

$$\mathbf{F}(x,y) = \sum_{i=0}^{n} \binom{n}{i} P_f^i(1-P_f)^{n-i}[(1-x)\,\mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y)) + x\mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y))]$$

.

$\mathrm{int}\{\cdot\}$ — Smallest integer larger than or equal to x: $x \le \mathrm{int}\{x\} < x+1$.

LDM — Local decision maker.

n — Number of local decision makers.

P_f — False alarm probability of an LDM.

$P_{fg}^{(t)}$ — Global probability of false alarm for the detection system at time t.

P_d — Detection probability of an LDM.

$P_{dg}^{(t)}$ — Global probability of detection for the detection system at time t.

r — Average signal-energy-to-noise ratio of an LDM.

S — Closed unit square $\{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1\}$.

$\mathbf{sgn}\{\cdot\}$ — Sign function: $\mathbf{sgn}\{x\} = \begin{cases} 1, & x \ge 0 \\ -1, & x < 0. \end{cases}$.

t_s — Stopping-time of the system with DFC memory.

$\mathbf{U_{-1}}[\cdot]$ — Unit step function: $\mathbf{U_{-1}}[x] = \begin{cases} 1, & x \ge 0 \\ 0, & x < 0 \end{cases}$.

$u_g^{(t)}$ — Global decision of the detection system at time instant t.

$u_i^{(t)}$ — Decision of the ith LDM at time t.

Manuscript received July 5, 1997; revised August 7, 1998. This work was supported by the National Science Foundation under Grants ECS 9057587, ECS 9216588, and ECS 9512363, and by the Electric Power Research Institute under Grant SF 958.

The authors are with the Data Fusion Laboratory, Drexel University, Philadelphia, PA 19104 USA (e-mail: kam@lorelei.ece.drexel.edu).

Publisher Item Identifier S 1083-4427(99)00238-6.

I. INTRODUCTION

BOTH parallel and serial decentralized distributed detection systems have been studied extensively since the early 1980's for applications in multisensor decision systems [2], [5], [10], [15]–[17]. Common optimal criteria used in analyzing the overall system performance are the Bayes risk and Neyman–Pearson (N-P) criterion. Most studies involving parallel architecture assume memoryless local decision makers (LDM's) and a memoryless data fusion center (DFC). It is, however, possible for components in the system to remember and use past decisions for synthesis of new decisions with potentially significant performance advantages [1], [4], [8], [9], [11]–[13], [19]. This observation is further motivated by the structure of centralized decision makers in target tracking (e.g., [3] and [7]). In this paper, the DFC employs a 1-bit memory through which its past output is used. The structure of the system is shown in Fig. 1. The n LDM's survey a common phenomenon and use a local performance index to determine whether or not a target is in sight. At each time instant t $(t = 0, 1, \ldots)$, the ith LDM obtains new measurements and makes a binary decision $u_i^{(t)}$, where $u_i^{(t)} = -1$ means that hypothesis H_0 (target-absent) is accepted and $u_i^{(t)} = 1$ means that hypothesis H_1 (target-present) is accepted. The local decisions are transmitted over noiseless communication channels to the DFC, and combined with the DFC's most recent global decision $u_g^{(t-1)}$ to form a new global binary decision $u_g^{(t)}$ at time t. Again, $u_g^{(t)} = -1$ means "accept H_0" and $u_g^{(t)} = 1$ means "accept H_1". The global decision is made optimally by minimizing a Bayes risk of the form

$$\beta\left(P_{fg}^{(t)}, P_{dg}^{(t)}\right) = C_{00}P_0\left(1 - P_{fg}^{(t)}\right) + C_{10}P_0P_{fg}^{(t)} + C_{01}P_1\left(1 - P_{dg}^{(t)}\right) + C_{11}P_1P_{dg}^{(t)} \tag{1}$$

where C_{ij} is the cost of accepting H_i when H_j is true $(i, j = 0, 1)$, P_j is the a priori probability of hypothesis H_j $(j = 0, 1)$, and $P_{fg}^{(t)} = P_r\{u_g^{(t)} = 1 \mid H_0\}$ and $P_{dg}^{(t)} = P_r$ $\{u_g^{(t)} = 1 \mid H_1\}$ are the global probabilities of false alarm and detection, respectively, at time instant t. In the sequel we assume that $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$, so that the Bayes risk is the global probability of error (GPE). As we shall show, the performance of the 1-bit memory DFC improves uniformly until the DFC converges to a decision and needs no more samples from the (stationary) LDM's[1].

The use of local objective functions and an independent global performance index is known to result in suboptimal performance when compared to designs which simultaneously design the local and global decision rules. In some cases, the imposition of global objective functions on a set of local detectors with specific structure may even result in counterproductive decisions (see [15, Example 1, pp. 506–507]). However, practical considerations (such as local detector modules that cannot be modified) and considerations of storage and compatibility often make the use of independent local detectors mandatory. For a discussion of the resulting tradeoffs, see [8] and [9].

To analyze the performance of this distributed detection system with memory, we make the following assumptions.

1) Every decision made by a LDM is conditionally statistically independent of all other decisions under hypothesis H_j $(j = 0, 1)$.
2) All LDM's are fixed with the same known time-invariant probabilities of false alarm $P_{fi} = P_f = P_r\{u_i^{(t)} = 1 \mid H_0\}$ and detection $P_{di} = P_d = P_r\{u_i^{(t)} = 1 \mid H_1\}$ for $i = 1, 2, \ldots, n$. We also assume that $0 < P_f < P_d < 1$.

[1]The problem described by Fig. 1 can be cast as a serial detection problem. In a binary serial-detection architecture each decision maker (say, the ith) obtains a set of observations y_i from the "environment," and a binary decision u from the previous decision maker, the $(i-1)$th. On this basis, it creates its own decision, $u_i^{(k)}$. One can assume that at time t we have built a serial architecture of t decision makers in tandem; denoting them $k = 1, 2, \ldots, t$, the kth decision maker obtains $y_k = \{u_i^{(k)}\}_{i=1}^n$ as inputs. Here, $u_i^{(k)}$ is the output at time k of an auxiliary binary local decision maker which makes decisions about the hypothesis. In this formulation, results from [8], [9], [12], [13], and [19] become applicable.

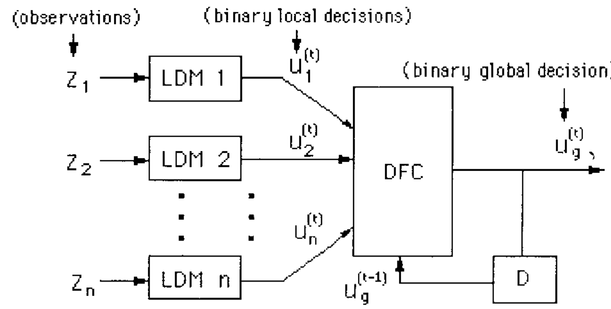

Fig. 1. The structure of the distributed detection system with feedback at the DFC.

Given $P_f, P_d, \{u_i^{(t)}\}_{i=1}^n$ and $u_g^{(t-1)}$ where $\{u_i^{(t)}\}_{i=1}^n$ and $u_g^{(t-1)}$ are conditionally statistically independent, we can obtain $u_g^{(t)}$ the global optimal decision at time t by using a generalization of the well-known Chair–Varshney fusion rule for distributed detection architectures with fixed LDM's ([2]). Our generalized fusion rule is thus expressed as

$$u_g^{(t)} = \mathbf{sgn}\left\{a\sum_{i=1}^{n} u_i^{(t)} + a_g^{(t-1)}u_g^{(t-1)} - \left(2\log\tau + b + b_g^{(t-1)}\right)\right\} \tag{2}$$

where $\mathbf{sgn}\{\cdot\}$ is the algebraic sign function: $\mathbf{sgn}\{x\} = \begin{cases} 1, & x \geq 0 \\ -1 & x < 0 \end{cases}$

$$a = \log\left[\frac{P_d(1-P_f)}{P_f(1-P_d)}\right],$$

$$b = n\log\left[\frac{P_f(1-P_f)}{P_d(1-P_d)}\right],$$

$$a_g^{(t-1)} = \log\left[\frac{P_{dg}^{(t-1)}\left(1-P_{fg}^{(t-1)}\right)}{P_{fg}^{(t-1)}\left(1-P_{dg}^{(t-1)}\right)}\right],$$

$$b_g^{(t-1)} = \log\left[\frac{P_{fg}^{(t-1)}\left(1-P_{fg}^{(t-1)}\right)}{P_{dg}^{(t-1)}\left(1-P_{dg}^{(t-1)}\right)}\right], \quad \text{and}$$

$\tau = P_0/P_1$ (the decision threshold of the global detector).

Equation (2) is obtained from the data fusion rule in [2, Eq. (12)]. At time t, the data fusion center's decision at time $(t-1)$ is used as the output of $(n+1)$st sensor.

