s.t.

$$
\begin{aligned}
x_2 &= 1 \\
x_1 - 2.2x_2 + x_4 - \delta_1 &\leq 12 \\
-x_1 + 2.2x_2 - x_4 &\leq -12 \\
bx_3 - ax_4 - \delta_3 &\leq 14.9 \\
-bx_3 + ax_4 &\leq -14.9 \\
bx_1 - ax_2 + x_3 - 2.2x_4 &\leq 21.49 \\
-bx_1 + ax_2 - x_3 + 2.2x_4 + \delta_2 &\leq -21.49 \\
\delta_1 &\geq 0 \\
\delta_2 &\leq 0 \\
\delta_3 &\geq 0 \\
g_1 - \gamma_1^2 &\leq 0 \\
g_2 - \gamma_2^2 &\leq 0.
\end{aligned}
$$

Choosing $Q = I$, the convergence was achieved, after two iterations, to the following set of controller parameters:

$$k_1 = \frac{5.701s + 36.508}{s + 8.477}$$

and

$$\delta_1 = 0, \qquad \delta_2 = -0.4956, \qquad \text{and} \qquad \delta_3 = 3.2598.$$

The objective function of this solution is 10.8719, and the poles of the closed loop are in the region of Fig. 3 for all possible plants.

VI. Conclusions

A computational method for designing controllers which attempt to place the roots of the characteristic polynomial of an uncertain system inside some prescribed regions has been presented. The approach is an extension of a previous work on robust characteristic polynomial assignment (Rotstein *et al.* [6]) by extending the original formulation through the addition of constraints that relate the pole position in the open left-half plane to the real variation of the coefficients in the characteristic polynomial.

The general problem is formulated as a semi-infinite programming problem which can be solved using standard techniques. Therefore, we feel that our solution to this problem is an important addition to the "tool kit" of the process control engineer.

References

[1] V. Kučera, *Discrete Linear Control: The Polynomial Approach.* Chichester: Wiley-Interscience, 1979.

[2] K. Åström and B. Wittenmark, *Computer Controlled Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[3] Y. Soh, R. Evans, I. Petersen, and R. Betz, "Robust pole assignment," *Automatica*, vol. 23, no. 5, pp. 601–610, 1987.

[4] Y. Soh, "Robust pole-placement for uncertain interval systems," *Proc. IEE*, vol. 136, pt. D, pp. 301–306, 1989.

[5] H. Rotstein, J. Bandoni, A. Desages, and J. Romagnoli, "Mathematical programming with linear uncertain constraints—Application to robust control," *Comput. Chem. Eng.*, vol. 14, no. 4/5, pp. 373–379, 1990.

[6] H. Rotstein, R. Sanchez Peña, J. Bandoni, A. Desages, and J. Romagnoli, "Robust characteristic polynomial assignment," *Automatica*, vol. 27, no. 4, pp. 711–715, 1991.

[7] C. Chen and Y. Hsu, "Pole assignment using dynamic output feedback compensators," *Int. J. Contr.*, vol. 45, pp. 1985–1994, 1987.

[8] S. Bhattacharyya, *Robust Stabilization Against Structured Perturbations*, Lecture Notes in Control and Information Sciences. Berlin: Springer, 1987.

[9] A. J. Markushevich, *The Theory of Analytic Functions.* India: Hindustan, 1963.

[10] G. Bortolotto, J. Romagnoli, and A. Desages, "Robust performance determination in the presence of plant uncertainty," *Comput. Chem. Eng.*, vol. 16, suppl., pp. S173–S180, 1992.

[11] E. Polak, "On the mathematical foundation of non-differentiable optimization in engineering design," *SIAM Rev.*, vol. 29, pp. 21–89, 1987.

[12] C. Gonzaga and E. Polak, "On constraint dropping schemes and optimality functions for a class of outer approximation algorithms," *SIAM J. Contr. Optimiz.*, vol. 17, pp. 477–493, 1979.

