SLAC – PUB – 4510
December 1987
(N)

DISTRIBUTED DETECTOR SYSTEMS AS A PROBLEM IN OPTIMIZATION WITH NONLINEAR CONSTRAINTS

R. BLANKENBECLER*
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305
and
PAUL B. KANTOR†
Tantalus, Inc., 3257 Ormond Road, Cleveland, Ohio 44118

Summary

In this paper, optimal control theory is applied to the design of decentralized sensor systems. Lagrange inequality multipliers are used to determine the optimal design parameters. Several models of possible response functions are fully discussed as examples of our technique.

1. Introduction and Definition of the Problem

There are many situations in science and engineering in which information is gathered from a variety of sensors and must be abstracted or summarized for future processing in order to comply with communication, storage, or processing constraints. The simplest example is the case in which a binary decision must be made based upon information sources that are constrained to transmit a binary signal. Examples include data from devices monitoring the performance of a power network, data from an array of elementary particle detectors, the coordination of radar or infrared signals, and so on.

In general, the communication restrictions may be lifted with some increase in cost; thus the examples under discussion represent a special case. As we shall see, even this simple case (two alternative states, two possible actions, two-fold signals, and two detectors) presents challenging problems of analysis. Discussions have been given by Srinivasan for more than two detectors[1] and with applications to a specific choice of the detector characteristics.[2]

Discussion of a case with distributed action is given by Tenney and Sandell.[3] A discussion for specific (series) topologies is given by Ekchian and Tenneyl.[4] Related problems have been discussed by Chair and Varshney,[5] by Reibman and Nolte,[6] and by Sadjadi.[7]

Quite generally, the performance of an entire network is summarized by four probabilities $p_r(y,H)$, of which only two are independent. (Here, $H = H_0, H_1$ represents two hypothesis about the world and $y = y_0, y_1$ represents two possible actions or determinations. This notation will be made more precise shortly.)

Several problems may be formulated, including
(i) min $p_r(y_1, H_0)$ subject to $p_r(y_0, H_1) \le p_m^0$.
(ii) min $p_r(y_0, H_1)$ subject to $p_r(y_1, H_0) \le p_f^0$.
(iii) min $Ap_r(y_1, H_0) + Bp_r(y_0, H_1)$.

The first and second problems correspond to setting acceptable error rates; the third arises when there is a tradeoff between the two types of error. The coefficients A, B may be positive or negative.

The physical characteristics of an individual detector constrain the achievable values of $p_r(y_1, H_0)$ and $p_r(y_0, H_1)$. The design of a network is then a selection from among a discrete set of topologies, with each topology tuned to give its best possible performance. The tuning is a constrained optimization, with the constraints determined by the achievable values of $p_r(y_1, H_0)$ and $p_r(y_0, H_1)$.

*Work supported by the Department of Energy, contract DE-AC03-76SF00515.
†Supported in part by the Office of Naval Research, contract N00014-87-C-0695.

Our work has many points of contact with previous work. We utilize a Lagrangian formulation to deal with the optimization problem involving equality and inequality constraints. Three problems are presented in detail, involving the cases of exponential response functions, special sums of exponentials, and block functions. We trace the behavior of the system tuning and the optimal cost as a function of the detector discrimination.

This paper is the first of a series whose goal is to clarify the relations between topics in distributed detection, optimal control, and experimental design, thereby leading to a more intuitive or "physical" understanding of the problems of distributed detection and sensing.

1.1 General Introduction

There are two possible states (of the world) H_0 and H_1. The prior probabilities of these two states are p_0 and p_1, where

$$\begin{aligned} p_0 &= p_{rior}(H_0) \\ p_1 &= p_{rior}(H_1) \ . \end{aligned} \tag{1}$$

There are two possible courses of action ("measures") denoted by m_0 and m_1.

The assumed *cost* function is $C(m, H)$, where

$$\begin{aligned} C(m_0,H_0) &= u_0 & C(m_0,H_1) &= u_1 + w_{01} \\ C(m_1,H_0) &= u_0 + w_{10} & C(m_1,H_1) &= u_1 \ . \end{aligned} \tag{2}$$

The expectation value of the cost function is to be minimized over the various design parameters, those in the response functions and those in the probability functions. As will become clear later in our discussion, the separate cost parameters u_0 and u_1 do not matter when the expected cost is minimized; the minimum depends only on a ratio involving the differences in the cost for a given H_j, namely w_{10} and w_{01}.

The essential point is that for the case of only two possible states of the world, the preferred action is determined by a single real number, determined by the posterior odds for the H_j. This is true because, using linear cost theory, the information in Eq. (2) is summarized by the intersection point of two straight lines; one describes the cost of action m_0 as a function of p_0 while the other describes the cost of action m_1.

1.2 Properties of the Integrator

For our model we choose a fusion structure in which signals are processed locally at each detector, with messages fed to a single integrator

$$A \longrightarrow C \longleftarrow B \ .$$

The problem is to design an integrator C and tune the sensors (A, B). Each of the two sensors detects some signal (y) and sends the central integrator a signal u_i. In general, these signals need not be binary. The integrator then chooses action m_0 or m_1, and this choice is determined by the fusion rules. The rules for both the sensors and the integrator are to be chosen so that the expected cost is minimized.