In Section II, we prove that when the system operates in a stationary environment, the DFC converges to a decision with probability one. Convergence is shown to be equivalent to having the global probability of false alarm and global probability of detection enter a certain polygon (the "stopping set") in the P_{fg}-P_{dg} plane. Once the probabilities $(P_{fg}^{(t)}, P_{dg}^{(t)})$ have entered this region, the DFC "freezes" and no further observations need be collected.

Knowing the boundaries of the stopping set allows the calculation of an upper bound on the global Bayes risk [(20)] that the system will achieve at convergence. In Section III, the global performance of this distributed detection system with memory is exemplified using a specific set of LDM's. We study the detection of signals with random phase and amplitude in Gaussian noise [20], and obtain a convenient upper bound on the steady-state probability of error versus the signal-to-noise ratio. The performance of the 1-bit memory DFC system is significantly superior to that of a system with a memoryless DFC.

II. Performance Characteristics of a 1-Bit Memory Data-Fusion Rule

Using the optimal fusion rule (2), the global probability of false alarm $P_{fg}^{(t)}$ and the global probability of detection $P_{dg}^{(t)}$ at time $t = 1, 2, \ldots$ can be expressed in terms of their initial values $(P_{fg}^{(0)}, P_{dg}^{(0)})$ through the following iterative scheme:

$$P_{fg}^{(t)} = \mathbf{F}\big(P_{fg}^{(t-1)}, P_{dg}^{(t-1)}\big) \tag{3a}$$

$$P_{dg}^{(t)} = \mathbf{D}\big(P_{fg}^{(t-1)}, P_{dg}^{(t-1)}\big) \tag{3b}$$

where

$$\mathbf{F}(x,y) = \sum_{i=0}^{n} \binom{n}{i} P_f^i (1-P_f)^{n-i} [(1-x)\mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y)) + x\mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y))] \tag{4a}$$

$$\mathbf{D}(x,y) = \sum_{i=0}^{n} \binom{n}{i} P_d^i (1-P_d)^{n-i} [(1-y)\mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y)) + y\mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y))] \tag{4b}$$

and

$$\gamma_{\mathbf{i}}(x,y) = P_d^i (1-P_d)^{n-i}(1-y) - \tau P_f^i (1-P_f)^{n-i}(1-x) \tag{5a}$$

$$\delta_{\mathbf{i}}(x,y) = P_d^i (1-P_d)^{n-i} y - \tau P_f^i (1-P_f)^{n-i} x. \tag{5b}$$

Here, $\mathbf{U_{-1}}(\cdot)$ is the unit step function defined by $\mathbf{U_{-1}}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$. These equations are based on the definitions

$$\begin{aligned} P_{fg}^{(t)} &= P_r\{u_g^{(t)} = 1 \mid H_0\} \\ &= \mathbf{E}\big[P_r\{u_g^{(t)} = 1 \mid H_0, u_1^{(t)}, u_2^{(t)}, \ldots, u_n^{(t)}, u_g^{(t-1)}\}\big] \end{aligned} \tag{6a}$$

$$\begin{aligned} P_{dg}^{(t)} &= P_r\{u_g^{(t)} = 1 \mid H_1\} \\ &= \mathbf{E}\big[P_r\{u_g^{(t)} = 1 \mid H_1, u_1^{(t)}, u_2^{(t)}, \ldots, u_n^{(t)}, u_g^{(t-1)}\}\big] \end{aligned} \tag{6b}$$

where $\mathbf{E}[\cdot]$ is an expectation function evaluated over all local decisions and the previous global decision.

The geometric interpretation of the above iterative scheme is as follows: let S be the closed unit square S in the xy-plane

$$S = \{(x,y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}. \tag{7}$$

Then the mapping $(x,y) \to (\mathbf{F}(x,y), \mathbf{D}(x,y))$ maps S into itself (Lemma 1 in the appendix). For each $i = 0, 1, \ldots, n$, the equation $\gamma_{\mathbf{i}}(x,y) = 0$ is the line $(1-y) = s_i(1-x)$ through the point $(1,1)$ with slope s_i given by

$$s_i = \tau \left(\frac{P_f}{P_d}\right)^i \left(\frac{1-P_f}{1-P_d}\right)^{n-i}. \tag{8}$$

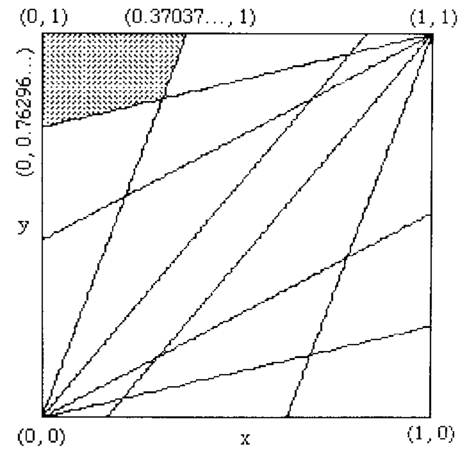

Fig. 2. The unit square S and its partition into polygons by the lines $\gamma_{\mathbf{i}}(x,y) = 0$ and $\delta_{\mathbf{i}}(x,y) = 0$ for $i = 0, 1, \ldots, n$. Here $n = 3$, $P_f = 0.4$, $P_d = 0.6$, and $\tau = 0.8$. The shaded region is the stopping set $\Re_S$ defined in (10).

Because of our assumption that $0 < P_f < P_d < 1$, we have

$$0 < s_n < s_{n-1} < \cdots < s_1 < s_0 < \infty. \tag{9}$$

Similarly, the equation $\delta_{\mathbf{i}}(x,y) = 0$ is the line $y = s_i x$ through the point $(0,0)$. These $n+1$ pairs of parallel lines partition the unit square into three- or four-sided convex polygons (Fig. 2). In each polygon, $\mathbf{F}(x,y)$ is a linear function of x alone and $\mathbf{D}(x,y)$ is a linear function of y alone. Both functions change discontinuously as (x,y) moves from one polygon to another.

Among the polygons that partition S, the following is of particular interest:

$$\Re_S = \{(x,y) \in S \mid \gamma_{\mathbf{n}}(x,y) < 0,\ \delta_{\mathbf{0}}(x,y) \geq 0\}. \tag{10}$$

This is the polygon that always contains the point $(0,1)$ (the shaded polygon in Fig. 2). In this polygon we have $\mathbf{F}(x,y) = x$ and $\mathbf{D}(x,y) = y$ (Lemma 2 in the appendix). We shall consequently call this polygon the *stopping set* because if an iterate $(P_{fg}^{(s)}, P_{dg}^{(s)})$ ever enters it, then $(P_{fg}^{(t)}, P_{dg}^{(t)}) = (P_{fg}^{(s)}, P_{dg}^{(s)})$ for all $t \geq s$, and so the iterates stop changing. In addition, as shown in Lemma 3 in the Appendix, the decisions of the DFC do not change from that time on; that is, $u_g^{(t)} = u_g^{(s)}$ for all $t \geq s$, regardless of the values of the local decisions $u_i^{(t)}$ for $i = 1, 2, \ldots, n$ and $t \geq s$.

Next, suppose we choose the initial operating point $(P_{fg}^{(0)}, P_{dg}^{(0)})$ as $(\frac{1}{2}, \frac{1}{2})$. Then from (2) we have $a_g^{(0)} = b_g^{(0)} = 0$ and so we need not know $u_g^{(0)}$ to compute $u_g^{(1)}$. In fact, the first iterate $(P_{fg}^{(1)}, P_{dg}^{(1)})$ coincides with the unchanging operating point of a memoryless system and the first global decision $u_g^{(1)}$ is the one of a memoryless system (see [3, Eq. (6)]). We shall denote this particular point by $(P_{fg}^{(M)}, P_{dg}^{(M)})$. (In fact, all points in the polygon containing $(\frac{1}{2}, \frac{1}{2})$ are mapped

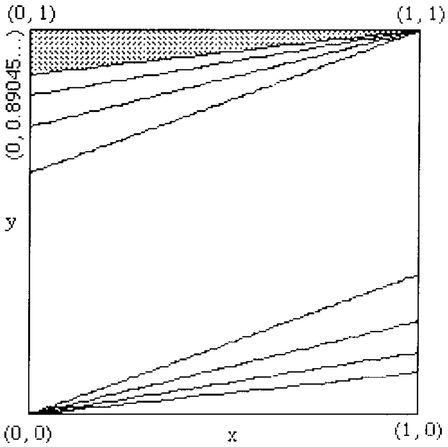

Fig. 3. The unit square S and its partition into polygons for a typical constant target-present decision case. Here $n = 3$, $P_f = 0.45$, $P_d = 0.55$, and $\tau = 0.2$. The shaded region is the stopping set $\Re_S$.