Some Properties of Optimal Thresholds in Decentralized Detection

William W. Irving and John N. Tsitsiklis

Abstract— **A decentralized Bayesian hypothesis testing problem is considered. It is analytically demonstrated that for the known signal in the Gaussian noise binary hypothesis problem, when there are two sensors with statistically independent identically distributed Gaussian observations (conditioned on the true hypothesis), there is no loss in optimality in using the same decision rule at both sensors. Also, a multiple hypothesis problem is considered; some structure is analytically established for an optimal set of decision rules.**

I. Introduction

The (static) decentralized detection problem is defined as follows. There are M hypotheses, $H_1, \cdots, H_M$, with known prior probabilities $P(H_j) > 0$ $(j = 1, \cdots, M)$, and there are N peripheral sensors. Let y_i $(i = 1, \cdots, N)$ be a random variable, denoting the observation of the ith sensor. The y_i's are conditionally independent and identically distributed given any hypothesis, with a known conditional distribution $P(y|H_j)$ $(j = 1, \cdots, M)$. Let D be a positive integer. Each peripheral sensor, upon receiving its observation, evaluates a message $u_i = \gamma_i(y_i) \in \{1, \cdots, D\}$. The messages $u_1, \cdots, u_N$ are all transmitted to a fusion center, where a decision rule γ_0: $\{1, \cdots, D\}^N \rightarrow \{1, \cdots, M\}$ is used to decide in favor of one of the M hypotheses. The objective is to choose the decision rules $\gamma_0, \gamma_1, \cdots, \gamma_N$ (collectively known as a *strategy*) of the sensors and fusion center so that the fusion center's probability of error is minimized.

Over the past decade, this problem and its variants have received a fair amount of attention in the literature [2]–[6]. In this paper, we study the structure of optimal strategies for two specific instances of the problem. By applying novel analytical techniques, we prove some modestly interesting properties of the optimal strategies.

First, we consider a binary hypothesis $(M = 2)$, binary messages $(D = 2)$ instance. It is well known that for the $M = 2/D = 2$ case, any optimal strategy has the following structure. Each one of the sensors evaluates its message u_i using a likelihood ratio test with an appropriate threshold. Then, the fusion center combines the

Manuscript received October 13, 1992; revised February 24, 1993. The work of W. W. Irving was supported by the Defense Advanced Research Projects Agency under Contract F19628-90-C-002. The work of J. N. Tsitsiklis was supported by the ARO under Contract DAAL03-86-K-0171.

The authors are with the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9216439.

0018-9286/94$04.00 © 1994 IEEE

sensor messages into a final decision by performing its own likelihood ratio test. The optimal value of the threshold of each sensor is obtained by first finding all solutions of a set of coupled algebraic equations, and by then selecting the solution that results in minimum probability of error. Unfortunately (and contrary to intuition), it is not necessarily true that all sensors should use the same threshold, even though the observations of the sensors are identically distributed and conditionally independent (see [5] and [6] for examples). Despite this caveat, most of the problems analyzed in the literature have been found to have globally optimal solutions in which each sensor uses the same threshold [2], [4]. However, global optimality has virtually always been established by numerical (as opposed to analytical) methods. In [6], some general analytical results are developed, but these are with respect to only local optimality. In this paper, we analytically demonstrate that under certain assumptions, global optimality can be achieved even when all of the sensors are restricted to use the same threshold.

Second, we consider a multiple-hypothesis ($M \geq 2$), D-message ($D \geq 2$) instance. Little is known about the structure of optimal strategies for the $M > 2$ case, even for specific instances. We use a bounding argument to establish some structure to an optimal strategy for our instance.

