January 1990 LIDS–P–1928

On Optimal Distributed Decision Architectures In A Hypothesis Testing Environment [1, 2]

Jason D. Papastavrou and Michael Athans [3]

Abstract. We consider the distributed detection problem, in which a set of decision makers (DMs) receive observations of the environment and transmit finite-valued messages to other DMs according to prespecified communication protocols. A designated primary DM makes the final decision on one out of two alternative hypotheses. All DMs make decisions, in order to maximize a measure of organizational performance. Given the DMs and the communication resources, the problem is to find an architecture for the organization which remains optimal for a variety of operating conditions (if it exists). We show that even for very small organizations this problem is quite complex, because the optimal architecture depends on variables external to the team like the prior probabilities of the hypotheses and the misclassification costs, so that global conslusions on optimal organizational structures cannot be drawn. We thus also consider suboptimal solutions and obtain bounds on their performance.

---

[1] This research was supported by the National Science Foundation under grant NSF/IRI–8902755 (under a subcontract from the University of Connecticut) and by the Office of Naval Research under grant ONR/N00014–84–K–0519.

[2] This paper has been submitted for publication in the IEEE Transactions on Automatic Control.

[3] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

1. INTRODUCTION AND PROBLEM DEFINITION

Scientists of very different disciplines have joined forces in order to attempt to model decision making by humans in cooperative organizations (teams). Systems theorists try to obtain *normative/ prescriptive* models which demonstrate how decisions should be made; the conclusions obtained from these models form hypotheses the validity of which should be investigated in practice. Cognitive psychologists gather data and develop *empirical/descriptive* models which they believe demonstrate how decisions are made; but, these models are "data dependent" and usually demonstrate how certain decisions were made by certain DMs under certain conditions, rather than how decisions are made. Research efforts have tried to combine the fruits of both approaches into *normative/descriptive* models which will be more realistic and accurate (for example, [GS66], [P85] and [K89]). These models should assist in improving decision making by demonstrating how decisions are indeed made.

It is worthwhile to note that although systems engineers and cognitive psychologists tackle the problem from diametrically opposed perspectives, they also share many common analytical tools in their modeling. For example, consider the *Receiver Operating Characteristic* (ROC) curve which is the cornerstone of mathematical binary hypothesis testing, since it offers a complete description of the DMs [V68]. The ROC curve may look like an artificial mathematical creation, but there is evidence from psychologists to suggest that human DMs can also be characterized by such a curve; see the pioneering book by Green and Swets [GS66]. In fact, human ROC curves have been constructed experimentally and have been employed in decision analysis. We thus concentrate in our normative models realizing that they constitute an incomplete, yet integral, part of the effort of modeling decision making; we will try to derive conclusions which subsequently can be tested in practice to determine whether they can improve actual human decisions.

The problem of distributed decision making in a hypothesis testing environment has attracted considerable interest during the past decade. This framework was selected because it combines two desirable attributes; the mathematical problems are easy to describe so that researchers from diverse disciplines can understand the models and their conclusions; also, the problems have trivial centralized counterparts, so that all the difficulties arise because of the decentralization of the decision making process. On the other hand, these problems are also known to become computationally intractable (NP-hard) even for a small number of DMs and a small number of communication messages [TA85]. Thus, in order to overcome the limitations caused by the combinatorial complexity, it would be desirable to combine DMs into more compact and aggregated decision making units so as to obtain building blocks for larger organizations, which hopefully have tractable quantitative descriptions.

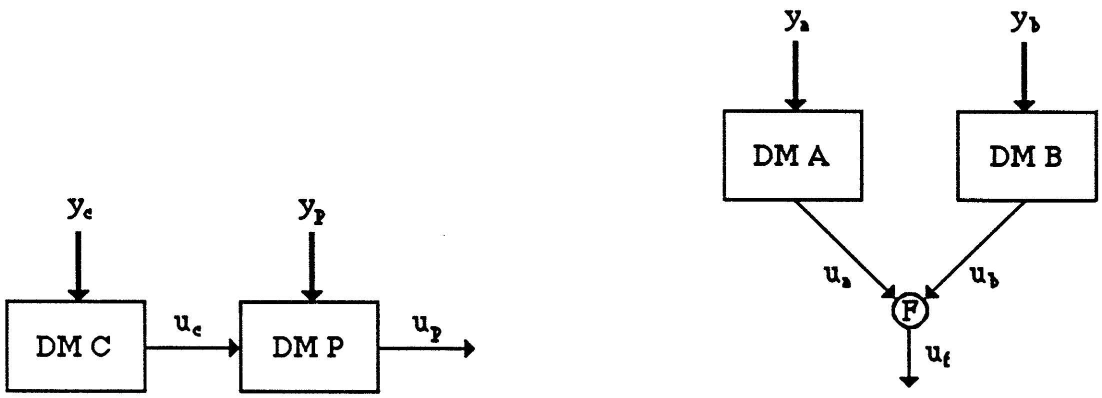
We examine problems of small organizations, that is cooperative organizations which consist of two or three DMs and perform binary hypothesis testing, because we want to keep the combinatorial complexity under control, so that the difficulties arise only from the intrinsic complexity of the distributed problems. We present different architectures for these organizations, analyze them in a quantitative manner and compare their performance. We investigate whether some "common sense" and "intuitively appealing" beliefs are indeed always true in this framework. As a concrete example, consider an organization consisting of two DMs, one better than the other. We determine whether, as has been suggested, the final team decision should be made by the better DM; it would have been desirable for organizational design to prove that this is always true, independent of parameters external to the team, such as the number and nature of communication messages available and/or the prior probabilities of the hypotheses. But, we show that the optimal assignment of the DM responsible for the final decision is depended on these external parameters, and that the better DM should not always make the team decision, although in many (but not all) situations this is optimal.

The Bayesian decentralized detection problem was first considered in [TS81], where the optimality of constant threshold strategies was established; this was formalized and generalized in [T89a]. Several generalizations of the basic detection model have appeared in [ET82], [S86], [CV86], [CV88], [HV88] and [TP89]. The parallel architecture with identical sensors has been analyzed in [R87], [RN87], [HV88] and in [T88], where the asymptotic results were established. The Neyman-Pearson formulation of similar problems is considered in [S86a], [HV86], [R87], [TV87] and [VT88], where different team architectures are compared; different team architectures, for the Bayesian case, are also compared in [E82], [RN87a] (numerically) and [PA88] (analytically). An excellent and thorough overview of the field was presented in [T89].

The distributed binary hypothesis testing model is defined as follows. There are two hypotheses H_0 and H_1 with known prior probabilities $P(H_0) > 0$ and $P(H_1) > 0$ respectively, and the team (organization) consists of $N \geq 2$ DMs. Let y_n, the observation of the nth DM, be a random variable taking values in a set $\boldsymbol{Y}_n$, $n = 1, ..., N$. We assume that the y_n's are conditionally independent given either hypothesis, with a known conditional distribution $P(y_n \mid H_j)$, $j = 0, 1$. We also assume that the communication protocols are given and known to all the team members. Let D_n be a positive integer, $n = 1, ..., N$. Each DM n evaluates a D_n-valued message $u_n \in \{1, ..., D_n\}$ as a function of its own observation and of some (possibly none) messages from other DMs; that is $u_n = \gamma_n(y_n, \boldsymbol{u}_n)$, where the measurable function γ_n: $\boldsymbol{Y}_n \times \boldsymbol{U}_n$ is the decision rule of DM n. The vector $\boldsymbol{u}_n$ ($\boldsymbol{u}_n \in \boldsymbol{U}_n$) consists of the messages transmited to DM n and the scalar decision u_n is the message transmitted to a single DM, according to the communication protocols. The decision of a designated DM, called the *primary* DM, is the final team decision and declares one of the hypotheses to be true. The objective is to choose the decision rules γ_n for the DMs ($n = 1, ..., N$), which minimize the probability of error of the team decision, taking into account different costs for hypothesis misclassifications; let $J(u, H)$ be the cost of the team deciding u when the true hypothesis is H [1,2] and define the decision threshold η, as follows:

$$\eta \equiv \frac{P(H_0)\,[J(1, H_0) - J(0, H_0)]}{P(H_1)\,[J(0, H_1) - J(1, H_1)]} \tag{1}$$

We begin in section 2 with the simplest non-trivial example and compare the performance of the two possible architectures for the two DM case; the tandem architecture (Figure 1(a)) and the parallel architecture (Figure 1(b)). In section 3, we analyze the team consisting of two DMs in tandem and examine the effects of variables external to the team on the optimal team configuration. In section 4, we discuss the team which consists of two DMs in parallel and pay special attention to the case where both DMs are identical. In section 5, we compare the architectures for the teams consisting of three DMs. Finally in section 6, we present our conclusions.



(a). Two DMs in *Tandem*

(b). Two DMs in *Parallel*

Figure 1. Architectures for Teams Consisting of Two DMs

2. TANDEM VERSUS PARALLEL ARCHITECTURE

Consider a team which consists of two DMs and performs binary hypothesis testing. There are two alternative architectures for this team. In the *tandem* architecture (Figure 1(a)) one DM, called the *consulting* DM, makes a decision based on its observation and transmits a binary message to the other DM, called the *primary* DM. Then, the primary DM makes the final team

[1] Throughout this discussion we assume that the cost function is such that it is more costly for the team to err than to be correct (i.e., $J(0, H_1) > J(1, H_1)$, $J(1, H_0) > J(0, H_0)$). This logical assumption is made in order to express the optimal decision rules in the convenient likelihood ratio form.

[2] As mentioned above, in our analysis we very often employ the ROC curve and its properties; for the reader's convenience a summary of these is included in appendix A.

decision based on its own observation and the message from the consulting DM. In the *parallel* architecture (Figure 1(b)) each DMs makes a decision and transmits a binary message to the fusion center which makes the final team decision so as to minimize the expected cost. The prior probabilities and the costs are assumed to be known by all the DMs in both architectures. In the following problem the best architecture is to be determined:

PROBLEM 1. Consider a team which consists of two DMs and performs binary hypothesis testing; determine which of the two possible architectures for the team, the tandem architecture (Figure 1(a)) or the parallel architecture (Figure 2(b)), is the dominant architecture; that is, the architecture that achieves superior performance for any prior probabilities and costs.

It is known that the tandem architecture is better than the parallel architecture. Several comparisons have appeared in the literature demonstrating that the tandem architecture achieves superior performance than the parallel architecture [ET82], [R87], [VT88]. We include a proof because it is more formal and because we are going to employ this result as a lemma in the proof of subsequent results.

LEMMA 1. *Consider a team consisting of two DMs. Then, the tandem architecture achieves at least as good performance as the parallel architecture.*

Proof. The proof is simple and straightforward. Consider a team consisting of two DMs in parallel and denote by $\Gamma^* = \{\gamma_a, \gamma_b, \gamma_t\}$ the set of the optimal decision rules for the two DMs and for the fusion center respectively. The optimal decision of each DM n depends exclusively on the observation y_n of the DM, for $n = a, b$; the decision of the fusion center depends on the decisions u_a and u_b of the two DMs.