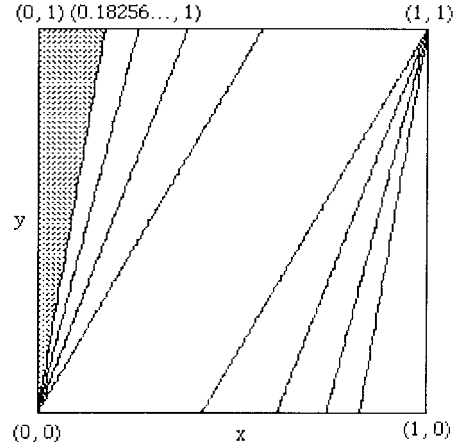

Fig. 4. The unit square S and its partition into polygons for a typical constant target-absent decision case. Here $n = 3$, $P_f = 0.45$, $P_d = 0.55$, and $\tau = 3$. The shaded region is the stopping set $\Re_S$.

onto $(P_{fg}^{(M)}, P_{dg}^{(M)})$.) This memoryless operating point is the reasonable initial point to begin with, and so for all the examples we shall assume that $(P_{fg}^{(0)}, P_{dg}^{(0)}) = (\frac{1}{2}, \frac{1}{2})$.

For certain extreme values of the system parameters, $(P_{fg}^{(M)}, P_{dg}^{(M)})$ and the stopping set are as follows.

Constant target-present decision case: If $\tau \le (\frac{1-P_d}{1-P_f})^n$, then from (8) and (9) we have $0 < s_n < \cdots < s_0 \le 1$. The stopping set reduces to a triangular region as shown in Fig. 3. In addition, we have $\gamma_{\mathbf{i}}(\frac{1}{2}, \frac{1}{2}) \ge 0$ and $\delta_{\mathbf{i}}(\frac{1}{2}, \frac{1}{2}) \ge 0$ for all $i = 0, 1, \ldots, n$, and so from (4) we have $(P_{fg}^{(M)}, P_{dg}^{(M)}) = (1,1)$. Although $(1,1) \notin \Re_S$, we nevertheless have $\mathbf{F}(1,1) = 1$ and $\mathbf{D}(1,1) = 1$ from (4). Thus, $(P_{fg}^{(t)}, P_{dg}^{(t)}) = (1,1)$ for all t. This means that fusion rule (2) always returns a target-present decision ($u_g^{(t)} = 1$ for $t = 1, 2, \ldots$) regardless of the inputs from the local detectors.

Constant target-absent decision case: If $\tau > (\frac{P_d}{P_f})^n$, then from (8) and (9) we have $1 < s_n < \cdots < s_0 < \infty$. The stopping set reduces to a triangular region as shown in Fig. 4. In addition, we have $\gamma_{\mathbf{i}}(\frac{1}{2}, \frac{1}{2}) < 0$ and $\delta_{\mathbf{i}}(\frac{1}{2}, \frac{1}{2}) < 0$ for all $i = 0, 1, \ldots, n$, and so from (4) we have $(P_{fg}^{(M)}, P_{dg}^{(M)}) = (0,0)$. Because $(0,0) \in \Re_S$ in this case, it follows that $(P_{fg}^{(t)}, P_{dg}^{(t)}) = (0,0)$ for all t. This means that fusion rule (2) always returns a target-absent decision ($u_g^{(t)} = -1$ for $t = 1, 2, \ldots$) regardless of the inputs from the local detectors.

For both of the above cases, our 1-bit memory accomplishes nothing. We shall consequently make the following assumption throughout this paper to exclude these extreme cases (we also collect some of our previous assumptions as items 1–3).

Assumption 1: The system parameters P_f, P_d, τ, and n satisfy

$$\begin{aligned} &1)\ n \ge 1; \\ &2)\ \tau > 0; \\ &3)\ 0 < P_f < P_d < 1; \\ &4)\ \left(\frac{1-P_d}{1-P_f}\right)^n < \tau \le \left(\frac{P_d}{P_f}\right)^n. \end{aligned} \tag{11}$$

Under Assumption 1, there must exist an integer k ($1 \le k \le n$) such that $0 < s_n \cdots < s_k \le 1 < s_{k-1} < \cdots < s_0 < \infty$. From (4), the memoryless initial point is then given by

$$P_{fg}^{(M)} = \sum_{i=k}^{n} \binom{n}{i} P_f^i (1-P_f)^{n-i} \tag{12a}$$

$$P_{dg}^{(M)} = \sum_{i=k}^{n} \binom{n}{i} P_d^i (1-P_d)^{n-i}. \tag{12b}$$

As is shown in Lemma 4 in the appendix, this point lies in the interior of S and $(P_{fg}^{(M)}, P_{dg}^{(M)}) \notin \Re_S$.

The stopping set under Assumption 1 is a 4-sided polygon, as illustrated in Fig. 2, with a vertex lying in the interior of S, which we shall denote by (P_{fg}^*, P_{dg}^*). This vertex is at the intersection of the lines $(1-y) = s_n(1-x)$ and $y = s_0 x$ and has coordinates

$$P_{fg}^* = \frac{P_d^n(1-P_d)^n - \tau P_f^n (1-P_d)^n}{\tau\left[P_d^n(1-P_f)^n - P_f^n(1-P_d)^n\right]} \tag{13a}$$

$$P_{dg}^* = \frac{P_d^n(1-P_f)^n - \tau P_f^n (1-P_f)^n}{P_d^n(1-P_f)^n - P_f^n(1-P_d)^n}. \tag{13b}$$

Notice that (P_{fg}^*, P_{dg}^*) does not belong to $\Re_S$.

We next return to the global probability of error (GPE), which we had previously defined at the point (x,y) as

$$\beta(x,y) = P_0 x + P_1(1-y) = \frac{\tau x + (1-y)}{1+\tau}. \tag{14}$$

We define the change in the GPE from one iteration to the next as

$$\Delta\beta(x,y) = \beta(\mathbf{F}(x,y), \mathbf{D}(x,y)) - \beta(x,y). \tag{15}$$

We obviously have that $\Delta\beta(x,y) = 0$ if $(x,y) \in \Re_S$, since $(\mathbf{F}(x,y), \mathbf{D}(x,y)) = (x,y)$ in $\Re_S$. In addition, from Lemma 6 in the Appendix, we have that $\Delta\beta(x,y) < 0$ outside of the closure of $\Re_S$. Consequently, if an iterate $(P_{fg}^{(t)}, P_{dg}^{(t)})$ lies outside of the closure of $\Re_S$, then the GPE of the next iterate, $(P_{fg}^{(t+1)}, P_{dg}^{(t+1)})$, is strictly less than the GPE of $(P_{fg}^{(t)}, P_{dg}^{(t)})$. The GPE can thus serve as a Lyapunov function for the dynamical system from which we can prove the following theorem (see the Appendix for the proof):

Theorem 1: Under Assumption 1, for any given initial point $(P_{fg}^{(0)}, P_{dg}^{(0)})$ in the unit square, the iterates $(P_{fg}^{(t)}, P_{dg}^{(t)})$, $t = 1, 2, \ldots$, behave according to one of the following three cases.

1) (The *stopping-time* case): There exists an integer t_S (the *stopping time*) such that $(P_{fg}^{(t)}, P_{dg}^{(t)}) \in \Re_S$ for all $t \geq t_S$.

2) (The *fixed-point* case): $(P_{fg}^{(t)}, P_{dg}^{(t)}) \notin \Re_S$ for any t and $(P_{fg}^{(t)}, P_{dg}^{(t)}) \to (P_{fg}^*, P_{dg}^*)$ as $t \to \infty$. This case arises if and only if $n > 1$ and

$$\tau = \left[\frac{P_d(1-P_d)}{P_f(1-P_f)}\right]^n \left[\frac{P_f^n + (1-P_f)^n}{P_d^n + (1-P_d)^n}\right]. \tag{16}$$

3) (The *2-cycle* case): $(P_{fg}^{(t)}, P_{dg}^{(t)}) \notin \Re_S$ for any t and the iterates $(P_{fg}^{(t)}, P_{dg}^{(t)})$ alternately approach the two distinct points

$$\left(\frac{P_f^n[1-(1-P_f)^n]}{P_f^n + (1-P_f)^n - P_f^n(1-P_f)^n}, \frac{P_d^n[1-(1-P_d)^n]}{P_d^n + (1-P_d)^n - P_d^n(1-P_d)^n}\right) \tag{17}$$

and

$$\left(\frac{P_f^n}{P_f^n + (1-P_f)^n - P_f^n(1-P_f)^n}, \frac{P_d^n}{P_d^n + (1-P_d)^n - P_d^n(1-P_d)^n}\right) \tag{18}$$

on the boundary of $\Re_S$ as $t \to \infty$. This case arises if and only if

$$\tau = \left[\frac{P_d(1-P_d)}{P_f(1-P_f)}\right]^n \left[\frac{P_f^n + (1-P_f)^n - P_f^n(1-P_f)^n}{P_d^n + (1-P_d)^n - P_d^n(1-P_d)^n}\right]. \tag{19}$$

Case 1 in the above theorem is the generic case in that cases 2 and 3 require that the system parameters P_f, P_d, τ, and n satisfy (16) and/or (19). In addition, the limit points in cases 2 and 3 lie on the boundary of the stopping set $\Re_S$. Consequently, once the iterates have gotten sufficiently close to the limit points, then numerical error will perturb a future iterate into the stopping set with probability one. All of this is to say that cases 2 and 3 are mainly of theoretical interest and that, in practice, the iterates will enter the stopping set in finite time with probability one.