II. Binary Hypothesis, Gaussian Problem

A. Problem Formulation

We consider the following binary hypothesis testing problem:

$$H_1: \quad y_i = w_i$$

$$H_2: \quad y_i = s + w_i, \qquad i = 1, 2.$$

There are two sensors; the noise corrupting the observations of the sensors consists of a pair of statistically independent Gaussian random variables, with density

$$p_{w_i}(w) = \frac{\exp(-w^2/2)}{\sqrt{2\pi}}, \qquad i = 1, 2. \tag{1}$$

We only consider the case of $D = 2$. It is well known [4] that for this case, there is no loss in optimality in using decision rules of the form

$$\frac{P(y_i|H_2)}{P(y_i|H_1)} \underset{u_i=1}{\overset{u_i=2}{\gtrless}} \alpha_i, \qquad i = 1, 2 \tag{2}$$

where α_1, α_2 are scalar constants; an equivalent (and often more useful) form is

$$y_i \underset{u_i=1}{\overset{u_i=2}{\gtrless}} T_i, \qquad i = 1, 2 \tag{3}$$

where

$$T_i = \frac{1}{s}\left(\ln \alpha_i + \frac{s^2}{2}\right), \qquad i = 1, 2. \tag{4}$$

For these decision rules, we have used the notation

$$y_i \underset{u_i=1}{\overset{u_i=2}{\gtrless}} T_i \Leftrightarrow \gamma_i(y_i) = u_i = \begin{cases} 1, & y_i < T_i \\ 2, & y_i \geq T_i. \end{cases}$$

We have the following proposition. Although it might seem to be an intuitively obvious result, it is actually the first result of this kind to appear in the literature.

Proposition 1: For the hypothesis testing problem described above, there is no loss in optimality in imposing the constraint

$$T_1 = T_2.$$

B. Proof

The proof of the above proposition proceeds as follows.

1) Overview: First, it is straightforward to show that there always exists a globally optimal strategy under which the fusion center uses the OR rule,

$$\gamma_0(u_1, u_2) = \begin{cases} 1, & u_1 = u_2 = 1 \\ 2, & \text{otherwise} \end{cases}$$

or the AND rule,

$$\gamma_0(u_1, u_2) = \begin{cases} 2, & u_1 = u_2 = 2 \\ 1, & \text{otherwise} \end{cases}$$

or ignores at least one of the sensors. In the case that at least one of the sensors is ignored, the threshold of a sensor that is ignored can be set to the threshold of the other sensor without any loss of performance; once this is done, the optimal fusion rule is either the OR rule or the AND rule. Thus, we restrict consideration to these two fusion rules.

Suppose, now, that the fusion rule is either OR or AND. For a fixed fusion rule, the optimal values of T_1 and T_2 are coupled by equations of the form

$$T_1 = f(T_2), \qquad T_2 = f(T_1) \tag{5}$$

where $f(\cdot)$ depends on the particular fusion rule (AND or OR).

For the OR and AND fusion rules, one can show that

$$\frac{df(t)}{dt} > -1, \qquad \forall\, t. \tag{6}$$

Note the strict inequality. This inequality implies that

$$f(T_1) - f(T_2) \leq T_2 - T_1 \tag{7}$$

with equality iff $T_1 = T_2$. But from (5),

$$f(T_1) - f(T_2) = T_2 - T_1. \tag{8}$$

Combining (7) and (8), we see that all threshold pairs (T_1, T_2) that satisfy (5) must also satisfy $T_1 = T_2$. Since a globally optimal strategy must satisfy (5), we conclude that it too must satisfy $T_1 = T_2$. Note that we have characterized the structure of an optimal strategy without explicitly finding one. Technically, we must still demonstrate that (6) holds for our instance.

2) Details: In this section, we develop the form of $f(\cdot)$ and verify (6). We go through these details only for the case of OR fusion; the details for AND fusion are virtually identical, and so they are omitted.