Now consider the same two DMs in a tandem architecture; without loss of generality assume that DM b is the consulting DM. DM b can employ γ_b to make its decision based on its own observation. Moreover, DM a can employ γ_a and make a *preliminary* decision based on its own observation and then also employ γ_t to make the *team* decision based on DM b's message and on its own preliminary decision. The proposed decision rules $\bar{\gamma}_a$ and $\bar{\gamma}_b$ for the tandem architecture are thus defined by:

$$\bar{\gamma}_a(y_a, u_b) \equiv \gamma_t(\gamma_a(y_a), u_b) \qquad (2)$$

and:

$$\bar{\gamma}_b(y_b) \equiv \gamma_b(y_b) \qquad (3)$$

The proposed decision rules, though not optimal in general, enable the tandem team to always duplicate the optimal performance of the parallel team. This implies that the tandem architecture can achieve at least as good performance as the parallel. **Q.E.D.**

Note that the above result neither depends on the DMs involved, nor on which of the two DMs is the primary DM in the tandem architecture, nor on the prior probabilities and the costs. Furthermore, the result can be generalized for any number of messages which can be transmitted within a team, as long as the consulting DM in the tandem configuration is allowed to transmit to the primary DM the same number of messages, as it (the consulting DM) is allowed to transmit to the fusion center in the parallel configuration.

3. TWO DMs IN TANDEM

3.1. Configuration Comparisons

Since the tandem architecture is superior to the parallel architecture for teams of two DMs, it is worthwhile to analyze it further. Given two DMs, we would like to determine the optimal configuration for the tandem team (i.e., determine which DM should be made the primary one). If one DM is better than the other, it is intuitively appealing that the better DM be made the primary DM. Given two DMs one would expect to have the better DM make the team decision, independent of the prior probabilities and the cost assignments. If this was the case, then the optimal way of organizing two DMs would not change, say, as the prior probabilities of the underlying hypotheses vary. This has also been supported with explanations on data compression [E82] and with numerical results [R87]. But, it is not true in general; we show that the optimal configuration depends on the prior probabilities, on the cost assignments and, in a counterintuitive manner, on the number of messages which the consulting DM is allowed to transmit to the primary DM. The *necessary conditions* which characterize the optimal decision rules of the two DMs were obtained in [ET82] and are presented for completeness in appendix B.

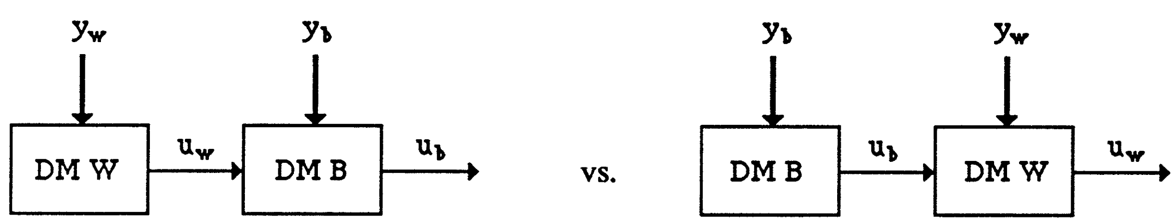


Figure 2. Different Configurations for the Two DM Tandem Team

Suppose that one of the DMs is *better* than the other, i.e., its ROC curve is higher than the ROC curve of the other DM. There exist two candidate configurations for the team; either make the better DM the primary DM, or, make the better DM the consulting DM (Figure 2). Recall that the primary DM makes the final team decision. Consider the following problem:

PROBLEM 2. Consider two DMs, one better than the other. Determine whether the optimal configuration of the tandem team which consists of these two DMs is independent of the external parameters of the problem (details of cost function, prior probabilities) which determine the value of the decision threshold η; that is,which of the two possible configurations (Figure 2) yields better performance than the other for all values of η.

The problem of the optimal configuration can be reduced to a simpler problem in which the worse DM will have a three piecewise linear ROC curve and the better DM will also have a piecewise linear ROC curve with at most four line segments[1]. The analysis of this restricted problem is presented in appendix C and yields the following *Example* 1. The ROC curves are presented in Figure 3 and the associated discrete probability density functions for the two DMs conditioned on the two hypotheses are presented in Figure 4. The team ROC curves are presented in Figure 5 for both configurations. It is interesting to note that for $\eta = 1.0$, performance is maximized by making DM B the consulting DM, while for $\eta = 0.4$, performance is maximized by making DM B the primary DM (Table 1). Thus, in this example, the optimal team architecture depends on the value η (i.e., the numerical values of the prior probabilities and the costs) [2].

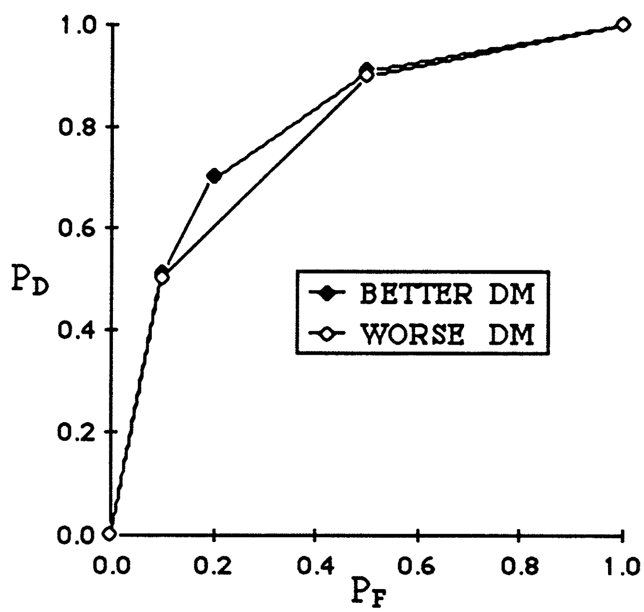
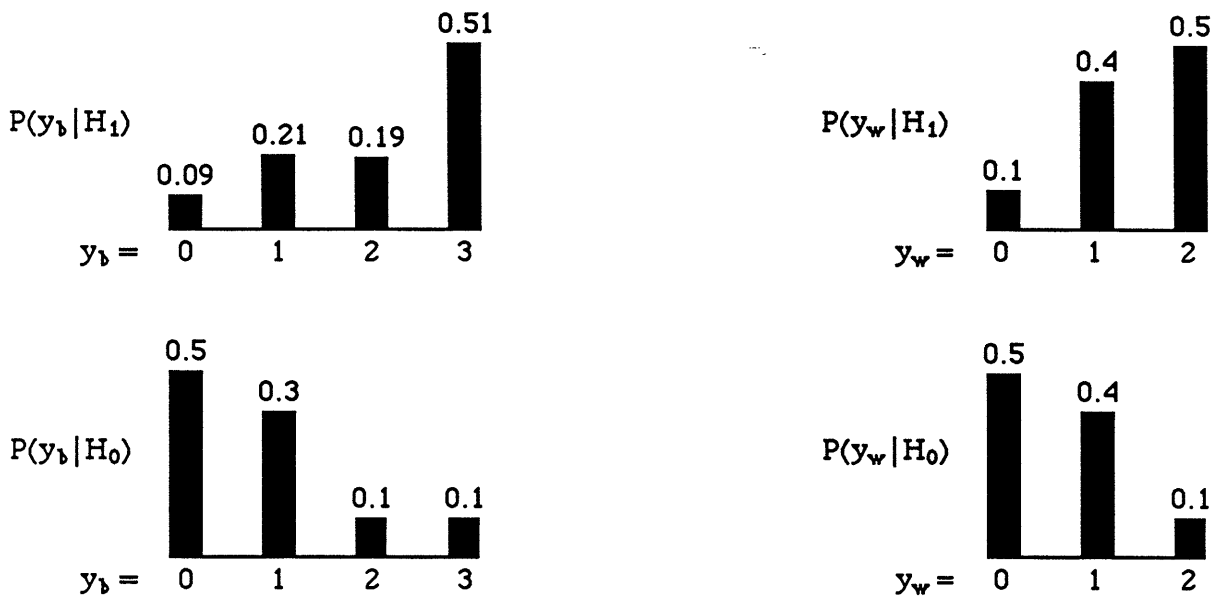


Figure 3. The DMs for *Example* 1

[1] This fact was formalized and generalized in [T89a].

[2] From the above discussion and the one following in section 3.2, one may conclude that the counterintuitive findings are a result of discrete distributions. They are not, because we can construct continuous distributions which have piecewise linear ROC curves (for example, a series of uniform distributions). Moreover, strictly concave ROC curves will not affect the results, since strictly concave curves which approximate the piecewise linear ROC curves within ε can be constructed.



(a). Better DM (b). Worse DM

Figure 4. The Probability Distributions for the DMs of Figure 3

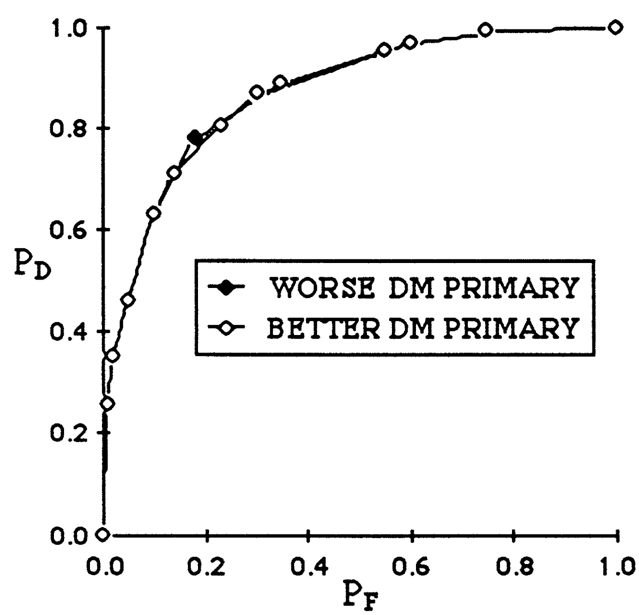


Figure 5(a). The Team ROC Curves for Both Configurations and Binary Messages

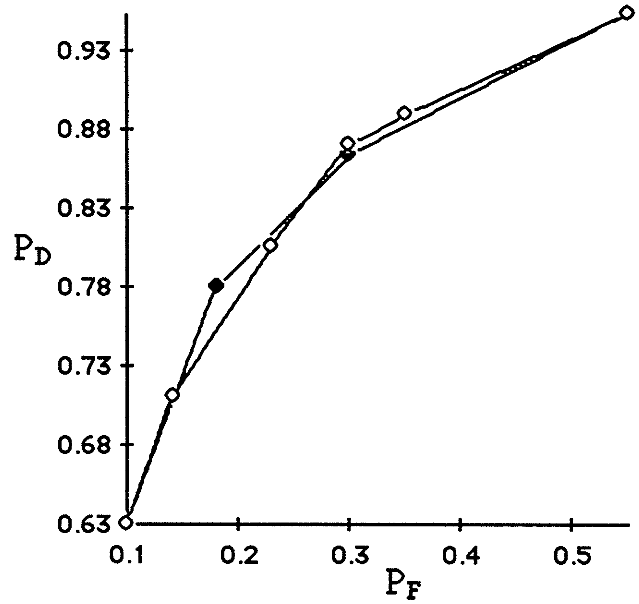


Figure 5(b). Close Up on Figure 5(a)