As was previously mentioned, the limiting point (P_{fg}^*, P_{dg}^*) in case 2 does not belong to the stopping set. It can also be verified that it is not a fixed point of the mapping. Rather, it can be described as a *potential* fixed point that certain sequences of iterates approach, but never reach. The same holds true for the 2-cycle case. As is shown in the proof of Theorem 1, one of the two limiting points in case 3 [the one in (18)] belongs to $\Re_S$ and the other does not, and both points have the same GPE. The two limiting points do not constitute a true 2-cycle of the mapping, but rather a potential 2-cycle. This situation can arise only because the mapping is not continuous.

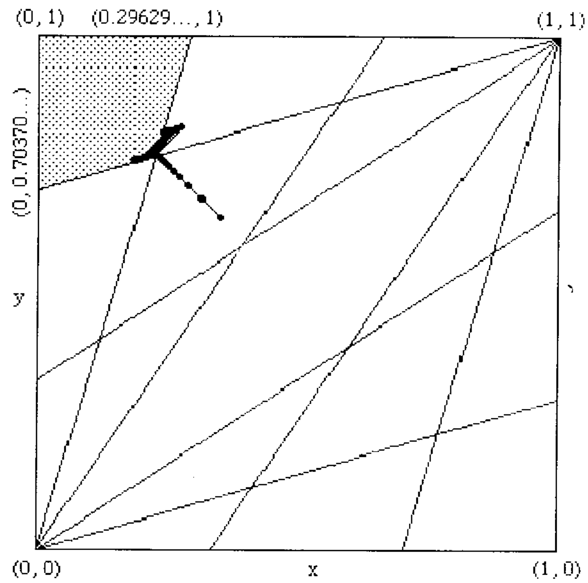

Fig. 5. The unit square S and its partition into polygons for which cases 2 and 3 of Theorem 1 both arise. Here $n = 3$, $P_f = 0.4$, $P_d = 0.6$, and $\tau = 1$. The shaded region is the stopping set $\Re_S$. The various iterates are shown beginning with the initial point $(P_{fg}^{(M)}, P_{dg}^{(M)}) = (44/125, 81/125) = (0.352, 0.648)$.

Notice that if $P_f + P_d = 1$, then the exceptional value of τ for both the fixed-point case and the 2-cycle case is 1. Conversely, the fixed-point and 2-cycle cases can coexist only if $P_f + P_d = 1$ and $\tau = 1$. All three cases in Theorem 1 will then occur for various choices of the initial point $(P_{fg}^{(0)}, P_{dg}^{(0)})$. Each of the three cases has its own basin (the set of $(P_{fg}^{(0)}, P_{dg}^{(0)})$ leading to that case) and the three basins meet at the vertex (P_{fg}^*, P_{dg}^*) of the stopping set.

The slope τ of the level curves of the linear function $\beta(x, y)$ lies between the slopes of the two lines whose intersection determines (P_{fg}^*, P_{dg}^*); that is, $s_n < \tau < s_0$ [see (8)]. By familiar linear-programming arguments, it follows that $\beta(x, y)$ is maximized within the closure of the convex polygon $\Re_S$ at the vertex (P_{fg}^*, P_{dg}^*). This maximum value is given by

$$\beta(P_{fg}^*, P_{dg}^*) = \frac{P_d^n(1-P_d)^n - (1+\tau)P_f^n(1-P_d)^n + \tau P_f^n(1-P_f)^n}{(1+\tau)\left[P_d^n(1-P_f)^n - P_f^n(1-P_d)^n\right]}. \tag{20}$$

This value serves as a convenient upper bound for the limiting value of the GPE of the iterates as $t \to \infty$. In other words, for any initial operating point we have $\lim_{t\to\infty} \beta(P_{fg}^{(t)}, P_{dg}^{(t)}) \leq \beta(P_{fg}^*, P_{dg}^*)$ since $(P_{fg}^{(t)}, P_{dg}^{(t)})$ either enters $\Re_S$ in finite time or asymptotically approaches its boundary as $t \to \infty$.

Fig. 5 illustrates some of the points discussed above. For a system with $n = 3$, $P_f = 0.4$, $P_d = 0.6$, and $\tau = 1$, Theorem 1 guarantees that all three cases will coexist since (16) and (19) are both satisfied. We begin with the initial point $(P_{fg}^{(M)}, P_{dg}^{(M)}) = (44/125, 81/125) = (0.352, 0.648)$, for which GPE $= 44/125 = 0.352$. This initial point lies in the basin of the potential fixed point

$(P^*_{fg}, P^*_{dg}) = (8/35, 27/35) = (0.228\,571\ldots, 0.771\,428\ldots)$, for which GPE $= 8/35 = 0.228\,571\ldots$. The iterates $(P^{(t)}_{fg}, P^{(t)}_{dg})$ first begin to converge to that point as t increases. However, because of truncation and round-off error in our computing machine, the 77th iterate is perturbed into the basin of the potential 2-cycle consisting of the two points $(784/4159, 3159/4159) = (0.188\,506\ldots, 0.759\,557\ldots)$ and $(1000/4159, 3375/4159) = (0.240\,442\ldots, 0.811\,493\ldots)$, for which GPE $= 892/4159 = 0.214\,474\ldots$. The iterates then begin to converge to this potential 2-cycle. But the 227th iterate is perturbed to a point in the fixed set near the second of the two listed points of the potential 2-cycle. Thereafter, the iterates are frozen at this point, $(0.240\,442\ldots, 0.811\,493\ldots)$, for which the GPE is $0.214\,474\ldots$. The global probability of error has thus decreased from 0.352 to $0.214\,474\ldots$, although with exact arithmetic it should have decreased only to $0.228\,571\ldots$.

III. Example

To exemplify the performance of the 1-bit DFC memory distributed detection system, we study the performance of a 4-sensor system characterized by the following specific LDM receiver operating characteristic (ROC):

$$P_d = P_f^{1/(1+\tau)}. \tag{21}$$

This kind of ROC occurs in many detection problems; e.g., in detecting signals with random phase and amplitude in Gaussian noise (where r is the average signal-energy-to-noise ratio) [20] and in classifying Gaussian signals with equal means and different variances (σ_0^2 under H_0, σ_1^2 under $H_1, \sigma_1^2 > \sigma_0^2$, $r = \frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2}$) ([18, p. 41]).

For this 4-sensor distributed detection system, we analyze and compare the performance—*the global probability of error*—achieved by a DFC with 1-bit memory and a DFC without memory.

For a DFC with 1-bit memory, the steady-state detection performance is shown in the sequel to be closely approximated by the upper bound given in (20). For a DFC without memory, the detection performance is characterized by

$$\beta\big(P_{fg}^{(M)}, P_{dg}^{(M)}\big) = P_0 P_{fg}^{(M)} + P_1\big(1 - P_{dg}^{(M)}\big) \tag{22}$$

where [cf. (11) and (12)]

$$P_{fg}^{(M)} = \sum_{i=k}^{4} \binom{4}{i} P_f^i (1-P_f)^{4-i} \quad \text{and}$$
$$P_{dg}^{(M)} = \sum_{i=k}^{4} \binom{4}{i} P_d^i (1-P_d)^{4-i} \tag{23}$$

with

$$k = \mathrm{int}\left\{ \frac{\ln\tau + 4\ln[(1-P_f)/(1-P_d)]}{\ln(P_d/P_f) + \ln[(1-P_f)/(1-P_d)]} \right\} \tag{24}$$

and $\mathrm{int}\{x\}$ is the smallest integer larger than or equal to x: $x \le \mathrm{int}\{x\} < x+1$.

In Fig. 6, the global probability of error is plotted versus the signal-to-noise ratio (SNR, r) of the local detectors for the actual global performance of a 1-bit memory DFC, that of its upper bound, and that of a memoryless DFC, with parameters $\tau = 0.5$, $n = 4$, $P_f = 0.2$, and $P_d = P_f^{1/(1+r)}$. It is seen that the performance of the upper bound for the 1-bit memory DFC is generally much better than that of a memoryless DFC. The performance of the 1-bit memory DFC can be tightly approximated by its upper bound for low and moderate SNR's (SNR $\le$ 30 dB). For SNR $\le -7.1122\ldots$ dB, the 1-bit memory DFC system performs the same as a memoryless DFC because both give a constant "target-present" decision regardless of local decisions. The discontinuous nature of the performance of the 1-bit memory DFC as r varies is a result of the discontinuities in the functions $\mathbf{F}$ and $\mathbf{D}$ [cf. (4)].