It has been shown [2], [4] that when decision rules of the form (2) are used, the optimal values of α, α_2 are coupled by

$$\alpha_i = \frac{P(H_1)}{P(H_2)} \frac{\Pr[U_{3-i} = 1|H_1]}{\Pr[U_{3-i} = 1|H_2]}, \qquad i = 1, 2. \tag{9}$$

In terms of T_i [see (3)], we have

$$\Pr[U_i = 1|H_j] = \begin{Bmatrix} \Phi(T_i), & j = 1 \\ \Phi(T_i - s), & j = 2 \end{Bmatrix} \tag{10}$$

where

$$\Phi(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2)\, dt.$$

Thus, combining (4), (9), and (10), we obtain

$$f(t) = \frac{1}{s}\left[\frac{s^2}{2} + \ln\frac{P(H_1)}{P(H_2)} + \ln\frac{\Phi(t)}{\Phi(t-s)}\right].$$

To establish the validity of (6), we first note that

$$\frac{df(t)}{dt} = \frac{1}{s\sqrt{2\pi}}[g(t) - g(t-s)] \qquad (11)$$

where

$$g(t) = \frac{\exp(-t^2/2)}{\Phi(t)}.$$

We will now show that

$$\frac{dg(u)}{du} > -\sqrt{2\pi}. \qquad (12)$$

This is useful because it implies that

$$g(t) - g(t-s) > -s\sqrt{2\pi}$$

which, in light of (11), implies the validity of (6).

To establish (12), first note that

$$\frac{dg(u)}{du} = -g(u)\left[u + \frac{1}{\sqrt{2\pi}}g(u)\right]. \qquad (13)$$

We bound this derivative by separately considering negative and nonnegative values of u.

For negative u, we exploit the bound [1]

$$\sqrt{\frac{2}{\pi}}\frac{\exp(-u^2/2)}{|u| + \sqrt{u^2+4}} < \Phi(u), \qquad u \le 0.$$

This bound implies that

$$g(u) < \sqrt{\frac{\pi}{2}}(|u| + \sqrt{u^2+4}), \qquad u \le 0, \qquad (14)$$

and

$$u + \frac{1}{\sqrt{2\pi}}g(u) < \frac{1}{2}(u + \sqrt{u^2+4}), \qquad u \le 0. \qquad (15)$$

Combining (13), (14), and (15), we obtain the simple bound

$$-\sqrt{2\pi} < \frac{dg(u)}{du}, \qquad u \le 0, \qquad (16)$$

thereby verifying (12) for negative u.

Now, we bound the derivative of $g(\cdot)$ for nonnegative u. We have

$$\begin{aligned} ug(u) &= \frac{u\exp(-u^2/2)}{\Phi(u)} \\ &< \frac{\max_{u\ge 0}[u\exp(-u^2/2)]}{\min_{u\ge 0}[\Phi(u)]} \\ &= \frac{1\exp(-1)}{\Phi(0)} \\ &= 2\exp(-1), \qquad u \ge 0. \end{aligned}$$

Also,

$$\begin{aligned} \frac{1}{\sqrt{2\pi}}g^2(u) &\le \frac{1}{\sqrt{2\pi}}\left(\frac{\max_{u\ge 0}[\exp(-u^2/2)]}{\min_{u\ge 0}[\Phi(u)]}\right)^2 \\ &= \frac{1}{\sqrt{2\pi}}\left(\frac{\exp(0)}{\Phi(0)}\right)^2 \\ &= \frac{4}{\sqrt{2\pi}}, \qquad u \ge 0. \end{aligned}$$

Combining these last two results, we conclude that

$$\begin{aligned} \frac{dg(u)}{du} &> -2\exp(-1) - \frac{4}{\sqrt{2\pi}} \\ &> -\sqrt{2\pi}, \qquad u \ge 0, \end{aligned} \qquad (17)$$

thereby verifying (12) for nonnegative u.

Together, (16) and (17) verify (12) for all u. Thus, from the discussion immediately following (12), we conclude that (6) is valid, which concludes the proof. Q.E.D.

C. Remarks

A review of the above proof will reveal that its success seems inextricably tied to the special structure of the hypothesis testing instance—that is, to the constraints $N = 2$ and $D = 2$. This state of affairs seems to reinforce the notion that analytical results are very difficult to develop in decentralized detection theory.