TABLE 1. Configuration Comparison for Binary Messages

(i). $\eta = \mathbf{0.4}$ $[P(H_0) = 0.2857]$	B → W →	W → B →
Operating Point of the Consulting DM:	(0.5, 0.91)	(0.5, 0.9)
Operating Point of the Primary DM when $u_c = 0$:	(0.1, 0.5)	(0.1, 0.51)
Operating Point of the Primary DM when $u_c = 1$:	(0.5, 0.9)	(0.5, 0.91)
Operating Point of the Team:	(0.3, 0.864)	(0.3, 0.87)
Probability of Error:	0.1829	0.1786*

(ii). $\eta = \mathbf{1.0}$ $[P(H_0) = 0.5]$	B → W →	W → B →
Operating Point of the Consulting DM:	(0.2, 0.7)	(0.1, 0.5)
Operating Point of the Primary DM when $u_c = 0$:	(0.1, 0.5)	(0.2, 0.7)
Operating Point of the Primary DM when $u_c = 1$:	(0.5, 0.9)	(0.5, 0.91)
Operating of the Team Point:	(0.18, 0.78)	(0.23, 0.805)
Probability of Error:	0.20*	0.2125

* Optimal

3.2. The Number of Messages

Consider again the example of the previous section and suppose that the number of messages that the consulting DM can transmit to the primary DM is increased to three. Then, if the worse DM is made the consulting DM, the team achieves the optimal centralized performance because the consulting DM can transmit its observation to the primary DM. The same result can **not** be achieved when the better DM is made the consulting DM. Hence, the ROC curve of the team with the worse DM as the consulting DM is higher than the one of the team with the better DM as the consulting DM (Figure 6); thus, for the case of three messages making the better DM the primary is optimal even for $\eta = 1.0$ (Table 2). Therefore, as can be seen from Tables 1(ii) and 2, the optimal team architecture depends also on the number of messages.

The fact that the optimal team architecture depends on the number of messages is not surprising, but the way it does is *counterintuitive*. Intuition suggests that, as the number of messages increases, it becomes more likely for the better DM to be placed as the consulting DM in the optimal configuration, because, as the number of messages increases, the loss of information caused by the fusion of the observation of the consulting DM to a message decreases. This is especially obvious in the two limit cases; in the zero message (isolation) case, the better DM should be the primary DM, thus making the team decision, and, in the infinite message (centralized) case, the better DM could be made the consulting DM without causing any deterioration in the team performance. In our particular example, though, assuming $\eta = 1.0$, increasing the number of messages from two to three makes the better DM change its role in the optimal team architecture from being the consulting DM to being the primary DM; a counterintuitive result.

TABLE 2. Configuration Comparison for Ternary Messages

η = **1.0** $[P(H_0) = 0.5]$	B → W →	W → B →
Operating Point (0 vs. 1) of the Consulting DM:	(0.5, 0.91)	(0.5, 0.9)
Operating Point (1 vs. 2) of the Consulting DM:	(0.2, 0.7)	(0.1, 0.5)
Operating Point of the Primary DM when $u_c = 0$:	(0.0, 0.0)	(0.1, 0.51)
Operating Point of the Primary DM when $u_c = 1$:	(0.1, 0.5)	(0.2, 0.7)
Operating Point of the Primary DM when $u_c = 2$:	(0.5, 0.9)	(0.5, 0.91)
Operating Point of the Team :	(0.13, 0.735)	(0.18, 0.786)
Probability of Error:	0.1975	0.197*

* Optimal

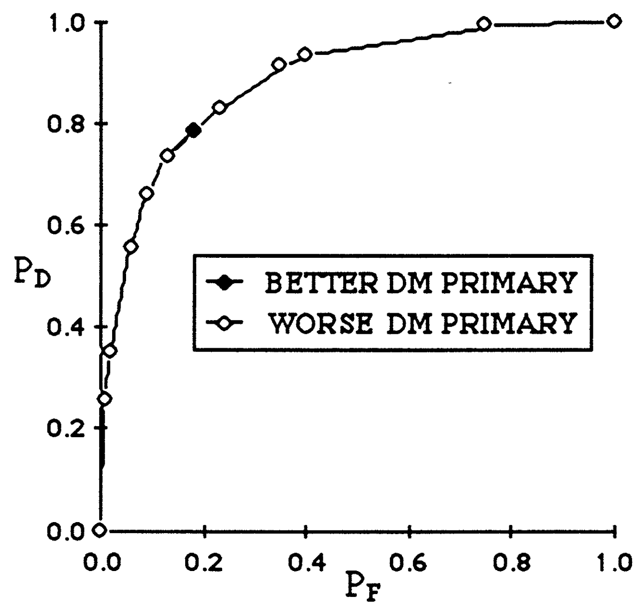


Figure 6(a). The Team ROC Curves for Both Configurations and Ternary Messages

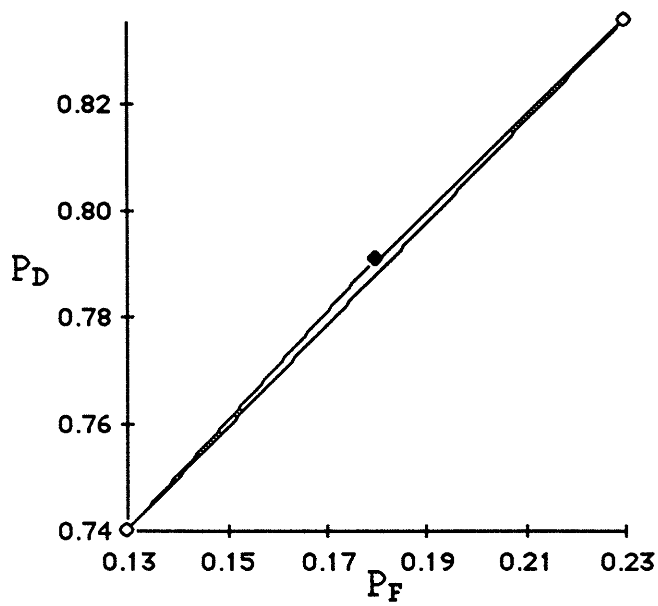


Figure 6(b). Close Up on Figure 6(a)

3.3. Performance Bounds

The main objective of this research is to obtain building blocks and design guidelines for large organizations. In section 3.1, it was shown that the intuitively appealing suggestion of designating the better DM as the primary DM is not optimal in general. Still, it is optimal for several probability density functions (as will be seen in section 3.4 below) and even in cases where it is not true (like the example presented above) both configurations have very similar performance. Therefore, it is logical to designate the better DM to be the primary DM and try to obtain a bound on the deterioration of the team performance. A meaningful bound is the deterioration of the team performance relatively to the optimal team performance.

But consider the following *Example* 2, for which the optimal team configuration requires that the better DM be the primary DM. The ROC curves for the two DMs are presented in Figure 7 and the associated discrete probability density functions conditioned on the two hypotheses are presented in Figure 8. Suppose that $\eta = 1.0$ (i.e., equal prior probabilities and minimum error cost function). The performance of the team is measured by the probability of error as usual.

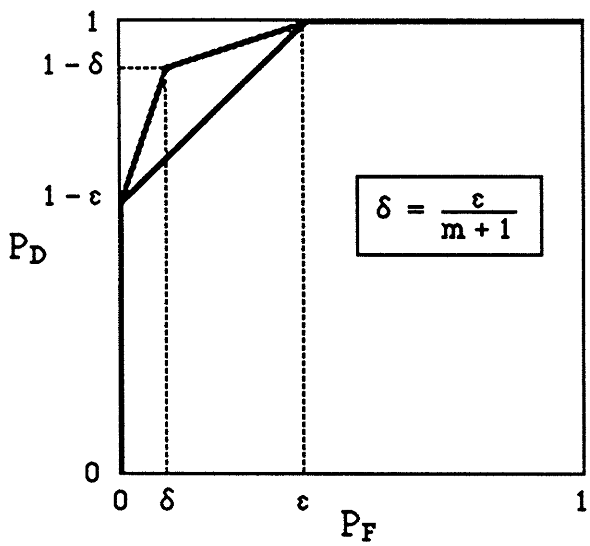


Figure 7. The DMs in *Example* 2

Consider any m such that:

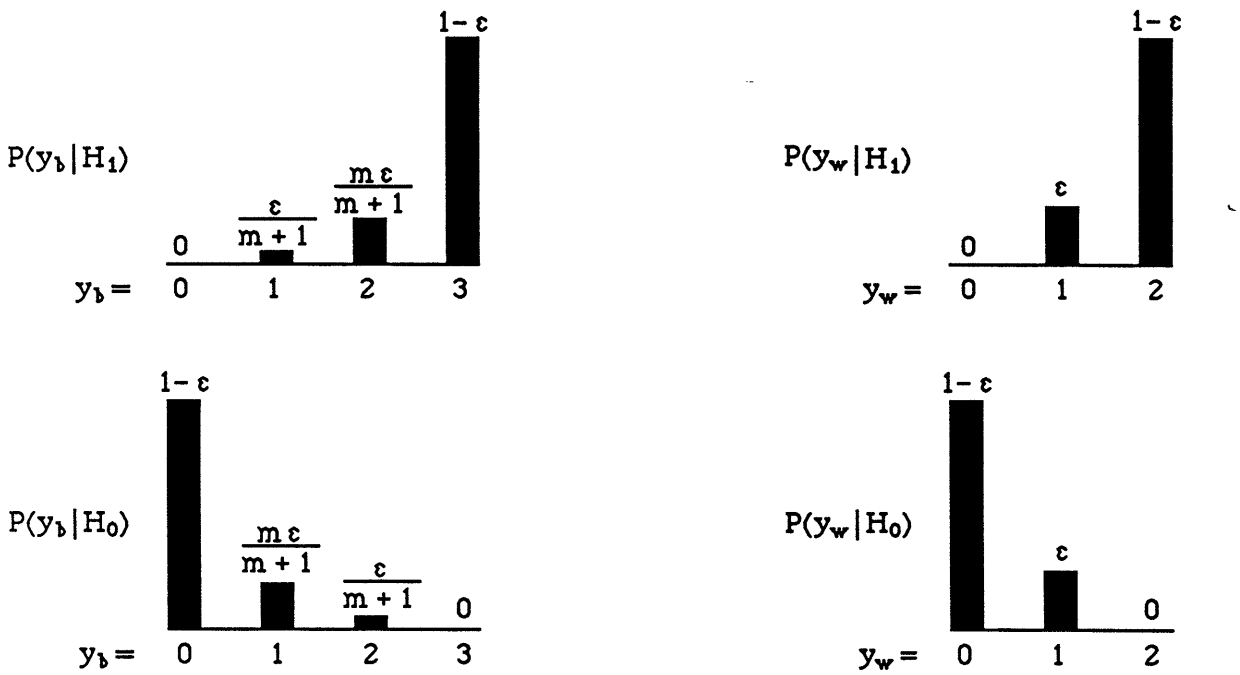
$$m > 1 \tag{4}$$

and any ε such that:

$$min\{1/2, 1/m\} > \varepsilon > 0 \tag{5}$$

Suppose that the better DM is the consulting DM. The optimal operating points can be found on Table 3 and the optimal probability of error is:

$$P_1^e = \frac{\varepsilon^2}{m+1} \tag{6}$$



(a). Better DM (b). Worse DM

Figure 8. The Probability Distributions for the DMs of Figure 7

Table 3. Configuration Comparisons for *Example* 2

η = **1.0** $[P(H_0) = 0.5]$	B → W →	W → B →
Operating Point of the Consulting DM:	$\left(\frac{\varepsilon}{m+1},\ 1-\frac{\varepsilon}{m+1}\right)$	$(0,\ 1-\varepsilon)$
Operating Point of the Primary DM when $u_c = 0$:	$(0,\ 1-\varepsilon)$	$(0,\ 1-\varepsilon)$
Operating Point of the Primary DM when $u_c = 1$:	$(\varepsilon,\ 1)$	$(\varepsilon,\ 1)$
Operating Point of the Team :	$\left(\frac{\varepsilon^2}{m+1},\ 1-\frac{\varepsilon^2}{m+1}\right)$	$(0,\ 1-\varepsilon^2)$
Probability of Error:	$\frac{\varepsilon^2}{m+1}$ *	$\frac{\varepsilon^2}{2}$