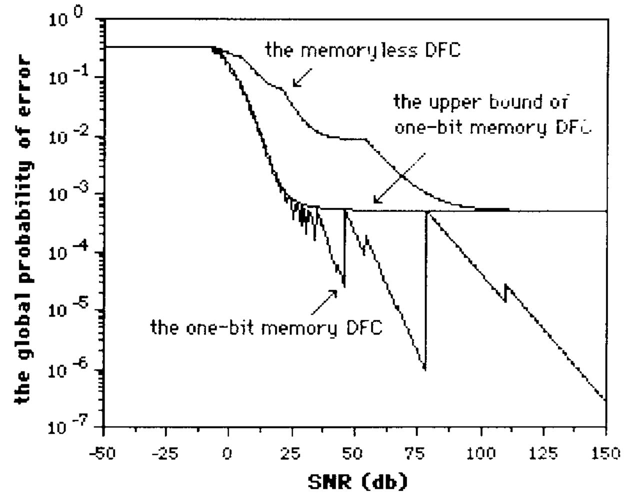

Fig. 6. The global probability of error versus the signal-to-noise ratio (r) for fixed local detectors with $P_f = 0.2$, $P_d = P_f^{1/(1+r)}$, $\tau = 0.5$, and $n = 4$.

TABLE I
THE PERFORMANCE OF 1-BIT MEMORY DFC FOR $n = 4$, $\tau = 0.5$ AND $r = 1$

Operating Point (P_f, P_d)	Number of Iterations	Memoryless DFC Prob. of Error	Upper Bound Eqs. (20)	1-bit memory DFC Prob. of Error
(0.1, 0.316228)	153	0.260365	0.231148	0.168530
(0.2, 0.447214)	51	0.259050	0.157629	0.131410
(0.3, 0.547723)	22	0.279122	0.132576	0.131016
(0.4, 0.632456)	25	0.270838	0.127551	0.126628
(0.5, 0.774597)	45	0.281451	0.136411	0.134241

In Table I, the performance of 1-bit memory DFC, along with the number of iterations required to reach the stopping set, is compared to that of its upper bound and to that of a memoryless DFC with $n = 4$, $\tau = 0.5$, $r = 1$, and five different set values for (P_f, P_d). Again, we observe that the performance of the upper bound for the 1-bit memory DFC approximates its actual performance and is superior than that of a memoryless DFC.

Another example is shown in Fig. 7 where the receiver-operating-characteristic (ROC: the global probability of detection versus the global probability of false alarm) is plotted as the threshold τ varies using $n = 4$, $P_f = 0.2$, and $r = 1$

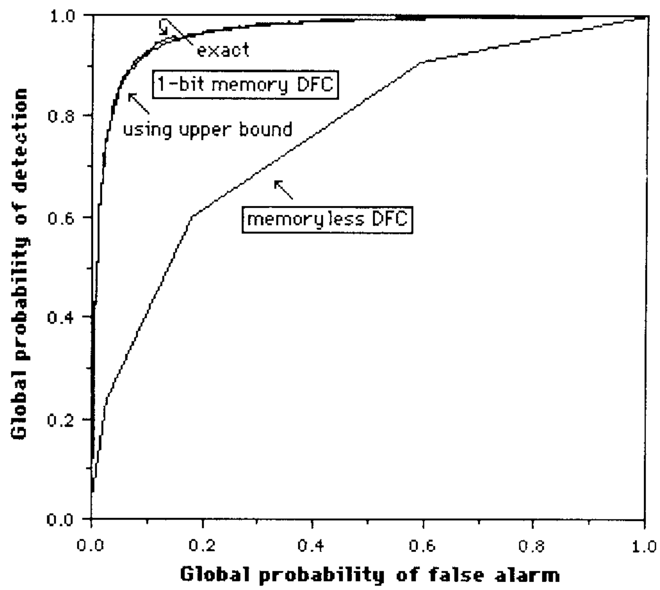

Fig. 7. The ROC of the 1-bit memory DFC, its upper bound and the memoryless DFC with the fixed local detector $P_f = 0.2$, $P_d = 0.447\,214$ for $n = 4$.

(i.e., $P_d = P_f^{1/2} = 0.447\,214$): 1) for the 1-bit memory DFC after the iterates have entered the stopping set; 2) for the 1-bit memory DFC using the upper bound (20); and 3) for a memoryless DFC. The ROC that uses the upper bound (20) tightly approximates the true global performance of the 1-bit memory DFC and is significantly better than the ROC of a memoryless DFC.

IV. Conclusion

The effect of a 1-bit DFC memory on the performance of a distributed detection parallel architecture was quantified and demonstrated. Specifically, we studied the use of the most recent decision in calculating the DFC's final decision. We showed that in stationary environments the 1-bit memory DFC will converge to a decision with probability one in finite time, and that the performance of this decision is quantified by way of a tight upper bound. The 1-bit DFC memory system is shown with specific examples to outperform that of a system without memory and its performance can be accurately predicted. The price for the improved performance is the memory hardware cost. Further research is needed to study the functional properties of the resulting discrete-time system with the aim to evaluate the performance of the 1-bit DFC memory system in nonstationary environments, and to assess the effect of DFC-LDM feedback on detection performance. Multibit memory is another interesting area of exploration, though it is clear that the correlation between past decisions would significantly complicate the calculations of [5].

Appendix

Lemma 1: If $(x,y) \in \Re_S$, then $(\mathbf{F}(x,y), \mathbf{D}(x,y)) \in \Re_S$.

Proof: From (4a), for each fixed y we have that $\mathbf{F}(x,y)$ is a linear function of x with

$$\mathbf{F}(0,y) = \sum_{i=0}^{n} \binom{n}{i} P_f^i (1-P_f)^{n-i} \mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y)) \qquad \text{(A1)}$$

and

$$\mathbf{F}(1,y) = \sum_{i=0}^{n} \binom{n}{i} P_f^i (1-P_f)^{n-i} \mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y)). \qquad \text{(A2)}$$

Because $\sum_{i=0}^{n}\binom{n}{i}P_f^i(1-P_f)^{n-i} = 1$, it is clear that both $\mathbf{F}(0,y)$ and $\mathbf{F}(1,y)$ are in $[0,1]$, and so $\mathbf{F}(x,y) \in [0,1]$ for all $x \in [0,1]$. A similar argument for each fixed x shows that $\mathbf{D}(x,y) \in [0,1]$ for all $y \in [0,1]$. □

Lemma 2: If $(x,y) \in \Re_S$, then $(\mathbf{F}(x,y), \mathbf{D}(x,y)) = (x,y)$.

Proof: If $(x,y) \in \Re_S$, then from (5), (8), and (9), we have $\gamma_{\mathbf{i}}(x,y) < 0$ and $\delta_{\mathbf{i}}(x,y) \geq 0$ for all $i = 0, 1, \ldots, n$. Thus, (4) reduces to $\mathbf{F}(x,y) = x$ and $\mathbf{D}(x,y) = y$. □

Lemma 3: An iterate $(P_{fg}^{(s)}, P_{dg}^{(s)})$ lies in $\Re_S$ if and only if $u_g^{(s+1)} = u_g^{(s)}$ for any possible values of local decisions $u_i^{(s+1)}$, $i = 1, 2, \ldots, n$.

Proof: We first show that if $(P_{fg}^{(s)}, P_{dg}^{(s)}) \in \Re_S$ then $u_g^{(s+1)} = u_g^{(s)}$ for any possible values of local decisions $u_i^{(s+1)}$, $i = 1, 2, \ldots, n$. Two cases arise according to the two possible values of the global decision $u_g^{(s)}$. In the case when $u_g^{(s)} = 1$, fusion rule (2) reduces to

$$u_g^{(s+1)} = \mathbf{sgn}\left\{ \frac{P_{dg}^{(s)}}{P_{fg}^{(s)}} - \tau \left(\frac{P_f}{P_d}\right)^{\left(\frac{n+m}{2}\right)} \left(\frac{1-P_f}{1-P_d}\right)^{\left(\frac{n-m}{2}\right)} \right\} \qquad \text{(A3)}$$

where $m = \sum_{i=1}^{n} u_i^{(s+1)}$. For any possible values of $u_i^{(s+1)}$, $i = 1, 2, \ldots, n$, we have $-n \leq m \leq n$. From the fact that $0 < P_f < P_d < 1$, we may derive the inequality

$$\left(\frac{1-P_f}{1-P_d}\right)^n \geq \left(\frac{P_f}{P_d}\right)^{\left(\frac{n+m}{2}\right)} \left(\frac{1-P_f}{1-P_d}\right)^{\left(\frac{n-m}{2}\right)} \qquad \text{(A4)}$$

for $-n \leq m \leq n$.

Now because $(P_{fg}^{(s)}, P_{dg}^{(s)}) \in \Re_S$, we have from (10) that $\delta_0(P_{fg}^{(s)}, P_{dg}^{(s)}) \geq 0$, and so from (5b)

$$\frac{P_{dg}^{(s)}}{P_{fg}^{(s)}} \geq \tau \left(\frac{1-P_f}{1-P_d}\right)^n. \qquad \text{(A5)}$$

Inequalities (A4) and (A5) combine to show that the argument of the $\mathbf{sgn}$ function in (A3) is nonnegative for $-n \leq m \leq n$. Consequently, $u_g^{(s+1)} = 1$ and so the desired conclusion $u_g^{(s+1)} = u_g^{(s)}$ is true.