A few comments are in order concerning the relation of this result to the work in [6]. There, the authors considered a generalization of this problem in which there are N sensors. They analytically established that for any fixed "k-out-of-N" fusion rule,

$$\gamma_0(u_1, u_2, \cdots, u_N) = \begin{cases} 1, & u_i = 2 \text{ for fewer than } k \text{ values of } i \\ 2, & \text{otherwise} \end{cases}$$

there exists some T^* for which it is *locally* optimal for the sensors to use decision rules of the form (3) with $T_1 = T_2 = \cdots = T_N = T^*$. In fact, this result was shown to hold for a class of binary hypothesis testing problems that extends beyond the case of Gaussian probability densities. However, we emphasize that only local optimality was established, and so in the context of our rather specialized hypothesis testing problem, our result is stronger.

III. M-ARY HYPOTHESIS GAUSSIAN PROBLEM

A. Problem Formulation

We now consider the following M-ary hypothesis testing problem:

$$H_j: \quad y_i = s_j + w_i, \qquad 1 \le i \le N, \quad 1 \le j \le M. \qquad (18)$$

There are N sensors; the noise corrupting the observations of the sensors is a collection of mutually independent Gaussian random variables, with density given by (1). For a fixed but arbitrary integer $D \ge 2$, we analyze the structure of an optimal strategy.

B. Structure of Optimal Decision Rules

It is clear that without loss of optimality, each of the sensors can use a decision rule of the form

$$\gamma_i(y_i) = \begin{cases} d_1, & y_i \le T_{1i} \\ d_2, & T_{1i} < y_i \le T_{2i} \\ \vdots & \\ d_{k_i+1}, & y_i > T_{k_i i} \end{cases} \qquad (19)$$

where

$$k_i \ge 1$$

$$d_l \in \{1, \cdots, D\}, \qquad d_l \ne d_{l+1}, \quad 1 \le l \le k_i$$

$$T_{1i} < T_{2i} < \cdots < T_{k_i i}.$$

This is just a formal way of saying that with respect to the real-axis observation space, any decision rule can be expressed as a set of decision regions. For example, if $D = 2$, then any decision rule can be expressed as alternating regions of "send message 1" and "send message 2." In this formalism, sensor i has k_i different thresholds, each acting as an alternation point from one message region to the next message region.

In general, there is no known bound on the number of regions needed for an optimal decision rule for the multiple hypothesis testing problem. However, for the Gaussian problem just described, there is an upper bound on k_i. In particular, we have the following proposition.

Proposition 2: For the hypothesis testing problem described above, there always exists an optimal set of decision rules of the form in (19) for which

$$k_i \le (M-1)\frac{D(D-1)}{2}, \qquad 1 \le i \le N.$$

C. Proof

We will need the following lemma for the proof of the proposition.

Lemma: Let $\alpha_1, \beta_1, \cdots, \alpha_N, \beta_N$ be an arbitrary collection of finite, real scalar, where $\alpha_i \neq 0$ for at least one value of i, $1 \le i \le N$. The equation

$$\sum_{i=1}^{N} \alpha_i \exp(\beta_i x) = 0$$

has no more than $N-1$ finite real roots.

Proof: We establish the lemma by induction. For $N = 1$, there are clearly no finite real roots. Now, assume that the lemma is true for $N = k$, $k \ge 1$, and consider the case of $N = k+1$. Then, it is easy to see that

$$\begin{aligned} &R_x\left\{\sum_{n=1}^{k+1}\alpha_n \exp(\beta_n x)\right\} \\ &= R_x\left\{\exp(\beta_{k+1}x)\left[\alpha_{k+1} + \sum_{n=1}^{k}\alpha_n \exp[(\beta_n - \beta_{k+1})x]\right]\right\} \\ &= R_x\left\{\alpha_{k+1} + \sum_{n=1}^{k}\alpha_n \exp[(\beta_n - \beta_{k+1})x]\right\} \\ &\le R_x\left\{\sum_{n=1}^{k}(\beta_n - \beta_{k+1})\alpha_n \exp[(\beta_n - \beta_k + _1)x]\right\} + 1 \\ &\le (k-1) + 1. \end{aligned}$$