* Optimal

Now suppose that the better DM is the primary DM. The optimal operating points can also be found on Table 3 and the optimal probability of error is:

$$P_2^e = \frac{\varepsilon^2}{2} \tag{7}$$

Then the deterioration of the team performance is:

$$\Delta P^e = P_2^e - P_1^e = \frac{m-1}{m+1}\frac{\varepsilon^2}{2} > 0 \tag{8}$$

and the relative deterioration of the team performance is:

$$\omega = \frac{\Delta P^e}{P_1^e} = \frac{m-1}{2} \tag{9}$$

Since we can choose any $m > 1$, we conclude that the relative deterioration of the team performance can **not** be bounded this way. But also note that, as $m \to \infty$, the absolute magnitude of the deterioration of the team performance goes to zero; thus as the relative deterioration increases, the absolute magnitude of the deterioration decreases.

3.4. Special Probability Distributions

There exist certain probability distributions for which the configuration with the better DM as the primary DM seems always to be superior. As we already saw in section 3.1 this is not necessarily true for discrete distributions. But, our numerical analysis suggests that it is true for the case of comparing means of Gaussian distributions with equal variance. Unfortunately, due to the complexity of the improper integrals involved, no theoretical results were obtained to substantiate our numerical findings.

On the other hand, in the case of exponential distributions with different rates (or equivalently of comparing variances of Gaussian distributions with equal means, when each DM receives two observations) not only it is always better to designate the better DM as the primary DM, but also we were able to obtain a proof. Thus consider two such DMs; from [V68] we know that the ROC curve of the worse DM is given by:

$$P_D^w = (P_F)^\alpha \tag{10}$$

and the ROC curve of the better DM is given by:

$$P_D^b = (P_F)^{\alpha/k} \tag{11}$$

where the exponential rates are $1/\alpha$ for the worse DM and k/α for the better DM, with $k > 1$ (for the Gaussian case $\alpha = \sigma_0^2/\sigma_1^2$ with $\sigma_1 > \sigma_0$).

Suppose that the better DM is made the primary. Then from (B.1)–(B.3) and the property of the tangent to the ROC curve:

$$\eta_c = \frac{P_F^1 - P_F^0}{P_D^1 - P_D^0}\,\eta = \left.\frac{dP_D^b}{dP_F}\right|_{(P_F^c, P_D^c)} = \frac{\alpha}{k}\frac{P_D^c}{P_F^c} \tag{12}$$

$$\eta_0 = \frac{1 - P_F^c}{1 - P_D^c}\,\eta = \left.\frac{dP_D^w}{dP_F}\right|_{(P_F^0, P_D^0)} = \alpha\frac{P_D^0}{P_F^0} \tag{13}$$

and:

$$\eta_1 = \frac{P_F^c}{P_D^c}\,\eta = \left.\frac{dP_D^w}{dP_F}\right|_{(P_F^1, P_D^1)} = \alpha\frac{P_D^1}{P_F^1} \tag{14}$$

Solving the system of (12)-(14) and recalling the concavity of the ROC curve, we obtain that:

$$(P_F^1, P_D^1) = (1,1)$$

which implies that whenever $u_c = 1$ is received from the consulting DM, the primary DM decides $u_p = 1$ independent of its own observation. Substituting into (B.5) and (B.6), we obtain that the point (P_F^t, P_D^t) of the team ROC curve, in this case, is given parametrically by:

$$P_F^t = P_F^0 + P_F^c - P_F^0 P_F^c \tag{15}$$

and:

$$P_D^t = P_D^0 + P_D^c - P_D^0 P_D^c \tag{16}$$

for some optimal operating points (P_F^0, P_D^0) of the worse DM and (P_F^c, P_D^c) of the better DM.

Now suppose that the better DM is made the primary DM. Then, we can arbitrarily assign to the DMs the following operating points:

(P_F^0, P_D^0): to the new consulting (worse) DM
(P_F^c, P_D^c): to the new primary (better) DM when $u_c = 0$ is received
$(1,1)$: to the new primary (better) DM when $u_c = 1$ is received

Substituting into (B.5) and (B.6) we obtain (15) and (16) again. Since for this arbitrary assignment of operating points, the configuration with the better DM as the primary DM can achieve performance equal to the optimal performance of the other configuration, *the better DM should always be the primary DM.*

3.5. Propagation of ROC Curves

Suppose that the better DM is made the consulting DM in the example of section 3.4 above. Then from the system of (10)-(14), we can solve for P_F^c to obtain:

$$P_F^c = \left[k\, \sigma_1^2 \left(1 - P_D^c\right) \eta \; \frac{\left[\sigma_0^2 \left(1 - P_D^c\right)\right]^{-\gamma} - \left[\sigma_1^2 \left(1 - P_F^c\right) \eta\right]^{-\gamma}}{\left[\sigma_0^2 \left(1 - P_D^c\right)\right]^{-\beta} - \left[\sigma_1^2 \left(1 - P_F^c\right) \eta\right]^{-\beta}} \right]^{-\delta} \tag{17}$$

where:

$$\beta = \frac{\sigma_0^2}{\sigma_1^2 - \sigma_0^2} \qquad \gamma = \frac{\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \qquad \delta = \frac{k\,\sigma_1^2}{k\,\sigma_1^2 - \sigma_0^2} \tag{18}$$

(17) is an equation of just P_F^c. We could have substituted for P_D^c from (15), but did not do it because of space limitations. If the equation is solved P_F^c is obtained. Moreover:

$$P_F^0 = \left[\alpha \, \frac{1 - P_F^c}{1 - P_D^c} \, \eta \right]^{-\gamma} \tag{19}$$

By substituting into (13), P_D^0 is obtained. Finally by substituting for all the probabilities into (15) and (16), the team ROC curve is obtained as a function of η, the variances and k.

It should be clear that the team ROC curve will not be of the same form as the ROC curves of the individual DMs. In fact, it can not be given in a closed form expression. Thus, the results cannot be easily extended to the case of three DMs in tandem even for these particular probability distributions.

4. TWO DMs IN PARALLEL

The team which consists of two DMs in parallel (Figure 1(b)) was the first one to be studied in this framework [TS81]. It was shown that even if the DMs are identical and the cost structure symmetric, the optimal decision rules of the two DMs do not have to be symmetric. This was attributed to the explicit dependence of the cost function on the decisions of the DMs. Subsequently, this result was also demonstrated with a simple example for the case where the cost function depends implicitly on the decisions of the DMs through the fusion center [T88]. The necessary conditions which characterize the optimal decision rules of the two DMs are presented in appendix D. We further analyze the problem in order to understand the reasons of the counterintutive behavior of the team members.

4.1. Second Order Optimality Conditions

As was mentioned above the optimality conditions of appendix D are necessary conditions. We also derive the second order necessary conditions for optimality in order to enhance our understanding of and intuition about the team behavior. These conditions depend, as the first order conditions do, on whether the AND or the OR decision rule is the optimal rule employed by the fusion center.

PROPOSITION 1. *Consider the team which consists of two DMs in parallel and a fusion center, and performs binary hypothesis testing (Figure* 1(b)).

(i). If the fusion center employs the AND rule as its optimal decision rule, the second order necessary optimality conditions are given by the following inequalities:

$$-\eta + \eta_a \eta_b \geq 0 \tag{20a}$$

or:

$$\alpha_a \alpha_b P_D^a P_D^b \geq (-\eta + \eta_a \eta_b)^2 \tag{20b}$$

where α_n is the second derivative of the ROC curve at the operating point (P_F^n, P_D^n), for $n = a, b$.

(ii). If the fusion center employs the OR rule as its optimal decision rule, the second order necessary optimality conditions are given by the following inequalities:

$$\eta - \eta_a \eta_b \geq 0 \tag{21a}$$

or:

$$\alpha_a \alpha_b (1 - P_D^a)(1 - P_D^b) \geq (\eta - \eta_a \eta_b)^2 \tag{21b}$$

Proof. Only (20) needs to be proved because (20) and (21) are symmetric. To prove (20), consider the two operating points (P_F^a, P_D^a) and (P_F^b, P_D^b) which satisfy the first order conditions and without loss of generality assume that $P_F^a \leq P_F^b$. Perturb P_F^a to $P_F^a + \varepsilon$, where ε is a real number of very small magnitude; then, by Taylor's theorem, the perturbed probability of detection will be $P_D^a + \varepsilon \eta_a + 0.5\varepsilon^2 \alpha_a$. Similarly, perturb the operating point of the other DM to $(P_F^b - \delta, P_D^b - \delta \eta_b + 0.5\delta^2 \alpha_b)$, where δ is a real number of very small magnitude with $\varepsilon\delta > 0$; note that the two operating points should be perturbed in opposite directions so that they continue to satisfy the first order conditions.

For (P_F^a, P_D^a) and (P_F^b, P_D^b) to be globally optimal, they need to satisfy the following necessary condition:

$$\frac{\eta}{\eta+1} P_F^a P_F^b + \frac{1}{\eta+1}(1 - P_D^a P_D^b) \leq \frac{\eta}{\eta+1}(P_F^a + \varepsilon)(P_F^b - \delta) +$$

$$+ \frac{1}{\eta+1}\left[1 - (P_D^a + \varepsilon \eta_a + 0.5\varepsilon^2 \alpha_a)(P_D^b - \delta \eta_b + 0.5\delta^2 \alpha_b)\right] \Rightarrow$$

$$0 \leq \varepsilon(\eta P_F^b - \eta_a P_D^b) + \delta(-\eta P_F^a + \eta_b P_D^a) + \varepsilon\delta(-\eta + \eta_a \eta_b) +$$

$$+\varepsilon^2(-0.5\alpha_a P_D^b) + \delta^2(-0.5\alpha_b P_D^a) + \text{H.O.T.} \quad (22)$$

For global optimality, (22) should hold for both positive and negative perturbations ε and δ which implies that the coefficients of ε and of δ should be zero; these, as expected, are the first order conditions of (D.1) and (D.2) respectively. Moreover, the coefficients of ε^2 and of δ^2 are non-negative, because the ROC curve is concave and thus has a non-positive second derivative. Hence, if the coefficient of $\varepsilon\delta$ is non-negative, that is if (20a) holds, then (22) holds as well.