Next, consider the case when $u_g^{(s)} = -1$. Then fusion rule (2) can be reduced to

$$u_g^{(s+1)} = \mathbf{sgn}\left\{ \frac{1-P_{dg}^{(s)}}{1-P_{fg}^{(s)}} - \tau \left(\frac{P_f}{P_d}\right)^{\left(\frac{n+m}{2}\right)} \left(\frac{1-P_f}{1-P_d}\right)^{\left(\frac{n-m}{2}\right)} \right\}. \qquad \text{(A6)}$$

Analogous to (A4), from $0 < P_f < P_d < 1$, we may derive the inequality

$$\left(\frac{P_f}{P_d}\right)^n \le \left(\frac{P_f}{P_d}\right)^{(\frac{n+m}{2})}\left(\frac{1-P_f}{1-P_d}\right)^{(\frac{n-m}{2})} \tag{A7}$$

for $-n \le m \le n$.

Now because $(P_{fg}^{(s)}, P_{dg}^{(s)}) \in \Re_S$, we have from (10) that $\gamma_n(P_{fg}^{(s)}, P_{dg}^{(s)}) < 0$, and so from (5a)

$$\frac{1-P_{dg}^{(s)}}{1-P_{fg}^{(s)}} < \tau\left(\frac{P_f}{P_d}\right)^n. \tag{A8}$$

Inequalities (A7) and (A8) combine to show that the argument of the $\mathbf{sgn}$ function in (A6) is negative for $-n \le m \le n$. Consequently, $u_g^{(s+1)} = -1$ and so the desired conclusion $u_g^{(s+1)} = u_g^{(s)}$ is again true.

To prove the converse of the lemma we must show that if $u_g^{(s+1)} = u_g^{(s)}$ for any possible values of the local decisions $u_i^{(s+1)}$, $i = 1, 2, \ldots, n$, then $(P_{fg}^{(s)}, P_{dg}^{(s)}) \in \Re_S$. Again, two cases arise according to the two possible values of the global decision $u_g^{(s)}$. In the case when $u_g^{(s)} = 1$, we have by hypothesis that $u_g^{(s+1)} = 1$ for any possible values of $u_i^{(s+1)}, i = 1, 2, \ldots, n$. It follows that the argument of the $\mathbf{sgn}$ function in (A3) is nonnegative for any m with $-n \le m \le n$. In particular, for $m = -n$ we have that (A5) must be satisfied and so $\delta_0(P_{fg}^{(s)}, P_{dg}^{(s)}) \ge 0$ from (5b). In the case when $u_g^{(s)} = -1$, we have by hypothesis that $u_g^{(s+1)} = -1$ for any possible values of $u_i^{(s+1)}, i = 1, 2, \ldots, n$. Then the argument of the $\mathbf{sgn}$ function in (A6) must be negative for any m with $-n \le m \le n$. In particular, for $m = n$ we have that (A8) must be satisfied and so $\gamma_n(P_{fg}^{(s)}, P_{dg}^{(s)}) < 0$ from (5a). Thus the two possible cases $u_g^{(s)} = 1$ and $u_g^{(s)} = -1$ lead to the two inequalities $\delta_0(P_{fg}^{(s)}, P_{dg}^{(s)}) \ge 0$ and $\gamma_n(P_{fg}^{(s)}, P_{dg}^{(s)}) < 0$ and these two inequalities must simultaneously be satisfied according to the statement of the lemma. The definition of $\Re_S$ in (10) then shows that $(P_{fg}^{(s)}, P_{dg}^{(s)}) \in \Re_S$. □

Lemma 4: Under Assumption 1: $0 < P_{fg}^{(M)} < 1$, $0 < P_{dg}^{(M)} < 1$, and $(P_{fg}^{(M)}, P_{dg}^{(M)}) \notin \Re_S$.

Proof: Because $1 \le k \le n$ and $\sum_{i=0}^{n}\binom{n}{i}P_f^i(1-P_f)^{n-i} = 1$, (12) immediately give us $0 < P_{fg}^{(M)} < 1$, $0 < P_{dg}^{(M)} < 1$. Next, from (12) we also have

$$\begin{aligned}\frac{P_{dg}^{(M)}}{P_{fg}^{(M)}} &= \frac{\sum_{i=k}^{n}\binom{n}{i}P_d^i(1-P_d)^{n-i}}{\sum_{i=k}^{n}\binom{n}{i}P_f^i(1-P_f)^{n-i}}\\ &= \frac{\sum_{i=k}^{n}\binom{n}{i}P_f^i(1-P_f)^{n-i}\left[\frac{P_d^i(1-P_d)^{n-i}}{P_f^i(1-P_f)^{n-i}}\right]}{\sum_{i=k}^{n}\binom{n}{i}P_f^i(1-P_f)^{n-i}}\\ &\le \left(\frac{P_d}{P_f}\right)^n.\end{aligned} \tag{A9}$$

The last inequality in the above follows by considering the expression to its immediate left as a weighted average of the quantities $[\frac{P_d^i(1-P_d)^{n-i}}{P_f^i(1-P_f)^{n-i}}]$ for $i = k, \ldots, n$. The term to the right of the inequality is the largest of the quantities being averaged. A similar argument gives us the following:

$$\begin{aligned}\frac{1-P_{dg}^{(M)}}{1-P_{fg}^{(M)}} &= \frac{\sum_{i=0}^{k-1}\binom{n}{i}P_d^i(1-P_d)^{n-i}}{\sum_{i=0}^{k-1}\binom{n}{i}P_f^i(1-P_f)^{n-i}}\\ &= \frac{\sum_{i=0}^{k-1}\binom{n}{i}P_f^i(1-P_f)^{n-i}\left[\frac{P_d^i(1-P_d)^{n-i}}{P_f^i(1-P_f)^{n-i}}\right]}{\sum_{i=0}^{k-1}\binom{n}{i}P_f^i(1-P_f)^{n-i}}\\ &\ge \left(\frac{1-P_d}{1-P_f}\right)^n.\end{aligned} \tag{A10}$$

If $(P_{fg}^{(M)}, P_{dg}^{(M)}) \in \Re_S$, then we first must have $P_{dg}^{(M)} \ge s_0 P_{fg}^{(M)}$. Together with (A9) this implies $s_0 \le \frac{P_{dg}^{(M)}}{P_{fg}^{(M)}} \le (\frac{P_d}{P_f})^n$. Using the definition of s_0 in (8), this leads to

$$\tau \le \left[\frac{P_d(1-P_d)}{P_f(1-P_f)}\right]^n. \tag{A11}$$

In addition, if $(P_{fg}^{(M)}, P_{dg}^{(M)}) \in \Re_S$ we must also have $(1-P_{dg}^{(M)}) < s_n(1-P_{fg}^{(M)})$. From (A10) we thus have $s_n > \frac{1-P_{dg}^{(M)}}{1-P_{fg}^{(M)}} \ge (\frac{1-P_d}{1-P_f})^n$. Using the definition of s_n in (8), this implies

$$\tau > \left[\frac{P_d(1-P_d)}{P_f(1-P_f)}\right]^n. \tag{A12}$$

Because (A11) and (A12) are contradictory, it follows that $(P_{fg}^{(M)}, P_{dg}^{(M)}) \notin \Re_S$. □

Lemma 5: The real-valued functions $\beta(x,y)$ and $\Delta\beta(x,y)$ are continuous for all $(x,y) \in S$.

Proof: That $\beta(x,y)$ is continuous is immediate from its definition in (14). As for $\Delta\beta(x,y)$, from (4), (5), (14), and (15) and a bit of algebra we have

$$\begin{aligned}\Delta\beta(x,y) = -\left(\frac{1}{1+\tau}\right)\sum_{i=0}^{n}\binom{\mathbf{n}}{\mathbf{i}}\{&\gamma_{\mathbf{i}}(x,y)\mathbf{U}_{-1}(\gamma_{\mathbf{i}}(x,y))\\ &- \delta_{\mathbf{i}}(x,y)[1-\mathbf{U}_{-1}(\delta_{\mathbf{i}}(x,y))]\}.\end{aligned} \tag{A13}$$

The only discontinuous function in (A13) is $\mathbf{U_{-1}}(\cdot)$, which appears only in the combination $\xi\mathbf{U_{-1}}(\xi)$ with $\xi = \gamma_{\mathbf{i}}(x,y)$ or $\delta_{\mathbf{i}}(x,y)$. Because $\xi\mathbf{U_{-1}}(\xi)$ is continuous for all ξ, it follows that $\Delta\beta(x,y)$ is continuous for all (x,y) in S. (Remark: Because $\mathbf{F}(x,y)$ and $\mathbf{D}(x,y)$ are discontinuous, the continuity of $\Delta\beta(x,y)$ may seem to be of minor interest. However, this continuity is essential in proving Theorem 1.) □

Lemma 6: If $(x,y) \notin \bar{\Re}_S$, then $\Delta\beta(x,y) < 0$; and if $(x,y) \in \partial\Re_S$, then $\Delta\beta(x,y) = 0$. ($\bar{\Re}_S$ denotes the closure of $\Re_S$ and $\partial\Re_S$ denotes the boundary of $\Re_S$.)