The integrator's actions are completely described by a matrix (with two adjustable parameters) that describes the probability of choosing measure m_0, given the signals u_i from detector $i = a, b$. This matrix will be denoted by $p(m_i|u_a, u_b)$, where

Submitted to Transactions on Aerospace and Electronic Systems

$$\begin{array}{ll} p(m_0|0,0) = 1 & p(m_0|0,1) = g \\ p(m_0|1,0) = d & p(m_0|1,1) = 0 \end{array} \tag{3}$$

or, in an alternative matrix notation,

$$p(m_0|\vec{u}) = \begin{array}{c} \\ u_a = 0 \\ u_a = 1 \end{array} \begin{array}{c} u_b = 0 \quad u_b = 1 \\ \begin{pmatrix} 1 & g \\ d & 0 \end{pmatrix} \end{array} . \tag{4}$$

The probability of choosing m_1 must be the complement, element by element,

$$p(m_1|\vec{u}) = 1 - p(m_0|\vec{u}) . \tag{5}$$

We exclude the possibility of a third course of action. The design parameters g and d are to be fixed by the optimization; they define the rule to follow when the two detectors disagree. If the two detectors are identical, then we expect that $d = g$ and that they will be 0 or 1 depending on the costs and the details of the sensitivities of the detectors.

1.3 Definition of the Detectors

Now consider the detectors in more detail. Each detector, labeled a or b, produces a single "meter reading" y_i, $(i = a, b)$, in response to the state of nature. The probabilities, $p_d(y, H)$, that the value of the reading is y for the state of the environment H for each detector is

Detector→	a	b
$p_d(y; H_0)$	$f_0^a(y)$	$f_0^b(y)$
$p_d(y; H_1)$	$f_1^a(y)$	$f_1^b(y)$.

The quantity y must now be converted to a yes or no signal ($u = 0$ or $u = 1$).

The effect of the decision process at each detector may be summarized *completely* by a table giving the decision strategy or probability of response $p_i(u; H)$ for each of the detectors. For detector a:

$$p_a(u_a; H) = \begin{array}{c} \\ u_a = 0 \\ u_a = 1 \end{array} \begin{array}{c} H_0 \quad\quad H_1 \\ \begin{pmatrix} a & 1-A \\ 1-a & A \end{pmatrix} \end{array} , \tag{6}$$

while for detector b:

$$p_b(u_b; H) = \begin{array}{c} \\ u_b = 0 \\ u_b = 1 \end{array} \begin{array}{c} H_0 \quad\quad H_1 \\ \begin{pmatrix} b & 1-B \\ 1-b & B \end{pmatrix} \end{array} . \tag{7}$$

1.4 Design Parameters

Therefore, the full set of parameters to be determined by optimization is

$$g, d \qquad a, A, b, B .$$

The first two describe the operation of the "integrator" that processes the two signals $\vec{u}$ of the sensor stations to form the output decision m. The last four describe the operation of the "sensors" — they take the detected signals, apply their respective detection criteria, and form their individual output signals $\vec{u}$.

1.5 Properties of the Detectors

A generalized detector uses the rule: if the signal y is in the region R, then the signal u_0 is sent to the integrator. Similarly, if the signal is in the complement of R, i.e., if $y \in \bar{R}$, then u_1 is sent.

If the external state is indeed H_0, then the response function of the detector is $f_0(y)$, but if it is H_1, the detector responds with $f_1(y)$ [see the table below Eq. (5)].

For any choice of R, the detection probability [see Eq. (6)] is

$$a = \int_R dy\, f_0^a(y) , \tag{8}$$

and this then implies for A,

$$A = \int_R dy\, f_1^a(y) . \tag{9}$$

Similar relations hold for b and B. If the response functions have interlaced maxima, then the region R (and $\bar{R}$) may be disconnected.

As R expands, clearly $\bar{R}$ contracts. For any fixed value of a there is a maximum and a minimum possible value for A. If the response functions $f_0(y)$ and $f_1(y)$ overlap, which is the general and expected case, then these limits on the value of A have important consequences.

The possible values A for a fixed value of a, are traversed as the region R is varied. It is clear that to make A as large as possible for a given value of a, R should contain those points whose contribution to A would be as small as possible (i.e., the ratio f_1/f_0 small) while the complement contains those points with large values of this ratio. This is the familiar likelihood ratio threshold rule.

If a goes to 1, then A goes to zero. Also, if A is 1, then a must vanish. This follows trivially from the unit normalization of the response functions.

Finally, note that an ideal detector with perfect discrimination has response functions that satisfy $f_0(y) \times f_1(y) = 0$ for *all* y. In this case, the values of a and A are independent. We will return to this limiting case shortly.

2. The Cost Function

The expected value of the cost function is

$$\langle C \rangle = \sum_{i,j} C(m_i, H_j) p_{rior}(H_j) p(m_i|H_j) , \tag{10}$$

where $p(m|H)$ is directly expressed in terms of the detector properties, and we assume that the signals received by the detectors are stochastically independent:

$$p(m_i|H_j) = \sum_{u_a, u_b} p(m_i|u_a, u_b) p(u_a|H_j)\, p(u_b|H_j) . \tag{11}$$

Using the explicit form of the cost matrix, Eq. (2), (10) can be expressed as

$$\langle C \rangle = w_{01}\; p(m_0|H_1)\, p_1 + w_{10}\; p(m_1|H_0)\, p_0 + u_0\, p_0 + u_1\, p_1 . \tag{12}$$

Additive constants do not matter in the minimization; the last two terms are fixed, and are the cost for an ideal system. For such a system with perfect discrimination, the off-diagonal probabilities $p(m_0|H_1)$ and $p(m_1|H_0)$ both vanish since $A = 1 - a$. The cost must be a minimum:

$$\langle C \rangle_{min} = u_0\, p_0 + u_1\, p_1 . \tag{13}$$

The quantity that we want to minimize is the additional cost due to imperfections in the system; this has the form

$$\begin{aligned} \langle \delta C \rangle &= \langle C \rangle - \langle C \rangle_{min} \\ &= w_{01}\; p(m_0|H_1)\, p_1 + w_{10}\; p(m_1|H_0)\, p_0 . \end{aligned} \tag{14}$$

Note that the position of the minimum will depend on the ratio

$$W = \frac{w_{10}\, p_0}{w_{01}\, p_1} , \tag{15}$$

which is the relative expected cost of being wrong if the state of the environment is H_0 (and responding with m_1) compared to the cost of being wrong if it is H_1 (and responding with m_0). The magnitude of the minimum cost will depend multiplicatively on the factor $w_{01}\, p_1$.