Assume that (20a) does not hold; (22) may hold even in this case as long as:

$$\phi(\varepsilon, \delta) \equiv \varepsilon^2(-0.5\,\alpha_a P_D^b) + \varepsilon\delta(-\eta + \eta_a\eta_b) + \delta^2(-0.5\,\alpha_b P_D^a) \geq 0 \quad (23)$$

for ε and δ sufficiently small with $\varepsilon\delta > 0$.

We minimize $\phi(\varepsilon, \delta)$ with respect to ε and to δ. Using simple calculus, the function is minimized with respect to ε if and only if:

$$\varepsilon = \delta\,\frac{(-\eta + \eta_a\eta_b)}{\alpha_a P_D^b} \quad (24)$$

Observe that the coefficient of δ in (24) is positive (because we assumed that (20a) does not hold), so that $\varepsilon\delta > 0$ as required. Substituting for ε from (24) into (23), we obtain that (23) is true if and only if (20b) is true. **Q.E.D.**

CORROLARY 1. *Consider again the team of the previous proposition. Then, the second order necessary conditions for optimality can be written in the following equivalent form:*

(i). If the fusion center employs the AND decision rule as its optimal decision rule:

$$-\eta + \eta_a\eta_b \geq -(\alpha_a\alpha_b P_D^a P_D^b)^{0.5} \quad (25)$$

(ii). If the fusion center employs the OR decision rule as its optimal decision rule:

$$\eta - \eta_a\eta_b \geq -\left[\alpha_a\alpha_b(1-P_D^a)(1-P_D^b)\right]^{0.5} \quad (26)$$

Proof. Again only (25) needs to be shown, since the proof of (26) follows by symmetry. We show first that if (20) holds, then (25) holds as well; if (20a) is true, then (25) is obviously also true because the right hand side of (25) is non-positive. Suppose that (20a) does not hold and that (20b) holds; then by taking the square root of each side of (20b), we obtain (25).

To complete the proof we have to demonstrate that if (25) is true, then (20) is true as well; so suppose that (25) holds and distinguish between two cases depending on whether $-\eta + \eta_a\eta_b \geq 0$ or $-\eta + \eta_a\eta_b < 0$. If the former is true, then (20a) is obviously true; if the latter is true, take the square of each side of (25) to obtain (20b). **Q.E.D.**

4.2. Identical DMs

We now focus on the special case where both DMs are identical. Because of the symmetry of the problem, it seems intuitive that the optimal decision rules of the two DMs will be identical or symmetric at least for the case where $\eta = 1$ (i.e., perfect symmetry of the variables external to the team). It is known that the optimal decision rules neither have to be identical nor have to be symmetric even if $\eta = 1$; a simple example with discrete probability density functions was presented in [T88]. We should note that such an example can be constructed for strictly concave ROC curves as well. These examples indicate that it is optimal in general for the team to split the risk unevenly among its members and suggest the need for different levels of command.

From an organizational designer's point of view it is desirable to assign identical decision rules to similar DMs having the same position in the organization; this not only significantly reduces the complexity of the problem, but it is also asymptotically optimal [T88]. So, we suppose that both DMs are restricted to employing identical decision rules and derive the first and second order optimality conditions; these follow directly from appendix D and from Corrolary 1 and thus the proofs are omitted.

CORROLARY 2. *Consider the team which consists of two identical DMs in parallel and a fusion center, and performs binary hypothesis testing (Figure 1(b)). Suppose the two DMs are restricted to employing identical decision rules. Then the optimal decision operating point* (P_F^n, P_D^n) *of the ROC curve satisfies the following conditions:*

(i). If the fusion center employs the AND decision rule as its optimal decision rule:

$$\left.\frac{dP_D}{dP_F}\right|_{(P_F^n, P_D^n)} = \frac{P_F^n}{P_D^n}\eta \tag{27}$$

(ii). If the fusion center employs the OR decision rule as its optimal decision rule:

$$\left.\frac{dP_D}{dP_F}\right|_{(P_F^n, P_D^n)} = \frac{1-P_F^n}{1-P_D^n}\eta \tag{28}$$

REMARK: Because of the concavity of the ROC curve there exists one and only one point on every ROC curve which satisfies each of (27) and of (28).

CORROLARY 3. *Consider again the team of Corrolary 2. Then, the second order necessary conditions for optimality are:*

(i). If the fusion center employs the AND decision rule as its optimal decision rule:

$$-\eta + (\eta_n)^2 \geq \alpha_n P_D^n \tag{29}$$

(ii). If the fusion center employs the AND decision rule as its optimal decision rule:

$$\eta - (\eta_n)^2 \geq \alpha_n (1 - P_D^n) \tag{30}$$

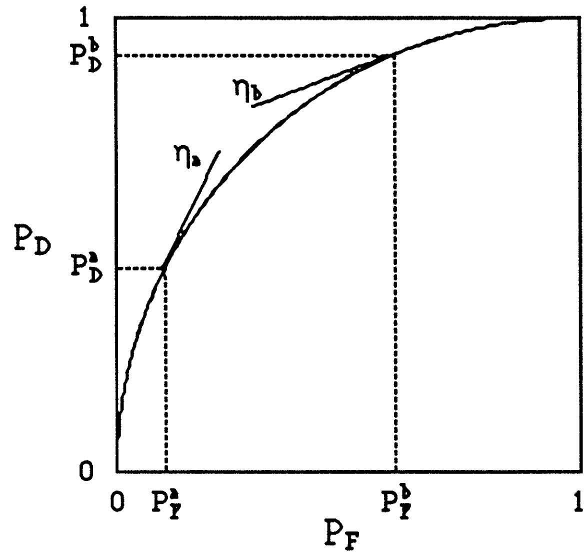
The above conditions suggest that DMs whose associated ROC curves have steep changes, implying large absolute values for the second derivative, are likely to have identical optimal decision rules (for example ROC curves for nearly uniform noise).

4.3. Performance Bounds

As was mentioned above, in [T88] it was shown that as the number of DMs in a parallel team, which performs binary hypothesis testing, increase to infinity, the decision rules of the DMs may be restricted to be identical without any deterioration in the team performance. Since this is not true for the parallel team which consists of two DMs, we would like to determine whether bounds exist for the deterioration of its performance.

(i). Absolute Bound

Consider two identical DMs whose optimal decision rules are not identical for some given η^*, as the ones in Figure 9(a); suppose also, without loss of generality, that the optimal decision rule for the fusion center is the AND decision rule. The operating points of the DMs have to satisfy the necessary optimality conditions of (D.1) and (D.2); without loss of generality assume $P_F^a < P_F^b$. Also consider two new identical DMs with the three-piecewise linear ROC curve of Figure 9(b). The new DMs are of course worse than the original DMs; still, for the particular η^*, the new (worse) team can achieve performance equal to the optimal performance of the original (better) team, since the operating points (P_F^a, P_D^a) and (P_F^b, P_D^b) can also be employed by the new (worse) DMs. Therefore, if the DMs are restricted to employing identical decision rules, an upper bound for the deterioration in the performance of the original (better) team, for the particular η^*, is given by the deterioration in the performance of the new (worse) team.



$$\eta_a = \frac{P_F^b}{P_D^b}\eta^*$$

$$\eta_b = \frac{P_F^a}{P_D^a}\eta^*$$

Figure 9(a). Identical DMs with Non-Identical Optimal Decision Rules

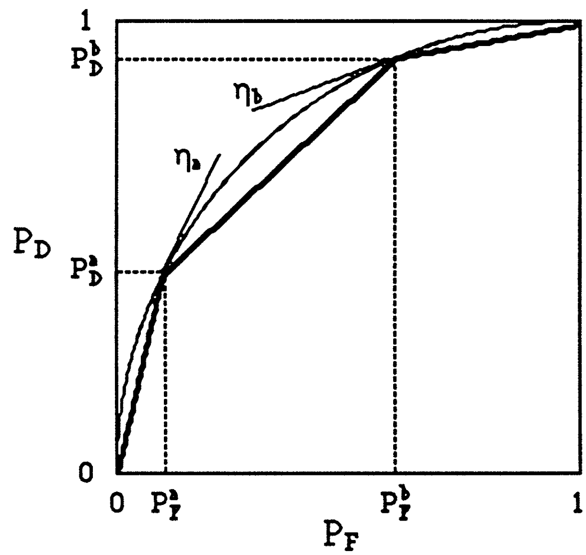


Figure 9(b). The 'Worse' Identical DMs with Non-Identical Optimal Decision Rules

y \ H	0	1	2
H_0	$1 - P_F^b$	$P_F^b - P_F^a$	P_F^a
H_1	$1 - P_D^b$	$P_D^b - P_D^a$	P_D^a

Table 4. The Probability Distributions for the DMs of Figure 9(b)

Thus, the problem is reduced to obtaining bounds for teams which consist of DMs with three-piecewise linear ROC curves[1]. Consider again the DMs of Figure 9(b), whose underlying probability density functions are presenred in Table 4. The set of possibly optimal decision rules may be found by exhaustive enumeration. Since each DM has to perform a likelihood ratio test, there are only two candidate decision rules for each DM n, $n = a, b$:

[i] $u_n = 1$ *if and only if* $y_n \in \{2\}$
[ii] $u_n = 1$ *if and only if* $y_n \in \{1, 2\}$

Thus, we need to consider six cases:

[1] *Both DMs employ* [i] *and the fusion center employs* AND
[2] *Both DMs employ* [ii] *and the fusion center employs* AND
[3] *DM a employs* [i], *DM b employs* [ii] *and the fusion center employs* AND
[4] *Both DMs employ* [i] *and the fusion center employs* OR
[5] *Both DMs employ* [ii] *and the fusion center employs* OR
[6] *DM a employs* [i], *DM b employs* [ii] *and the fusion center employs* OR

REMARK: In order to avoid the explicit dependence of the probability of error on both costs and prior probabilities, we define the (normalized) probability of error in terms of η as follows:

$$P^e = \frac{\eta}{\eta + 1} P_F + \frac{1}{\eta + 1}(1 - P_D) \tag{31}$$

---

[1] This fact was formalized and generalized in [T89a].