Proof: If $(x,y) \notin \bar{\Re}_S$, then because $\Re_S = \{(x,y) \in S \mid \gamma_n(x,y) < 0,\ \delta_0(x,y) \ge 0\}$ either $\gamma_n(x,y) > 0$ and/or $\delta_0(x,y) < 0$. From (A13) it then follows that $\Delta\beta(x,y) \le -\frac{\gamma_n(x,y)}{1+\tau}$ and/or $\Delta\beta(x,y) \le -\frac{\delta_0(x,y)}{1+\tau}$. In either case, we have $\Delta\beta(x,y) < 0$.

If $(x,y) \in \partial\Re_S$, then either 1) $\gamma_n(x,y) = 0$ and $\delta_0(x,y) \ge 0$; or 2) $\gamma_n(x,y) \le 0$ and $\delta_0(x,y) = 0$. In both case 1) and 2), we also have $\gamma_{\mathbf{i}}(x,y) < 0$ for $i = 0, 1, \ldots, n-1$, and $\delta_{\mathbf{i}}(x,y) > 0$ for $i = 1, 2, \ldots, n$. It then follows from (A13) that $\Delta\beta(x,y) = 0$. □

Proof of Theorem 1: Let $(x^{(0)}, y^{(0)})$ be any point in $\mathbf{S}$ and let $x^{(t)} = \mathbf{F}(x^{(t-1)}, y^{(t-1)})$ and $y^{(t)} = \mathbf{D}(x^{(t-1)}, y^{(t-1)})$ for $t = 1, 2, \ldots$. If there exists an integer t_S such that $(x^{(t_s)}, y^{(t_s)})$ is in $\Re_S$, then case 1 of the theorem is true, since from Lemma 2 we would have $(x^{(t)}, y^{(t)}) = (x^{(t_s)}, y^{(t_s)})$ for all $t \geq t_S$.

If no such t_S exists, then $(x^{(t)}, y^{(t)}) \notin \Re_S$ for all t. We now show that case 2 or case 3 of the theorem must follow. From Lemma 6, the sequence of real numbers $\{\beta(x^{(t)}, y^{(t)})\}$ is monotonically decreasing. It is also bounded from below by zero, the minimum value of $\beta(x,y)$ in $\Re_S$. Hence, $\lim_{t\to\infty}(x^{(t)}, y^{(t)})$ must exist, say, $\lim_{t\to\infty}(x^{(t)}, y^{(t)}) = \alpha$.

Next, because the infinite sequence $\{(x^{(t)}, y^{(t)})\}$ is contained in the compact set $\mathbf{S}$, it must have at least one accumulation point in $\mathbf{S}$, say (ξ^*, η^*). Let $\{(\xi^{(t)}, \eta^{(t)})\}$ be a subsequence of $\{(x^{(t)}, y^{(t)})\}$ that converges to (ξ^*, η^*). We must have that $\lim_{t\to\infty}\beta(\xi^{(t)}, \eta^{(t)}) = \alpha$ since $\{\beta(\xi^{(t)}, \eta^{(t)})\}$ is a subsequence of the convergent sequence of real numbers $\{\beta(x^{(t)}, y^{(t)})\}$. From the continuity of $\beta(x,y)$ (Lemma 5), we then have $\beta(\xi^*, \eta^*) = \lim_{t\to\infty}\beta(\xi^{(t)}, \eta^{(t)}) = \alpha$.

Note that $\{(\mathbf{F}(\xi^{(t)}, \eta^{(t)}),\ \mathbf{D}(\xi^{(t)}, \eta^{(t)})\}$ is also a subsequence of $\{(x^{(t)}, y^{(t)})\}$, though not necessarily convergent. We thus also have that $\lim_{t\to\infty}\beta(\mathbf{F}(\xi^{(t)}, \eta^{(t)}), \mathbf{D}(\xi^{(t)}, \eta^{(t)})) = \alpha$. Then from the definition of $\Delta\beta(x,y)$ in (15) and from its continuity (Lemma 5), it follows that

$$
\begin{aligned}
\Delta\beta(\xi^*, \eta^*) &= \lim_{t\to\infty}\Delta\beta\big(\xi^{(t)}, \eta^{(t)}\big) \\
&= \lim_{t\to\infty}\big[\beta\big(\mathbf{F}\big(\xi^{(t)}, \eta^{(t)}\big), \mathbf{D}\big(\xi^{(t)}, \eta^{(t)}\big)\big) \\
&\quad - \beta\big(\xi^{(t)}, \eta^{(t)}\big)\big] \\
&= \lim_{t\to\infty}\beta\big(\mathbf{F}\big(\xi^{(t)}, \eta^{(t)}\big), \mathbf{D}\big(\xi^{(t)}, \eta^{(t)}\big)\big) \\
&\quad - \lim_{t\to\infty}\beta\big(\xi^{(t)}, \eta^{(t)}\big) \\
&= \alpha - \alpha = 0.
\end{aligned}
$$

We now show that (ξ^*, η^*) lies on $\partial\Re_S$. If not, then because $(\xi^{(t)}, \eta^{(t)}) \notin \Re_S$ we also have $(\xi^*, \eta^*) \notin \bar{\Re}_S$. From Lemma 6, we then have $\Delta\beta(\xi^*, \eta^*) < 0$. But this is a contradiction to $\Delta\beta(\xi^*, \eta^*) = 0$. Hence, (ξ^*, η^*) must lie on $\partial\Re_S$.

So far, we have that (ξ^*, η^*) is a solution of $\beta(x,y) = \alpha$ and lies on $\partial\Re_S$. From (14), $\beta(x,y) = \alpha$ is just the straight line $\tau x + (1-y) = \alpha(1+\tau)$; and from (10), $\partial\Re_S$ consists of portions of the straight lines $(1-y) = s_n(1-x)$ and $y = s_0 x$ (Fig. 8). Under Assumption 1, the slope τ of $\beta(x,y) = \alpha$ must lie strictly between the slopes s_0 and s_n of the two straight line boundary segments (i.e., $s_n < \tau < s_0$). Thus, the intersection of $\beta(x,y) = \alpha$ and $\partial\Re_S$ can consist either of one point (the vertex of $\Re_S$) or two points (one on each of the two straight line segments of $\partial\Re_S$). In other words, any particular sequence of iterates $\{(x^{(t)}, y^{(t)})\}$ (with $(x^{(t)}, y^{(t)}) \notin \Re_S$ for all t) either has only one accumulation point (the vertex of $\Re_S$) or two accumulation points (one on each of the two straight line segments of $\partial\Re_S$). Because S is compact, there can only be a finite number of iterates outside any neighborhood of the one or two possible accumulation points.

Next, we must show that the above two possibilities can actually be attained by our particular mapping of S. We first show that there is exactly one value of τ for which the vertex of $\Re_S$ is an accumulation point for some sequences of iterates. From Fig. 8, we see that $\Re_S$ is surrounded by the three regions numbered I, II, and III. If $(x^{(t)}, y^{(t)}) \notin \Re_S$ for all t, then eventually all of the iterates lie in the union of these three regions. In each of the regions, the functions $\mathbf{F}(x,y)$ and $\mathbf{D}(x,y)$, as given in (4), reduce to the following.

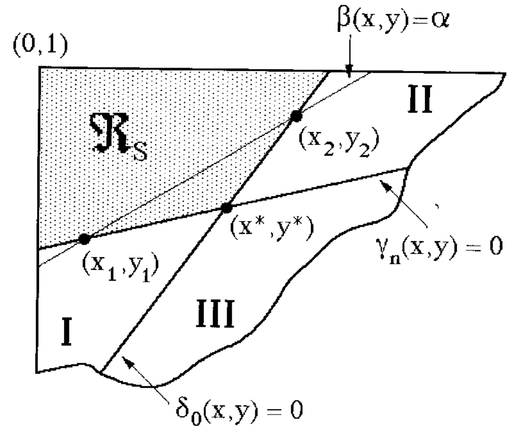

Fig. 8. Illustration for the proof of Theorem 1.