It is convenient to rewrite the cost function as

$$J \equiv \langle \delta C \rangle / (w_{01}\, p_1) , \tag{16}$$

or

$$J = p(m_0|H_1) + W[1 - p(m_0|H_0)] . \tag{17}$$

The minimization of the expected value of the cost is equivalent to minimizing J.

Some interesting limits on J can now be determined. The perfect detector has $J = 0$. It is amusing to note that a detector that is *always* wrong has $J = 1 + W$. (One would then use such a detector "backward.") A more interesting case follows from noting that if W is sufficiently small, i.e., the cost w_{10} (of erroneously choosing m_1) is small, then a good strategy is to always choose m_1. This implies that $p(m_0|H_1) = p(m_0|H_0) = 0$, and $J = W$ (and $g = d = 0$). If, on the other hand, W is larger than 1, then one wants to always choose m_0; in this limit, $J = 1$ (and $g = d = 1$). The final cost for this limiting case may be expressed in terms of the step function $\theta(x)$ ($\theta(x) = 1, x > 0, \theta(x) = 0, x < 0$)

$$\begin{aligned} J_{max} &= W\theta(1 - W) + \theta(W - 1) \\ g = d &= \theta(W - 1) . \end{aligned} \tag{18}$$

This result arises in another way. If the response functions are the same, $f_0(y) = f_1(y)$, then no discrimination is possible, and we find $A = 1 - a$. Using this relation in the probabilities, we find the above result by choosing the obvious optimum.

The general optimization problem consists of choosing the design parameters so the expected cost lies as far below J_{max} as possible and as close to the ideal case, $J = 0$, as possible. We now turn to a general discussion of the problem of finding extrema when the constraints define a connected subset of the real line for each variable.

3. General Minimization with Inequalities

Using the form of the probabilities defined in Eqs. (6) and (7), one finds the explicit expressions

$$\begin{aligned} p(m_0|H_1) &= 1 - p(m_1|H_1) , \\ &= (1 - A)(1 - B) + g\,(1 - A)B + d\,A(1 - B) , \end{aligned} \tag{19}$$

and

$$\begin{aligned} p(m_0|H_0) &= 1 - p(m_1|H_0) , \\ &= ab + g\,a(1 - b) + d\,(1 - a)b . \end{aligned} \tag{20}$$

Using Eqs. (19) and (20), the minimization problem can be re-cast explicitly as

$$J = W + [S + g\,T + d\,U] , \tag{21}$$

where

$$\begin{aligned} S &= (1 - A)(1 - B) - Wab \\ T &= (1 - A)B - Wa(1 - b) \\ U &= A(1 - B) - W(1 - a)b , \end{aligned} \tag{22}$$

and all the variables must satisfy inequality constraints. A complete mathematical treatment for problems of this type can be found in the excellent book by Hestenes.[8] A reference that discusses such variational problems in a language perhaps more familiar to physicists and engineers is available.[9]

To *minimize* J, in the case that the variables g and d occur linearly in J, but have a restricted range from zero to one, it is convenient to form the variational functional J_{var}, where

$$J_{var} = J - \gamma g(1 - g) - \delta d(1 - d) . \tag{23}$$

The optimum will be a saddle point in (γ, δ) versus (g, d). In this case J is a linear function of g and d, hence the extrema will occur at the endpoints. The Lagrange inequality multipliers γ and δ must be zero if their associated variable g or d is inside the allowed range, and non-negative if they are on the boundary.[9] As usual, the derivative with respect to g must vanish at the minimum and this yields the condition

$$0 = T - \gamma(1 - 2g) . \tag{24}$$

This takes the place of paired Kuhn–Tucker conditions for $g \geq 0$ and $g \leq 1$. If T is nonzero, which is the typical case, then the minimum must be on the boundary (γ cannot be zero). If T is positive, then g vanishes; if negative, then g is unity. A similar argument holds for d and U. The result can be expressed as

$$\begin{aligned} g &= \theta(-T) \\ d &= \theta(-U) , \end{aligned} \tag{25}$$

and the minimum of J_{var} becomes

$$J_m = W + [S + T\,\theta(-T) + U\,\theta(-U)] . \tag{26}$$

Note that if T or U vanish, there is no uncertainty in the minimum of J, even though g and d are not determined.

The variables left to consider are a, A, and b, B. Each of these variables has a restricted range, so inequality multipliers will again be used. As was noted before, the possible values of A are limited by the form of the response function and the value of a. This can be expressed as the statement that for any choice of the region R, with a given by (8), one must have

$$A_{min}(a) \leq A(R) \leq A_{max}(a) . \tag{27}$$

Of course, similar restrictions apply to B.

These inequalities can be treated as above. Write the variational functional in the form

$$J_{var} = J_m - F - f , \tag{28}$$

where

$$F \equiv \alpha_A\,(A - A_{min})(A_{max} - A) + \beta_B\,(B - B_{min})(B_{max} - B) , \tag{29}$$

and

$$f \equiv \alpha\,a(1 - a) + \beta\,b(1 - b) . \tag{30}$$

Again, the Lagrange inequality multipliers $\alpha_A, \beta_B, \alpha$ and β must be zero if their associated variable is inside the allowed range and non-negative if they are on the boundary.