For the given η^*, denote by P_j^e the team probability of error when the decision rules of case $[j]$, $j = 1, ..., 6$, are employed and by ΔP^e the minimum deterioration in the team performance, if both DMs are restricted to employing identical decision rules. Since [3] was assumed to be the optimal set of decision rules:

$$\Delta P^e \le P_1^e - P_3^e = \frac{\eta^*}{\eta^*+1}\left[-P_F^a\,(P_F^b - P_F^a)\right] + \frac{1}{\eta^*+1}\left[P_D^a\,(P_D^b - P_D^a)\right] \equiv \Delta P_1^e \qquad (32a)$$

$$\Leftrightarrow \frac{P_D^a\,(P_D^b - P_D^a)}{P_F^a\,(P_F^b - P_F^a)} \ge \eta^* \qquad (32b)$$

$$\Delta P^e \le P_2^e - P_3^e = \frac{\eta^*}{\eta^*+1}\left[P_F^b\,(P_F^b - P_F^a)\right] - \frac{1}{\eta^*+1}\left[P_D^b\,(P_D^b - P_D^a)\right] \equiv \Delta P_2^e \qquad (33a)$$

$$\Leftrightarrow \eta^* \ge \frac{P_D^b\,(P_D^b - P_D^a)}{P_F^b\,(P_F^b - P_F^a)} \qquad (33b)$$

$$\Delta P^e \le P_4^e - P_3^e = \frac{\eta^*}{\eta^*+1}\left[P_F^a\,(1 - P_F^b) + P_F^a\,(1 - P_F^a)\right] -$$

$$- \frac{1}{\eta^*+1}\left[P_D^a\,(1 - P_D^b) + P_D^a\,(1 - P_D^a)\right] \equiv \Delta P_4^e \qquad (34a)$$

$$\Leftrightarrow \eta^* \ge \frac{P_D^a\,(1 - P_D^b) + P_D^a\,(1 - P_D^a)}{P_F^a\,(1 - P_F^b) + P_F^a\,(1 - P_F^a)} \qquad (34b)$$

$$\Delta P^e \le P_5^e - P_3^e = \frac{\eta^*}{\eta^*+1}\left[P_F^b\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)\right] -$$

$$- \frac{1}{\eta^*+1}\left[P_D^b\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)\right] \equiv \Delta P_5^e \qquad (35a)$$

$$\Leftrightarrow \eta^* \ge \frac{P_D^b\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)}{P_F^b\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)} \qquad (35b)$$

$$0 \le P_6^e - P_3^e = \frac{\eta^*}{\eta^*+1}\left[P_F^a\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)\right] - \frac{1}{\eta^*+1}\left[P_D^a\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)\right]$$

$$\Leftrightarrow \eta^* \ge \frac{P_D^a\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)}{P_F^a\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)} \qquad (36)$$

REMARK: (36) does not provide a bound on the deterioration of the team performance incured by the restriction that both DMs employ identical decision rules, because, according to [6], the DMs of the team do *not* employ identical decision rules. Nevertheless, it provides a condition for η^* so that [3] be the optimal set of decision rules employed by the team.

Using simple algebra and since $P_D^a/P_F^a > P_D^b/P_F^b$ (Figure 9(b)), we obtain that, as long as (34b) holds true, (35b) holds true as well; this implies that the bound of (34a) is tighter than the bound of (35a). So, we need only to examine (32a), (33a) and (34a) in order to obtain the *maximum least upper bound* for the deterioration of the team performance.

Furthermore, note that the bound of (32a) is decreasing with η^*, while the bounds of (33a) and (34a) are both increasing with η^*. Therefore, the maximum least upper bound of (32a) and (33a) occurs when the bounds are equal. Thus:

$$P_1^e - P_3^e = P_2^e - P_3^e \iff \eta^* = \frac{P_D^b + P_D^a}{P_F^b + P_F^a}\frac{P_D^b - P_D^a}{P_F^b - P_F^a} \tag{37}$$

Similarly, the maximum least upper bound of (32a) and (34a) occurs when the bounds are equal. Thus:

$$P_1^e - P_3^e = P_4^e - P_3^e \iff \eta^* = \frac{P_D^a}{P_F^a}\frac{1 - P_D^a}{1 - P_F^a} \tag{38}$$

REMARK: Using simple algebra, we verify that the η^* of both (37) and (38) satisfy (36) because:

$$\frac{P_D^b + P_D^a}{P_F^b + P_F^a}\frac{P_D^b - P_D^a}{P_F^b - P_F^a} \geq max\left\{\frac{P_D^b}{P_F^b}\frac{P_D^b - P_D^a}{P_F^b - P_F^a}, \frac{P_D^b + P_D^a}{P_F^b + P_F^a}\frac{1 - P_D^b}{1 - P_F^b}\right\} \geq$$

$$\geq \frac{P_D^b\,(P_D^b - P_D^a) + (P_D^b + P_D^a)\,(1 - P_D^b)}{P_F^b\,(P_F^b - P_F^a) + (P_F^b + P_F^a)\,(1 - P_F^b)} = \frac{P_D^a\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)}{P_F^a\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)}$$

and:

$$\frac{P_D^a}{P_F^a}\frac{1 - P_D^a}{1 - P_F^a} \geq max\left\{\frac{P_D^a\,(1 - P_D^b)}{P_F^a\,(1 - P_F^b)}, \frac{P_D^b\,(1 - P_D^a)}{P_F^b\,(1 - P_F^a)}\right\} \geq \frac{P_D^a\,(1 - P_D^b) + P_D^b\,(1 - P_D^a)}{P_F^a\,(1 - P_F^b) + P_F^b\,(1 - P_F^a)}$$

Substituting for η^* from (37) into (32a) (or into (33a)), we obtain the maximum least upper bound of (32a) and (33a). Substituting for η^* from (38) into (32a) (or into (34a)), we obtain the maximum least upper bound of (32a) and (34a). The globally maximum least upper bound is obtained when both the above bounds are equal, that is when:

$$\frac{P_D^b + P_D^a}{P_F^b + P_F^a} \frac{P_D^b - P_D^a}{P_F^b - P_F^a} = \frac{P_D^a}{P_F^a} \frac{1 - P_D^a}{1 - P_F^a} \tag{39a}$$

$$\Leftrightarrow \quad (P_D^b)^2 = \frac{P_D^a}{P_F^a} \frac{1 - P_D^a}{1 - P_F^a} (P_F^b)^2 + \frac{P_D^a (P_D^a - P_F^a)}{1 - P_D^a} \tag{39b}$$

To maximize the least upper bound, the two operating points of the DMs have to satisfy (39b); it is not hard to see (Appendix E) that to maximize the least upper bound for a fixed (P_F^a, P_D^a), subject to (39b), the operating point (P_F^b, P_D^b) has to satisfy:

$$\begin{aligned} (P_F^b, P_D^b) &= \left(\left[(P_F^a)^2 + \frac{(1 + P_D^a)(1 - P_F^a) P_F^a}{P_D^a} \right]^{0.5}, 1 \right) \\ &= \left(\left[\frac{P_F^a}{P_D^a} (1 + P_D^a - P_F^a) \right]^{0.5}, 1 \right) \end{aligned} \tag{40}$$

Moreover, substituting from (38) and (40) into (32a), we obtain that the maximum least upper bound of the deterioration in the team performance as a function of (P_F^a, P_D^a) is given by:

$$\Delta P^e \leq \frac{P_D^a \dfrac{1 - P_D^a}{1 - P_F^a}}{1 + \dfrac{P_D^a}{P_F^a} \dfrac{1 - P_D^a}{1 - P_F^a}} \left[1 - \left[\frac{P_F^a}{P_D^a} (1 + P_D^a - P_F^a) \right]^{0.5} \right] \tag{41}$$

In order to obtain the largest absolute deviation from optimality, we have to maximize the bound of (41) with respect to P_F^a and P_D^a, subject to:

$$0 \leq P_F^a \leq P_D^a \leq 1 \tag{42}$$

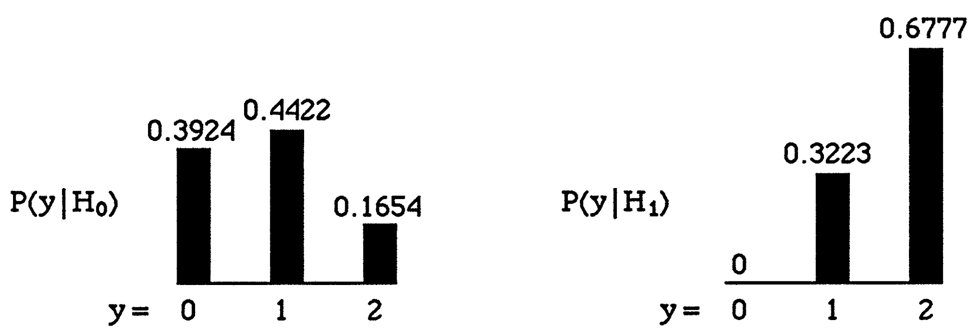


Figure 10. The Probability Distributions that Maximize the *Absolute* Bound

TABLE 5. The Probability of Error of Team of Identical DMs of Figure 10

Decision Rules of DMs	Fusion Center Decision Rule AND	OR
[i] and [i]	0.226179†	0.226179†
[ii] and [ii]	0.226179†	0.518340
[i] and [ii]	0.186403*	0.412035

$\Delta P^e = 0.039776$ $\eta = 1.581904$ $[P(H_0) = 0.612689]$
* Globally Optimal † Optimal for Identical Decision rules

Using calculus, the bound is maximized[1] for $(P_F^a, P_D^a) = (0.165438, 0.677735)$, giving $P_F^b = 0.607583$ and:

$$\Delta P^e \leq 0.039776 \tag{43}$$

Note that this is a tight bound; it is achieved by a parallel team which consists of two DMs, whose underlying probability distributions are presented in Figure 10, for $\eta^* = 1.581904$ obtained from (37) (Table 5). Thus, assuming minimum probability of error cost structure, the absolute deviation is not maximized for equal prior probabilities, but for: $P(H_0) = 0.612689$.

(ii). Relative Bound

Consider the exact same team used in the analysis of the absolute bound, the same given η^* and again denote by P_3^e the optimal probability of error of the two DM parallel team. Also, denote by P_s^e the probability of error when the DMs are employing some other decision rules; then define the *relative deterioration* of the team performance as:

$$\omega^e \equiv \frac{P_s^e - P_3^e}{P_3^e} \tag{44}$$

Proceeding with analysis similar to the above, in order to obtain the least upper bound on the relative deterioration, only the following three relative bounds, derived from (33a), (34a) and (35a) respectively, need to be considered:

[1] To be more specific, we expressed P_F^a and P_D^a in terms of the slopes $M = P_D^a/P_F^a$ and $m = (1 - P_D^a)/(1 - P_F^a)$, and then maximized with respect to M and m subject to: $M \geq 1 \geq m$.

$$\omega^e \le \frac{P_1^e - P_3^e}{P_3^e} = \frac{\eta^*(P_F^a)^2 + [1 - (P_D^a)^2]}{\eta^* P_F^a P_F^b + [1 - P_D^a P_D^b]} - 1 \equiv \omega_1^e \tag{45}$$

$$\omega^e \le \frac{P_2^e - P_3^e}{P_3^e} = \frac{\eta^*(P_F^b)^2 + [1 - (P_D^b)^2]}{\eta^* P_F^a P_F^b + [1 - P_D^a P_D^b]} - 1 \equiv \omega_2^e \tag{46}$$

$$\omega^e \le \frac{P_4^e - P_3^e}{P_3^e} = \frac{\eta^*[1 - (1 - P_F^a)^2] + (1 - P_D^a)^2}{\eta^* P_F^a P_F^b + [1 - P_D^a P_D^b]} - 1 \equiv \omega_4^e \tag{47}$$

To maximize the least upper bound of (45)-(47), (39) and (40) have to hold again (Appendix E); substituting from (39) and (40) into (45):

$$\omega^e \le \frac{1 + P_D^a - P_F^a}{1 - P_F^a + P_D^a \left[\frac{P_F^a}{P_D^a}(1 + P_D^a - P_F^a)\right]^{0.5}} - 1 \tag{48}$$

In order to obtain the largest absolute deviation from optimality, we have to maximize the bound of (48) with respect to P_F^a and P_D^a, subject to (42). Using calculus, the bound is maximized for $(P_F^a, P_D^a) \to (0, 1)$ giving:

$$\omega^e \le 1 \tag{49}$$

The bound of (49) is a tight bound, since it can be achieved by a team which consists of two DMs, whose underlying probability distributions are presented in Figure 11, as $\varepsilon \to 0$. It is achieved for $\eta^* = 1.0$ obtained from (38), as can be seen in Table 6. Note that this 100% bound is obtained as the team probability of error is going to zero. As the optimal probability of error increases the relative deterioration in the team perfromance is considerably smaller; in fact, the relative deterioration for the example of Table 5, in which the absolute deterioration in performance is maximized, is only 21.34%.