Here, we have used the notation

$$R_x\{f(x)\} \equiv |\{x \mid x \in \Re,\ x \text{ finite},\ f(x) = 0\}|, \tag{20}$$

that is, it is the number of finite real roots of the enclosed expression. The equalities on the first and second lines are straightforward. The third line follows because the number of roots of a function is upper bounded by one plus the number of roots of its derivative. The final line follows from the induction hypothesis. Q.E.D.

Returning now to the proposition, we first prove the result for the special case of $D = 2$. The generalization to arbitrary D, $D \ge 2$ will then readily follow.

For $D = 2$, the peripheral sensor person-by-person optimality condition can be expressed as [3]

$$\gamma_i(y) = \underset{d=1,2}{\text{argmin}} \sum_{j=1}^{M} b_i(d, H_j) P(y|H_j) \tag{21}$$

where

$$\begin{aligned} b_i(d, H_J) = \Pr[&\gamma_0(U_1, \cdots, U_{i-1}, d, U_{i+1}, \cdots, U_N) \\ &\neq j \mid H_j] P(H_j). \end{aligned} \tag{22}$$

The important point is that $b_i(d, H_j)$ is a scalar whose value depends on the decision rules employed by all of the other sensors; the specific form $b_i(d, H_j)$ is not important for this discussion.

777

A simple rearrangement of (21) yields the rule

$$\frac{1}{\sqrt{2\pi}}\left(\sum_{j=1}^{M} c_{ij}\exp[-(y-s_j)^2/2]\right) \underset{u_i=1}{\overset{u_i=2}{\gtrless}} 0$$

or, equivalently, after multiplying both sides by $\exp(y^2/2)$,

$$\left(\sum_{j=1}^{M}\alpha_{ij}\exp(s_j y)\right) \underset{u_i=1}{\overset{u_i=2}{\gtrless}} 0 \tag{23}$$

where

$$\begin{aligned} c_{ij} &= b_i(d=1, H_j) - b_i(d=2, H_j) \\ \alpha_{ij} &= c_{ij}\exp\left(\frac{-s_j^2}{2}\right). \end{aligned}$$

The form of the decision rule in (23) makes it clear how to find the thresholds T_{ij} for the decision (19). In particular, each real root (with respect to y_i) of the function on the left-hand side of (23) marks the location of a threshold. Thus, any upper bound that we can find for the number of real roots of that function is also an upper bound on the number of thresholds in an optimal rule. But from the lemma, we immediately obtain the upper bound $M-1$, thus establishing the proposition for the special case of $D = 2$.

Now, we generalize the result to arbitrary D. To motivate the generalization, consider the case of $D = 3$. It is straightforward to see that the number of decision region transitions cannot be more than the number of intersections [as in (23)] between decisions 1 and 2 plus between 1 and 3, plus between 2 and 3. But, from the $D = 2$ analysis, the maximum number of intersections for each of these is $M-1$; in general, then, we must consider $\binom{D}{2}$ pairs of intersections, which yields the upper bound in the proposition. Q.E.D.

References

[1] M. Abromowitz and I. Stegun, *Handbook of Mathematical Functions.* Washington, DC: National Bureau of Standards, 1964.

[2] A. R. Reibman and L. W. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 789–797, 1987.

[3] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing, Vol. 2: Signal Detection.*

[4] R. R. Tenney and N. R. Sandell, Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501–510, 1981.

[5] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Contr., Signals, Syst.*, vol. 1, pp. 167–182, 1988.

[6] P. Willet and D. Warren, "Decentralized detection: When are identical sensors identical?," in *Proc. 1991 Conf. Inform. Sci. Syst.*, Mar. 1991, pp. 287–292.