Now the variation with respect to A yields

$$2\,\alpha_A\,(A - A_{bar}) = -\frac{\partial J_m}{\partial A} , \tag{31}$$

where $A_{bar} \equiv (A_{min} + A_{max})/2$.

It is a straightforward task, though somewhat tedious, to discuss the general case. First note that the above equation becomes

$$2\,\alpha_A\,(A - A_{bar}) = +(1 - B) + B\,\theta(-T) - (1 - B)\,\theta(-U) . \tag{32}$$

Since the right-hand side is never negative, α_A cannot vanish, and hence A must be at its boundary. Since α_A must also be non-negative, it follows that A must be above A_{bar}. Repeating the same argument for B we find that

$$\begin{aligned} A &= A_{max}(a) , \\ B &= B_{max}(b) . \end{aligned} \tag{33}$$

These are computable functions of a and b given the response functions of the detectors. They correspond to the so-called Receiver Operating Characteristic used in several of the papers cited above. We shall term these functions the DOC, or Detector Operating Characteristic, and they will play a fundamental role in our analysis.

The next stage is to vary a and b within their allowed range to achieve the overall minimum. One can anticipate that there may be symmetric ($a = b$) and nonsymmetric minima; which particular one is the global minima must be determined from a more detailed examination using the explicit forms for the response functions. This will be carried out in the explicit examples discussed in the next section. First let us discuss the boundary behavior in a and b.

Double boundary: The boundary region in which both variables are at their limits consists of four terms. They will be denoted by $L(a,b)$, where a and b can take on the values zero or one.

L(0,0): For this case, $A = B = 1$, and $J_m = W$ for all W.

$\underline{L(1,0) \text{ and } L(0,1)}$: For these cases, $A = 0\,B = 1$, or the reverse, and $[S = 0 = U, T = 1 - W \text{ and } g = \theta(W-1)]$

$$J_m \;=\; W + (1-W)\,\theta(W-1)\;, \tag{34}$$

the J_{max} discussed earlier.

$\underline{L(1,1)}$: For this limit, $A = B = 0$,

$$J_m \;=\; 1\;. \tag{35}$$

Therefore, the minimum of J_m on this double boundary is *always* given by Eq. (34) which amounts to setting the detectors to always signal oppositely. Now let us turn to the single variable boundaries.

Single boundary: This boundary region is symmetric in both variables and hence we need only treat the case in which b is at its limits while a is in the interior. The reversed situation will yield the same minima. These will be denoted as:

$\underline{L(a,0)}$: For this case, $B = 1$, and $S = U = 0$. The quantity T is not zero, with

$$\begin{aligned} J_m &= W + T\,\theta(-T) \\ T &= 1 - A - Wa\;. \end{aligned} \tag{36}$$

As noted, A should be equal to its maximum value for a fixed value of a in order to achieve the minimum value of T, as was shown earlier. The limit cases of $a = 0$ and $a = 1$ are on the double boundary. Any minimum for a in the interior must satisfy

$$\begin{aligned} \frac{\partial T(a)}{\partial a} &= \frac{\partial A_{max}(a)}{\partial a} - W \\ &= 0\;. \end{aligned} \tag{37}$$

Since $A_{max}(a)$ is a decreasing function of a, there will in general be a solution in this region if W is in an appropriate range. This could yield a smaller minimum than that given by the double boundary result, Eq. (34); however, for this case, we have $g = 1$ and d is not determined, but its value does not matter since $U = 0$ and one can arbitrarily choose $d = 1$ also.

$\underline{L(a,1)}$: For this situation, $B = 0 = T$ and $S = 1 - A - Wa$, with $U = 1 - W - S$. If U is positive, then $J_m = 1$, while if it is negative, then $J_m = W + S$. Both these cases have arisen before, and there are no new minima of J_m.

Interior: In the interior region, the inequality multipliers must vanish and the standard variational equations become symmetric in form. One can safely assume that there will be minima in this region, but whether any is the global minimum requires detailed study. Note that generally there will be (local) minima with T(and/or U) both positive and negative with the corresponding limiting values of g (and d). Since this is a standard well-discussed variational problem, further general treatment here is not necessary. Let us now turn to a exhaustive discussion of some explicit examples.

4. Exponential Response Functions

Consider a detector with response functions given by

$$\begin{aligned} f_0 &= n\lambda\,\exp\{-n\lambda y\} \\ f_1 &= \lambda\,\exp\{-\lambda y\}\;. \end{aligned} \tag{38}$$

We will assume that n is greater than one without any loss of generality, so that the likelihood ratio (f_1/f_0) is less than one for $y \leq z$, where $\lambda z = (\ell n\, n)/(n-1)$. Using the above argument, to achieve the extrema of A for these monotonic response functions, the region R must be either the range below or the range above some point x whose value will be determined by the optimization process.[10]

Therefore, it is easy to see that there are two cases to discuss:

	Case I	*Case II*
R	$0 \leq y \leq x$	$x \leq y \leq \infty$
$\bar{R}$	$x \leq y \leq \infty$	$0 \leq y \leq x$
a	$1 - \exp\{-n\lambda x\}$	$\exp\{-n\lambda x\}$
A	$\exp\{-\lambda x\}$	$1 - \exp\{-\lambda x\}$
A	$(1-a)^{1/n}$	$1 - a^{1/n}$.