These results suggest, that it could be worthwhile for the designer of the team to incur some additional error and in the same time considerably simplify the complexity of the problem by restricting similar DMs to employ the same decision rules.

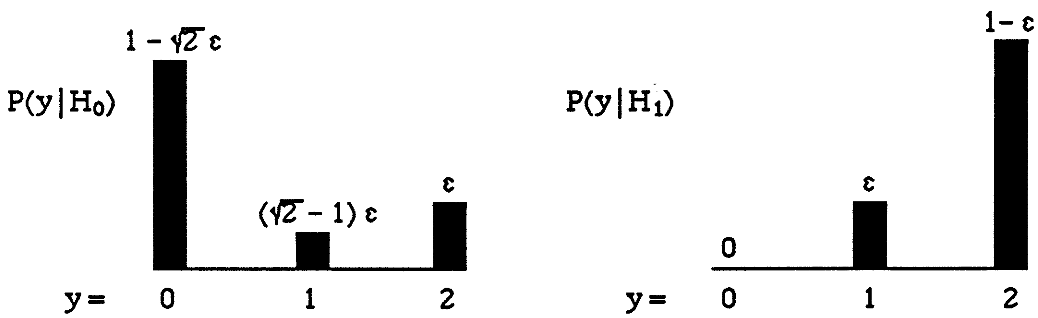


Figure 11. The Probability Distributions that Maximize the *Relative* Bound

TABLE 6. The Probability of Error of Team of Identical DMs of Figure 11

	Fusion Center Decision Rule	
	AND	OR
Decision Rules of DMs		
[i] and [i]	ε †	ε †
[ii] and [ii]	ε †	$1.41\varepsilon + \varepsilon^2$
[i] and [ii]	$0.5\varepsilon + 0.705\varepsilon^2$ *	$1.205\varepsilon - 0.707\varepsilon^2$

$\omega^{e} \approx 1.0$ $\eta = 1.0$ $[P(H_0) = 0.5]$

* Globally Optimal † Optimal for Identical Decision rules

5. TEAMS OF THREE DMs

There exist four different acyclic architectures for a team which consists of three DMs; the *two consultant* or *V*-architecture (Figure 12), the three DM *tandem* architecture (Figure 13), the three DM *parallel* architecture (Figure 14) and the *asymmetrical* architecture (Figure 15). We compare the performance of the different architectures to determine whether a *dominant* architecture exists.

It is worthwhile to understand the changes in the complexity of the decision rules when the number of team members increases from two to three. Three thesholds describe the decision rules of both the two DM tandem architecture (one for the consulting DM and two for the primary DM) and for the two DM parallel architecture (one for each DM and one for the fusion center). Six thresholds describe the decision rules for the *V*-architecture (one for each consulting DM and four for the primary DM), five thresholds describe the decision rules for both the three DM tandem

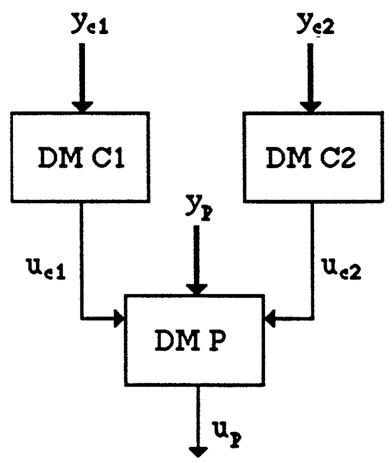


Figure 12. The *Two Consultant* or *V*-Architecture

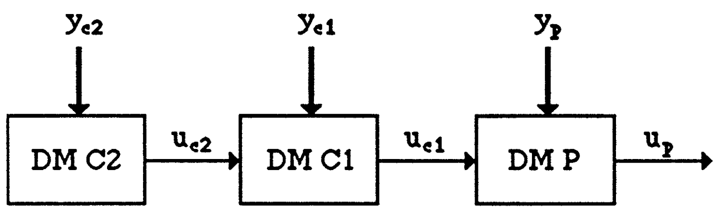


Figure 13. The Three DM *Tandem* Architecture

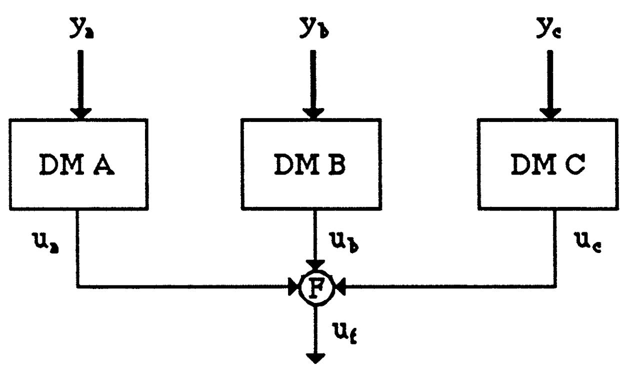


Figure 14. The Three DM *Parallel* Architecture

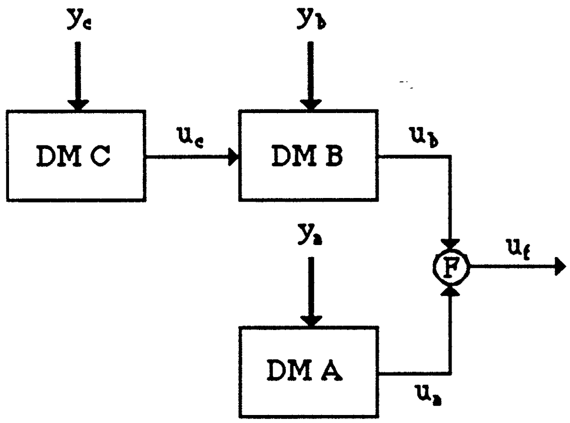


Figure 15. The Three DM *Asymmetrical* Architecture

architecture (one for the second consulting DM, two for the first consulting DM and two for the primary DM) and the three DM asymmetrical architecture (two for DM B and one for DM C, DM A and the fusion center), and four thresholds describe the decision rules for the three DM parallel architecture (one for each DM and one for the fusion center). It should be clear that the complexity of the decision rules depends not only on the number of the DMs in the team, but also on the particular architecture of the team. Therefore, we will compare the performance of the different architectures for the teams of three DMs, keeping in mind that the complexity and communication requirements are not the same for all architectures. We should also note that the equations which define the optimal decision rules for the members of the three DM teams are of the same form as, though more complex than, the equations which define the optimal decision rules for the members of the two DM teams[1].

As was shown in Lemma 1, the two DM tandem architecture is superior to the two DM parallel architecture; it will be interesting to determine whether a dominant architecture also exists for the teams which consist of three DMs. It is not hard to construct proofs, similar to the proof of Lemma 1, to show that the parallel and asymmetrical architectures cannot perform better than the *V*-architecture. Therefore, to determine the dominant three DM architecture, if such exists, we have to compare the *V*-architecture and the tandem architecture.

In several places in the literature it is conjectured and supported with numerical results that the *V*-architecture is better than the tandem architecture [E82], [R87], [RN87a]. But, consider these two architectures when the primary DM is a very bad DM, that is when the primary DM's observation is extremely unreliable; the *V*-architecture for all practical purposes reduces to the two

---

[1] The solutions for the optimal decision rules of three DM tandem architecture and of the *V*-architecture have appeared in [E82], [R87] and [RN87a].

DM parallel architecture (Figure 1(b)), while the three DM tandem architecture reduces to the two DM tandem architecture (Figure 1(a)). Then, according to Lemma 1, in that particular case the tandem architecture is superior to the *V*-architecture. Thus, the comparison of the three DM tandem and the *V*-architecture depends on the particular DMs involved.

Subsequently, it was suggested that the 'best' *V*-architecture achieves superior performance to the 'best' tandem architecture, where 'best' refers to the optimal configuration of the DMs in a particular architecture. This suggestion can not be tested in general because, as was shown in section 3, the 'best' optimal configuration of the DMs in a particular architecture depends not only on the DMs of the team, but also on variables external to the team (i.e., costs, prior probabilities). But in the special case where all three DMs of the team are identical, there exists (obviously) only one configuration for each architecture.

We would thus like to compare the *V*-architecture and the three DM tandem architecture for the case where all three DMs are *identical*. Then the problem can be reduced to comparing the performance of these two architectures for DMs whose ROC curve is piecewise linear with at most six line segments (since the tandem architecture has at most five distinct thresholds). Proceeding with analysis similar to section 3, we obtain the following examples.

If the identical DMs receive *binary* observations, the *V*-architecture achieves *centralized* performance, while the tandem team does not; thus in this case the *V*-architecture is superior to the tandem architecture.

But, the tandem architecture which consists of three identical DMs, whose probability distributions are presented in Figure 16, achieves better performance to the *V*-architecture for η = 2.2, and achieves worse performance for η = 1.0 (Table 6). Thus, a dominant architecture for a three DM team may not exist and the optimal architecture may depend on η, even if all the DMs are identical. The team ROC curves for both the *V*-architecture and the tandem architecture can be seen in Figure 17.