In region I:

($\gamma_{\mathbf{n}}(x,y) \geq 0$; $\gamma_{\mathbf{i}}(x,y) < 0$ for $i = 0, 1, \ldots, n-1$; $\delta_{\mathbf{i}}(x,y) \geq 0$ for $i = 0, 1, \ldots, n$)

$$
\begin{aligned}
\mathbf{F}(x,y) &= \mathbf{F_I}(x) = (1-x)P_f^n + x, \\
\mathbf{D}(x,y) &= \mathbf{D_I}(y) = (1-y)P_d^n + y.
\end{aligned}
$$

In region II:

($\gamma_{\mathbf{n}}(x,y) < 0$ for $i = 0, 1, \ldots, n$; $\delta_{\mathbf{0}}(x,y) < 0$; $\delta_{\mathbf{i}}(x,y) \geq 0$ for $i = 1, 2, \ldots, n$)

$$
\begin{aligned}
\mathbf{F}(x,y) &= \mathbf{F_{II}}(x) = [1 - (1-P_f)^n]x, \\
\mathbf{D}(x,y) &= \mathbf{D_{II}}(y) = [1 - (1-P_d)^n]y.
\end{aligned}
$$

In region III:

($\gamma_{\mathbf{n}}(x,y) \geq 0$; $\gamma_{\mathbf{i}}(x,y) < 0$ for $i = 0, 1, \ldots, n-1$; $\delta_{\mathbf{0}}(x,y) < 0$; $\delta_{\mathbf{i}}(x,y) \geq 0$ for $i = 1, 2, ., n$)

$$
\begin{aligned}
\mathbf{F}(x,y) &= \mathbf{F_{III}}(x) = P_f^n + \big[1 - P_f^n - (1-P_f)^n\big]x, \\
\mathbf{D}(x,y) &= \mathbf{D_{III}}(y) = P_d^n + \big[1 - P_d^n - (1-P_d)^n\big]y.
\end{aligned}
$$

Within region I, the iterates are tending to the point $(1,1)$, which is the attracting fixed point of the mapping $(x,y) \to (\mathbf{F_I}(x), \mathbf{D_I}(y))$. Within region II, the mapping $(x,y) \to (\mathbf{F_{II}}(x), \mathbf{D_{II}}(y))$ has the attracting fixed point $(0,0)$. And within region III, the mapping $(x,y) \to (\mathbf{F_{III}}(x), \mathbf{D_{III}}(y))$ has the attracting fixed point $(x^*, y^*) = \left(\frac{P_f^n}{P_f^n + (1-P_f)^n}, \frac{P_d^n}{P_d^n + (1-P_d)^n}\right)$. The only way that the iterates could have only one accumulation point is if that point is (x^*, y^*) and the iterates eventually all remain in region III. Equating the coordinates of the vertex as given in (13) to the coordinates (x^*, y^*) yields the expression for τ given in (16) of case 2 of the theorem. The basin of the vertex then includes part or all of region III.

For $n = 1$, however, there is a problem in that the mapping in region III reduces to the constant mapping $\mathbf{F_{III}}(x) = P_f^n$ and $\mathbf{D_{III}}(y) = P_d^n$ with fixed point $(x^*, y^*) = (P_f^n, P_d^n)$. However, (P_f^n, P_d^n), when it is the vertex of $\Re_S$, belongs to region I. Consequently, if an iterate reaches (P_f^n, P_d^n) in finite time, (which it will if it ever enters region III) it will then be driven by the mapping in region I and so will not converge to the vertex. Consequently, if $n = 1$, case 2 of the theorem cannot arise.

Finally, we show under what conditions case 3 of the theorem is possible. Let (x_1, y_1) and (x_2, y_2) be two points on the boundary of $\Re_S$ that might constitute a potential 2-cycle. We must have the following six equations satisfied:

$(1 - y_1) = s_n(1 - x_1)$—(x_1, y_1) lies on the common boundary of $\Re_S$ and region I.

$y_2 = s_0 x_2$—(x_2, y_2) lies on the common boundary of $\Re_S$ and region II.

$x_2 = \mathbf{F_I}(x_1)$—(x_1, y_1) maps to (x_2, y_2) under mapping I.

$y_2 = \mathbf{D_I}(y_1)$—(x_1, y_1) maps to (x_2, y_2) under mapping I.

$x_1 = F_{\mathbf{II}}(x_2)$—$(x_2, y_2)$ maps to (x_1, y_1) under mapping II.

$y_1 = \mathbf{D_{II}}(y_2)$—$(x_2, y_2)$ maps to (x_1, y_1) under mapping II.

Solving these six equations gives (17)–(19) for the two points of the potential 2-cycle and the corresponding value of τ in terms of P_f, P_d, and n. Notice that (x_1, y_1) belongs to region I, while (x_2, y_2) belongs to $\Re_S$. These two points are not an actual 2-cycle, but the iterates converge to them as if they were an attracting 2-cycle as long as the iterates remain in some neighborhood of the two points lying in regions I and II.

REFERENCES

[1] S. Alhakeem and P. K. Varshney, “Decentralized Bayesian detection with feedback,” *IEEE Trans. Syst., Man, Cybern. A*, vol. 26, pp. 503–513, July 1996.

[2] Z. Chair and P. K. Varshney, “Optimal data fusion in multiple sensor detection systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 98–101, 1986.

[3] C. Chong, S. Mori, and K. Chang, “Distributed multitarget multisensor tracking,” in *Multitarget-Multisensor Tracking, Advanced Applications*, Y. Bar-Shalom, Ed. Norwood, MA: Artech House, 1990, Section 8.2.1.

[4] H. R. Hashemi and I. B. Rhodes, “Decentralized sequential detection,” *IEEE Trans. Inform. Theory*, vol. 35, pp. 509–520, 1989.

[5] M. Kam, W. Chang, and Q. Zhu, “Hardware complexity of binary distributed detection systems with isolated local Bayesian detectors,” *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 565–571, 1991.

[6] M. Kam, Q. Zhu, and W. S. Gray, “Optimal data fusion of correlated local decisions in multiple sensor detection systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-28, no. 3, pp. 916–920, 1992.

[7] J. M. Manyika and H. F. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information—Theoretic Approach*. New York: Ellis Horwood, 1994, Section 3.2.

[8] J. Papastavrou and M. Athans, “Distributed detection by a large team of sensors in tandem,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 639–653, 1992.

[9] ——, “On optimal distributed decision architectures in a hypothesis testing environment,” *IEEE Trans. Automat. Contr.*, vol. 37, no. 8, pp. 1154–1169, 1992.

[10] A. R. Reibman and L. W. Nolte, “Optimal detection and performance of distributed sensor systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 1, pp. 24–30, 1987.

[11] H. M. H. Shalaby and A. Papamarcou, “A note on the asymptotics of distributed detection with feedback,” *IEEE Trans. Inform. Theory*, vol. 39, pp. 633–643, Mar. 1993.

[12] P. F. Swaszek, “On the performance of serial network in distributed detection,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 254–259, 1993.

[13] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, “Optimization of detection networks: Part I—Tandem structures,” *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 5, pp. 1044–1059, 1991.

[14] D. Teneketzis, “The decentralized Wald problem,” in *Proc. IEEE 1982 Int. Large-Scale Systems Symp.*, Virginia Beach, VA, 1982, pp. 423–430.

[15] R. R. Tenney and N. R. Sandell, “Detection with distributed sensors,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 501–509, 1981.

[16] J. N. Tsitsiklis, “Decentralized detection by a large number of sensors,” *Math. Contr., Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[17] J. N. Tsitsiklis and M. Athans, “On the complexity of distributed decision problems,” *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 5, pp. 440–446, 1985.

[18] H. V. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1969.

[19] R. Viswanathan, S. Thomopoulos, and R. Tumuluri, “Optimal serial distributed decision fusion,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-24, no. 4, pp. 366–375, 1988.

[20] A. D. Whalen, *Detection of Signals in Noise*. New York: Academic, 1971, Section 7.5.

Moshe Kam (S’75–M’77–SM’92) received the B.Sc. degree in electrical engineering from Tel Aviv University, Tel Aviv, Israel, in 1977, and the M.S. and Ph.D. degrees from Drexel University, Philadelphia, PA, in 1985 and 1987, respectively, under the supervision of Paul Kalata.

He is a Professor of Electrial and Computer Engineering at Drexel University. His research interests include detection and estimation, data fusion, robot navigation, and pattern recognition.

Dr. Kam is the recipient of a National Science Foundation Presidential Young Investigator Award and the C. Holmes MacDonald award for the outstanding young electrical engineering education.

Chris Rorres received the B.S. degree in physics from Drexel University, Philadelphia, PA, in 1963 and the Ph.D. degree in mathematics from New York University in 1969 under the supervision of Joseph B. Keller.

He is a Professor of Mathematics in the Department of Mathematics and Computer Science at Drexel University, Philadelphia, PA. His research interests include acoustical scattering, dynamical systems, and optimization theory. He is coauthor of the texts *Applications of Linear Algebra* and *Elementary Linear Algebra with Applications* and also maintains a Web site on Archimedes at http://www.mcs.drexel.edu/~crorres/Archimedes/contents.html.

Wei Chang received the B.S. and M.S. degrees in 1989 and the Ph.D. degree in 1992 from Drexel University, Philadelphia, PA, under the supervision of Moshe Kam. His master's thesis was on "Systems of random iterative mappings with a common fixed point," and his Ph.D. dissertation was on "distributed detection with limited computational resources."

He is a Principal Engineer at Liberty Technologies. Previously he was a Research Engineer at the U.S. Army Harry Diamond Laboratories. His research interests include industrial sensor design, signal detection and signal processing, and control of distributed systems.

Xiaoxun Zhu received the B.S. and M.S. degrees in 1994 and the Ph.D. degree in 1997, under the supervision of Moshe Kam. His master's thesis was on "extensions of the Chair–Varshney algorithm for decision fusion," and his Ph.D. dissertation was on "decision-level data fusion."

He is a Research Engineer at Metrologic Instruments, Blackwood, NJ. His research interests include image processing and pattern recognition, especially in laser scanning applications.

Dr. Zhu was a Hill Fellow at the College of Engineering of Drexel University.