Thus, A must lie in the region

$$1 - a^{1/n} \leq A \leq (1-a)^{1/n}\;, \tag{39}$$

and its position in this interval is determined by the particular choice of the region R. Similar relations hold for b and B.

As an example, consider the case $n = 2$, and then the feasible region for A as a function of a is labeled F in the graph shown in Fig. 1.

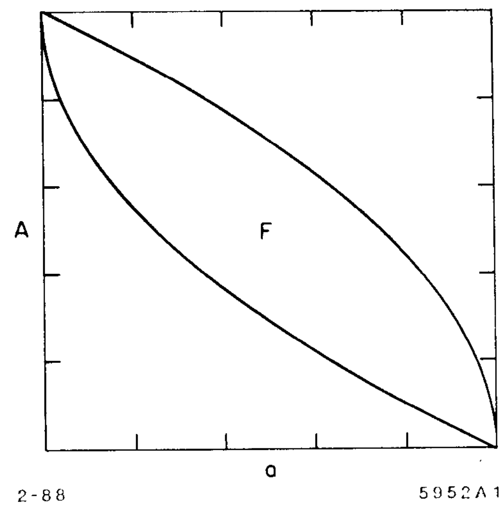

Fig. 1. The allowed region of A is plotted for $n = 2$ as a function of a and labeled F.

Figure 2 shows the graph for the value $n = 4$.

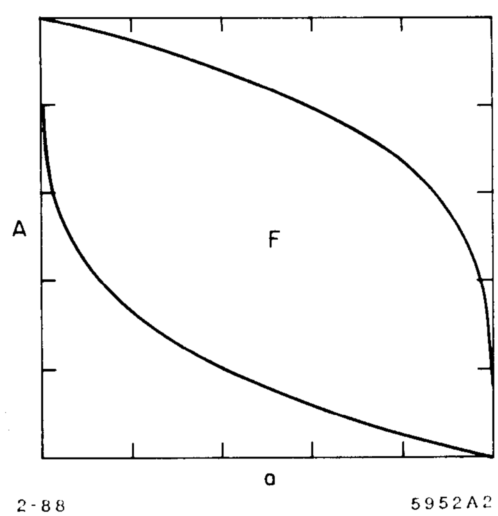

Fig. 2. Same as Fig. 1 but with $n = 4$.

We see that as n increases, the allowed region increases to eventually include all values of A between zero and one.

4.1 Explicit Minimization

Using the general results derived in the previous chapter, we have A and B at their maximum allowed values:

$$\begin{aligned} A &= (1-a)^{1/n} \\ B &= (1-b)^{1/n}\;. \end{aligned} \tag{40}$$

Varying J_m with respect to a and b and introducing inequality multipliers α and β to keep these variables between zero and one, we find the conditions

$$\begin{aligned} n\alpha(1-2a) &= A^{1-n}(1-B) - nWb \\ n\beta(1-2b) &= B^{1-n}(1-A) - nWa \end{aligned} \tag{41}$$

whose solution should contain all relevant minima. Let us examine the boundary and interior minima in that order. Recall that $n \geq 1$ in the following discussion, and we have assumed for the moment that T and U are positive. This will be proven shortly for our solutions.

Boundary: The double boundary region has been discussed in general and the result is a minimum of the form ($a = 0, b = 1$ or $a = 1, b = 0$)

$$J_m = W + (1-W)\,\theta(1-W) \equiv J_{max} . \tag{42}$$

L(a,0): For this single boundary problem, the task is to find the minimum of T, where $T = 1 - A_{max} - Wa$. The result is with $p = 1/(n-1)$

$$\begin{aligned} A_0 &= (\frac{1}{nW})^p \\ a_0 &= 1 - \frac{A_0}{nW} . \end{aligned} \tag{43}$$

For A_0 to be less than one, $nW \geq 1$. The value of T at this minimum is negative, and

$$J_m(Bnd) = 1 - \frac{n-1}{n} A_0 . \tag{44}$$

If nW is slightly larger than one, $nW = 1 + \epsilon$, then it is easy to see that to lowest order

$$J_m(Bnd) \sim J_{max} - \frac{\epsilon^2}{2(n-1)} . \tag{45}$$

L(a,1): For this case, $B = 0 = T$ and $U = 1 - W - S$. If U is negative, then the minimum of J_m is 1. If it is positive, then S must be minimized, and this is just the problem discussed above.

In summary, J_m has a minimum on the boundary given by Eq. (42) or Eq. (44), depending on the value of nW.

Interior: In the interior region, the inequality multipliers α and β must vanish and Eqs. (41) become symmetric in form. Thus there is a symmetric solution with $a = b$ and $A = B$. Unsymmetric solutions will be searched for later. In the symmetric case, the equation for the optimal probability A is

$$nWA^{n-1}(1 - A^n) = (1 - A) , \tag{46}$$

which does not have an analytic solution for general n. The limiting behavior of the solution is easily extracted. For large W, A approaches zero, and a approaches one with the behavior [recall that $p = 1/(n-1)$]

$$\begin{aligned} A &\simeq \left(\frac{1}{nW}\right)^p + \dots , \\ a &\simeq 1 - \left(\frac{1}{nW}\right)^{1+p} + \dots . \end{aligned} \tag{47}$$

This is similar to one of the boundary solutions. The minimum of J in this limit has the form

$$J \simeq 1 - 2\frac{n}{n-1}A + \dots . \tag{48}$$

Let us now discuss small values of W. Note that as W decreases, a decreases. The value of W where a vanishes is

$$W(a = 0) = 1/n^2 . \tag{49}$$

For values of W smaller than this value, there is no interior symmetric solution. At this critical value, $S = 0$. Finally, note that for this symmetric solution, $T = U$, and using the above equations,

$$T = (n-1)WaA^n , \tag{50}$$

which is positive definite. Therefore, the T and U terms do not contribute to this minimum because $g = d = 0$.

Using the equation for A, we find at the minimum

$$J_m(Int) = W + (1-A)^2 - W(1-A^n)^2 . \tag{51}$$

This is smaller than the minimum arising from the boundary.

To see this, study the difference of Eq. (44) and Eq. (51) for sufficiently large W (so that the former exists). If W is eliminated between Eqs. (46) and (43), the result is

$$A_0 = A\,(\frac{1-A^n}{1-A})^p , \tag{52}$$

which shows that $A_0 = A_0(A) \geq A$. The difference becomes

$$\begin{aligned} J_m(Bnd) - J_m(Int) &= [1-(1-A)^2] - (\frac{1}{n})A_0^{1-n} \\ &\quad [1-(1-A^n)^2] - \frac{n-1}{n}A_0 . \end{aligned} \tag{53}$$

For large W this difference approaches zero as $(n-1)/(n^2W)$. For all values of nW larger than one it is a simple matter to show that it is positive (a numerical proof is easiest).

Some sample numerical results are:

	$n = 2$		$n = 4$	
n^2W	a	J-W	a	J-W
0 − 1	0.00	-0.000	0.00	-0.00
1.0	0.00	-0.000	0.00	-0.00
1.1	0.120	-0.0001	0.082	-0.0000
1.2	0.218	-0.0009	0.150	-0.0001
1.3	0.299	-0.0026	0.210	-0.0003
1.4	0.367	-0.0054	0.261	-0.0007
1.6	0.475	-0.0144	0.346	-0.0018
1.8	0.556	-0.0278	0.413	-0.0036
2.0	0.618	-0.0451	0.467	-0.0061
3.0	0.791	-0.175	0.636	-0.0260
4.0	0.866	-0.348	0.725	-0.091
6.0	0.930	-0.757	0.815	-0.131
8.0	0.957	-1.203	0.862	-0.219
10.0	0.971	-1.669	0.890	-0.315
12.0	0.979	-2.144	0.909	-0.417
16.0	0.987	-3.117	0.933	-0.629

Recall that $g = d = 0$ for this global minimum. Therefore, if either detector signals 1, one should make the choice m_1 for any value of W.

Perhaps it is more understandable to present this data in another format:

		$n = 2$			$n = 4$	
W	a	J	J/J_{max}	a	J	J/J_{max}
0.25	0.250	0.250	1.0	0.725	0.195	.78
0.4	0.475	0.386	.97	0.827	0.252	.63
0.5	0.618	0.455	.91	0.862	0.281	.56
0.75	0.791	0.575	.77	0.909	0.333	.44
1.0	0.866	0.652	.65	0.933	0.371	.37
1.5	0.930	0.743	.74	0.957	0.423	.42
2.0	0.957	0.797	.80	0.969	0.459	.46
2.5	0.971	0.831	.83	0.976	0.486	.49
3.0	0.979	0.856	.86	0.980	0.508	.51
4.0	0.987	0.883	.88	0.986	0.541	.54
5.0	0.992	0.909	.91	0.989	0.567	.57
6.0	0.994	0.923	.92	0.991	0.586	.59
8.0	0.997	0.941	.94	0.994	0.616	.62
10.0	0.998	0.952	.95	0.995	0.639	.64

The column labeled J/J_{max} gives the ratio of the minimum J to the quantity J_{max} defined in (18). Again, for this global minimum, $g = d = 0$.

4.2 Global Minimum

As a check that the symmetric minimum is indeed the global minimum, we have evaluated J throughout the allowed region of the six variables g, d and a, A, b, B. We could find no point where J was below the value at the symmetric minimum given above.

Note that as a function of $W \equiv w_{10}\, p_0 / w_{01}\, p_1$, the largest fractional improvement in cost is achievable when $W = 1$. This is precisely the case in which the *prior* choice of action is a matter of indifference, that is,

$$u_0\, p_0 + u_1\, p_1 + w_{01}\, p_1 \;=\; u_0\, p_0 + w_{10}\, p_0 + u_1\, p_1\ . \tag{54}$$

This is intuitively reasonable, as one expects the information from the sensor to be the most valuable in this case.

5. Invariant Imbedding

We now consider a detector whose response functions allow superior discrimination between the two possible states of the environment and contains the previous example as a special case. The general form that allows a smooth limit back to the previous model is

$$\begin{aligned} f_0 &= n\lambda \exp\{-n\lambda y\} \\ f_1 &= \frac{n+1}{n+1-z}\lambda \exp\{-\lambda y\}\, [1 - z \exp\{-n\lambda y\}]\ . \end{aligned} \tag{55}$$

For values of z near 1 this allows the improved separation between f_0 and f_1 since the former is large at $y = 0$ while the latter is small there. On the other hand, for z equal to zero, this is the model of the previous section.

It will again be assumed that n is greater than 1, and proceeding as before we find:

	Case I	*Case II*
R	$0 \le y \le x$	$x \le y \le \infty$
$\bar{R}$	$x \le y \le \infty$	$0 \le y \le x$
a	$1 - \exp\{-n\lambda x\}$	$\exp\{-n\lambda x\}$
A	$\frac{1}{n+1-z}\exp\{-\lambda x\} \times [n+1-z\exp\{-n\lambda x\}]$	$1 - \frac{1}{n+1-z}\exp\{-\lambda x\} \times [n+1-z\exp\{-n\lambda x\}]$
A	$(1-a)^{1/n}\left[1 + \frac{za}{n+1-z}\right]$	$1 - a^{1/n}\left[1 + \frac{z(1-a)}{n+1-z}\right]$.

Thus, A must lie in the region

$$1 - a^{1/n}\left[1 + \frac{z(1-a)}{n+1-z}\right] \le A \le (1-a)^{1/n}\left[1 + \frac{za}{n+1-z}\right]\ . \tag{56}$$

Its position in this interval is determined by the particular choice of the region R. Similar relations hold for b and B. Note that the allowed region of A increases as z increases from zero to one.

To provide maximum contrast with the previous model we will present data for the value $z = 1$. For this case, the interior symmetric minimum exists for all W values. An interesting new behavior is found in this model for small enough W and n; the minimum cost occurs for $g = d = 0$, as before, but as W increases, these design parameters flip to $g = d = 1$. The value of a at the minimum jumps *discontinuously*.

	$n=2$			$n=4$		
W	a	J	J/J_{max}	a	J	J/J_{max}
0.1	0.46	0.088	.88	0.70	0.068	.68
0.2	0.58	0.160	.80	0.80	0.111	.56
0.25	0.63	0.191	.764	0.82	0.128	.51
0.26	0.64	0.197	.758	0.828	0.131	.504
0.27	0.24	0.203	.752	0.832	0.134	.496
0.3	0.26	0.22	.733	0.84	0.143	.48
0.5	0.35	0.31	.62	0.89	0.191	.38
1.0	0.50	0.47	.47	0.94	0.268	.27
1.5	0.60	0.56	.56	0.96	0.316	.32
2.0	0.68	0.63	.63	0.97	0.35	.35
4.0	0.79	0.77	.77	0.985	0.44	.44
6.0	0.85	0.83	.83	0.990	0.49	.49
8.0	0.88	0.87	.87	0.993	0.52	.52
10.0	0.90	0.89	.89	0.994	0.55	.55

The columns are the same as in the previous table. Note that the parameters for the minimum (for $n = 2$) shows a definite jump as W passes through the value ≈ 0.265. At this point, the optimum values of g and d change from zero to one; in fact, we find that $g = d = \theta(W - W_0)$, where $W_0 \approx 0.265$.

On the other hand, for $n = 4$, the quantities T and U are always positive, so that $g = d = 0$. There does not appear to be a discontinuity in a.

At the discontinuity, the *cost* varies smoothly. This is intuitively reasonable, since cost is, ultimately, determined by the position of some tangent hyperplane, along a normal to the feasible region, which is connected. However, the jump in design parameters could have serious consequences because a small variation in the (frequently subjective) data summarized by the parameter W could require a complete change of the system parameters g and d. This phenomena has important implications for the design of constant false alarm rate systems, which will be discussed elsewhere.[11]

6. A Step Function Example

We now consider a detector whose response functions, in a certain limit, allow a clean discrimination between the two possible states of the environment. In that limit, $A \to 1$ does not force a to zero. We will assume simple "square" response functions for ease of presentation. The response functions are chosen to be zero for $y > 3$ and, of course, normalized.

Proceeding as before we find for Case I, $0 < y < x$:

For the range	$0 \le x \le 1$	$1 \le x \le 2$	$2 \le x \le 3$
$f_0(x)$	$(1-\lambda_0)$	λ_0	0
$f_1(x)$	0	λ_1	$(1-\lambda_1)$
a	$(1-\lambda_0)x$	$1+\lambda_0(x-2)$	1
A	1	$1-\lambda_1(x-1)$	$(1-\lambda_1)(3-x)$.

A similar table can be evaluated for Case II, $x < y < 3$. If either λ_0 or λ_1 vanish, then this describes an ideal detector system.

We need the value of A_{max} for a fixed value of a which is

$$A_{max} = 1 - \frac{\lambda_1}{\lambda_0}(a - 1 + \lambda_0)\theta(a - 1 + \lambda_0)\ , \tag{57}$$

while the minimum value is

$$A_{min} = \frac{\lambda_1}{\lambda_0}(\lambda_0 - a)\theta(\lambda_0 - a)\ . \tag{58}$$

The feasible region for A as a function of a , labeled F in the graph (Fig. 3), is bounded by straight lines:

In the limit that either λ_0 or λ_1 vanishes, the allowed region for A covers the unit square.

It is a simple matter to analyze this problem for the minimum J corresponding to the maximum allowed A and B values as given above. Consider the cases:

1. $A = B = 1$, and $a = b = (1 - \lambda_0)$. For these values, T and U are negative and $J = W\lambda_0^2$.
2. $A = B = (1 - \lambda_1)$ and $a = b = 1$. For this case T and U are now positive and $J = \lambda_1^2$.

Thus the final result can be expressed as

$$\begin{aligned} J &= \min[W\lambda_0^2\ ,\ \lambda_1^2] \\ g = d &= \theta(\lambda_1^2 - \lambda_0^2 W) \\ a = b &= 1 - \lambda_0\, \theta\, (\lambda_1^2 - \lambda_0^2 W) \\ A = B &= 1 - \lambda_1\, \theta\, (\lambda_0^2 W - \lambda_1^2)\ . \end{aligned} \tag{59}$$

The limit of perfect discrimination, λ_0 and/or λ_1 going to zero, can be easily discussed from the above results.

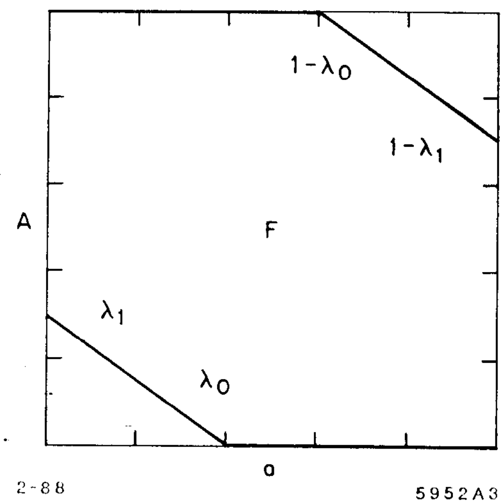

Fig. 3. The allowed region F for the parameter A as a function of a using the discrete three-step model.

The proof that the above minimum is indeed a global minimum follows simply by letting a and b deviate from the above values while keeping A and B as close to their optimum values as allowed by the constraint. For W small enough we have:

$$\begin{aligned} a &= b = 1 - \lambda_0(1+\epsilon) \\ A &= B = 1 \\ (S+T+U)_{min} &\equiv -W\,(1-\lambda_0{}^2)\ . \end{aligned} \tag{60}$$

It now follows for any positive ϵ that

$$(S+T+U)-(S+T+U)_{min} \;=\; W\,\lambda_0{}^2[(1+\epsilon)^2-1] \geq 0\ . \tag{61}$$

For a and b larger than their optimum values, the constraints on A and B come into play and

$$\begin{aligned} a &= b = 1 - \lambda_0(1-\epsilon) \\ A &= B = 1 - \lambda_1\epsilon \end{aligned} \tag{62}$$

and we find

$$\begin{aligned} (S+T+U) - (S+T+U)_{min} &= \\ 2\epsilon\lambda_1(1-\lambda_1) + (\lambda_1{}^2 - W\lambda_0{}^2)\,[1-(1-a)^2] & \\ \geq 0 & \end{aligned} \tag{63}$$

if $(\lambda_1{}^2 - W\lambda_0{}^2)$ is positive (and if ϵ is positive, of course).

When W grows so that $(\lambda_1{}^2 - W\lambda_0{}^2)$ becomes negative, one should repeat the above procedure around the values $a = b = 1$ and $A = B = 1 - \lambda_1$ to prove the global nature of the minimum in this region. Alternatively, one may argue that the feasible region is defined, in this case, by hyperplanes, so that the minimum must occur at a vertex, as given above.

7. Summary

We find that the problem of optimal design, with fusion and detector tuning, is difficult but tractable. Our simple examples yield some insight into how the best achievable cost varies between its bounds and how that best cost depends on the prior distribution and the cost function itself.

By utilizing the technique of invariant imbedding, that is by considering a general class of response functions that contain the exponential response model as a particular case, we can trace a discontinuous change in design parameters, even though the optimum cost varies smoothly. We cannot yet give a detailed explanation of the critical value at which this jump occurs. The third model studied also has such a discontinuity, and permitted a continuous transition to the state of complete information (perfect discrimination). In this case, the cost depends on the degree of ambiguity in a quadratic manner.

Finally, in all cases, we found that the best cost is achieved with a symmetric choice of parameters for the individual detectors. We do not yet have a general characterization of response functions for which this is always the case regardless of costs and prior probabilities.

The problem considered here is not only of theoretical interest, but has many practical applications ranging from optimal design of complex particle detector systems to the design of seismic and warning systems.

Acknowledgements

One of the authors (PBK) acknowledges with thanks the hospitality of the Department of Operations Research, Weatherheading School of Management, Case–Western Reserve University, and its chairmen Arnold Reisman and Hamilton Emmons. We both thank Dr. Rabindar Madan of the Office of Naval Research for calling this problem to our attention.

REFERENCES

1. R. Srinivasan, *Distributed Radar Detection Theory,* IEEE Proc. F (GB) **133**, Vol. 1, pp. 55-60 (1986). See also, R. Srinivasan, Signal Processing (Netherlands) **11,** Vol. 4, pp. 319–327 (1986).
2. R. Srinivasan, P. Sharma and V. Malik, *Distributed Detection of Swerling Targets,* IEEE Proc. (GB) **133**, Vol. 7, pp. 624–629 (1986).
3. R. R. Tenney and N. R. Sandell, *Detection With Distributed Sensors,* IEEE Transactions on Aerospace and Electronic Systems **17,** Vol. 4, pp. 501–509 (1981), and **17**, Vol. 5, p. 736 (1981).
4. L. K. Ekchian and R. R. Tenney, *Recursive Solution of Distributed Detection/Communication Problems,* Proc. of the 1983 American Control Conference, Vol. 3, San Francisco, CA, pp. 1338–1339.
5. Z. Chair and P. K. Varshney, *Optimal Data Fusion in Multiple Detector Systems,* IEEE Transactions on Aerospace and Electronic Systems **22,** Vol. 1, pp. 98–101 (1986).
6. A. R. Reibman and L. W. Nolte, *Optimal Detection and Performance of Distributed Sensor Systems,* IEEE Transactions on Aerospace and Electronic Systems **23,** Vol. 1, pp. 24–30 (1987).
7. F. A. Sadjadi, *Hypothesis Testing in a Distributed Environment,* IEEE Transactions on Aerospace and Electronic Systems **22,** Vol. 2, pp. 134–137 (1986).
8. M. R. Hestenes, *Calculus of Variations and Optimal Control Theory* (New York, Wiley, 1966).
9. M. Einhorn and R. Blankenbecler, *Bounds on Scattering Amplitudes,* Annals of Physics **67**, 480 (1971).
10. The key point here is that the likelihood ratio is a monotonic function of x, so R is an interval including one endpoint.
11. M. Cherikh, *Design Discontinuities in Distributed Sensors: Implications for CFAR Systems,* in preparation.