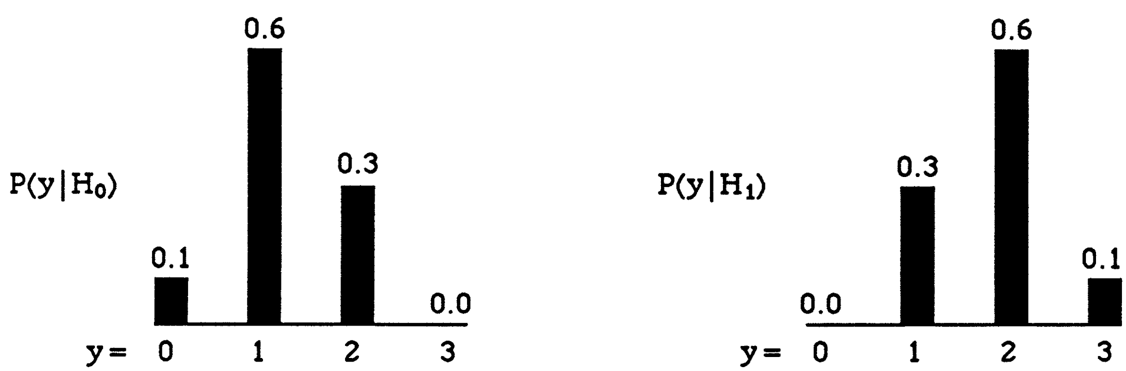


Figure 16. DM for which the Optimal Architecture Depends on η

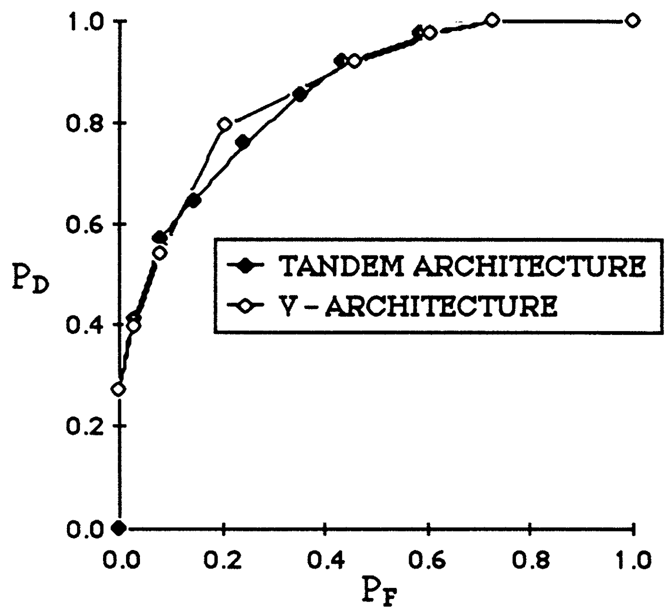


Figure 17(a). Team ROC Curves for DMs of Figure 16: No Dominant Architecture Exists

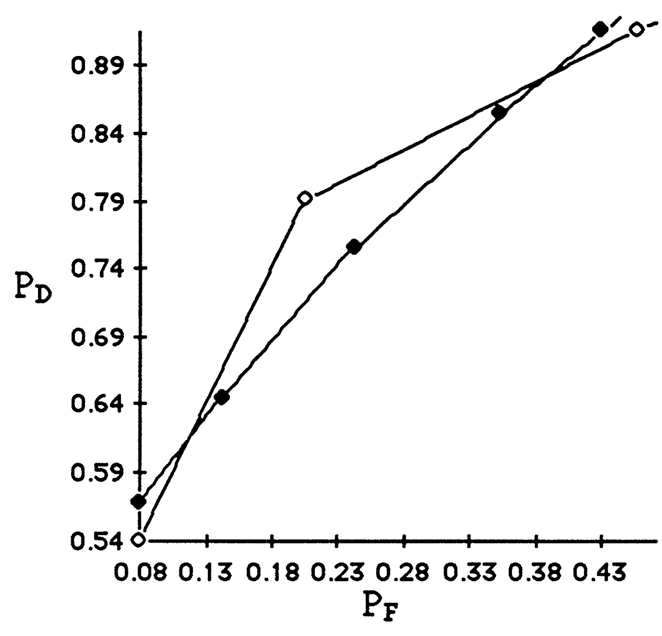


Figure 17(b). Close Up on Figure 17(a)

Table 6. Architecture Comparisons for Teams of Figure 17

(i). η = **2.2** [$P(H_0)$ = 0.3125]	**Tandem**	**Two Consultant (*V*)**
Operating Point of DM C2:	(0.3, 0.7)	(0.3, 0.7)
Operating Point(s) of DM C1:	(0.0, 0.1) if $u_{c2} = 0$	(0.3, 0.7)
	(0.3, 0.7) if $u_{c2} = 1$	
Operating Points of DM P:	(0.0, 0.1) if $u_{c1} = 0$	(0.0, 0.1) if $u_{c1} + u_{c2} = 0$
	(0.9, 1.0) if $u_{c1} = 1$	(0.0, 0.1) if $u_{c1} + u_{c2} = 1$
		(0.9, 1.0) if $u_{c1} + u_{c2} = 2$
Team Operating Point:	(0.081, 0.568)	(0.081, 0.541)
Probability of Error:	0.1906875*	0.199125

(ii). η = **1.0** [$P(H_0)$ = 0.5]	**Tandem**	**Two Consultant (*V*)**
Operating Point of DM C2:	(0.3, 0.7)	(0.3, 0.7)
Operating Point(s) of DM C1:	(0.0, 0.1) if $u_{c2} = 0$	(0.3, 0.7)
	(0.9, 1.0) if $u_{c2} = 1$	
Operating Points of DM P:	(0.0, 0.1) if $u_{c1} = 0$	(0.0, 0.1) if $u_{c1} + u_{c2} = 0$
	(0.9, 1.0) if $u_{c1} = 1$	(0.3, 0.7) if $u_{c1} + u_{c2} = 1$
		(0.9, 1.0) if $u_{c1} + u_{c2} = 2$
Team Operating Point:	(0.243, 0.757)	(0.207, 0.793)
Probability of Error:	0.243	0.207*

* Optimal

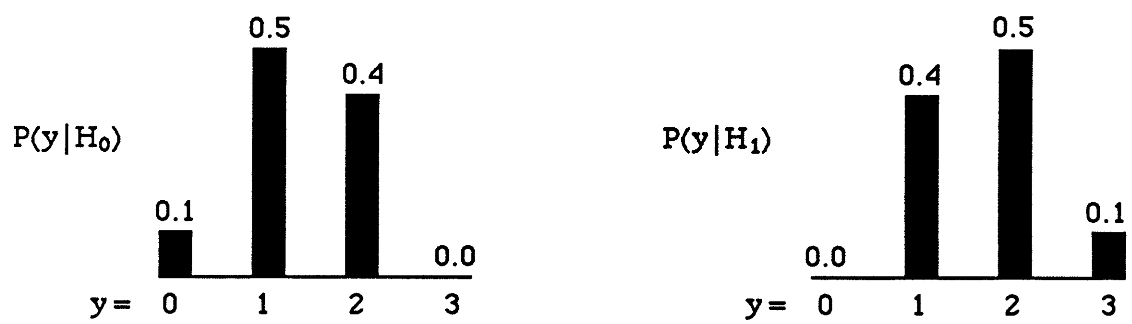


Figure 18. DM resulting in Dominant *Tandem* Architecture

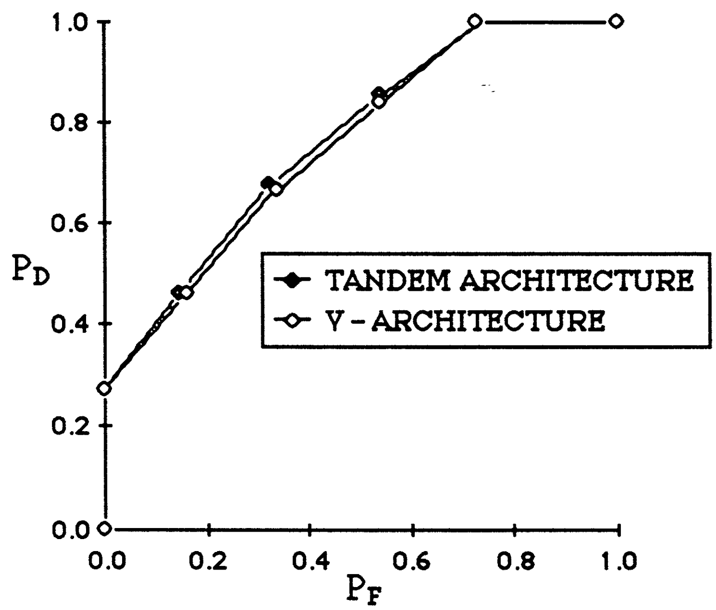


Figure 19. Team ROC Curves for DMs of Figure 18: *Tandem* is the Dominant Architecture

Furthermore, the tandem architecture which consists of the DMs, whose probability distributions are presented in Figure 18, is superior to the *V*-architecture for all values of the threshold η, as can be seen from the team ROC curves of Figure 19. Thus, the tandem architecture may be superior to the *V*-architecture.

Therefore, a "globally" dominant architecture for the teams which consist of three DMs does not exist, even if all the DMs are identical; the optimal architecture depends on the DMs involved *and* on parameters external to the team (i.e., prior probabilities and costs).

6. SUMMARY AND CONCLUSIONS

The architectures of some very simple organizations in a binary hypothesis testing environment were analyzed. The tandem architecture is dominant for teams consisting of two DMs. For teams of three DMs, a dominant architecture does not exist in general. The optimal architecture of an organization depends on parameters external to the team, like the prior probabilities and the cost structure. Nevertheless, there exist special probability distributions for which the optimal architecture can be unequivocally determined.

The problems in this framework should be approached cautiously because counterintuitive results are common. Still, the intuitive solutions, although not necessarily optimal, result in considerable reduction in the complexity of the problem and in relatively good performance. Thus, it may be advisable to sacrifice optimality in favor of simple but reliable suboptimal solutions, which is in accordance with the famous theory of *satisfiability*.

ACKNOWLEDGEMENT

The authors would like to thank Prof. John N. Tsitsiklis for his valuable suggestions.

REFERENCES

[AV89] Al-Ibrahim, M.M., and P.K. Varshney, "A Decentralized Sequential Test with Data Fusion", *Proceedings of the 1989 American Control Conference*, Pittsburg, PA, 1989.

[CV86] Chair, Z., and P.K. Varshney, "Optimum Data Fusion in Multiple Sensor Detection Systems", *IEEE Transactions on Aerospace and Electronic Systems*, AES-22, 1986, pp. 98-101.

[CV88] Chair, Z., and P.K. Varshney, "Distributed Bayesian Hypothesis Testing with Distributed Data Fusion", *IEEE Transactions on Systems, Man and Cybernetics*, SMC-18, 1988, pp. 695-699.

[E82] Ekchian, L.K., "Optimal Design of Distributed Detection Networks", Ph.D. dissertation, Dept. of Elec. Eng. and Computer Science, M.I.T., Cambridge, Ma., 1982.

[ET82] Ekchian, L.K., and R.R. Tenney, "Detection Networks", *Proceedings of the 21st IEEE Conference on Decision and Control*, 1982, pp. 686-691.

[GS66] Green, D.M., and J.A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley and Sons, 1966.

[HV86] Hoballah, I.V., and P.K. Varshney, "Neyman-Pearson Detection with Distributed Sensors", *Proceedings of the 25th Conference on Decision and Control*, Athens, Greece, 1986, pp. 237-241.

[K89] Kastner, M.P., et al., "Hypothesis Testing in a Team: A Normative-Descriptive Study", TR–420, ALPHATECH Inc., Burlington, Ma., 1989.

[P88] Polychronopoulos, G.H., "Solutions of Some Problems in Decentralized Detection by a Large Number of Sensors", TR-LIDS-TH-1781, LIDS, M.I.T.,Cambridge, Ma., June 1988.

[P89] Pothiawala, J., "Analysis of a Two-Sensor Tandem Distributed Detection Network", TR-LIDS-TH-1848, LIDS, M.I.T., Cambridge, Ma., 1989.

[PA86] Papastavrou, J.D., and M. Athans, "A Distributed Hypothesis-Testing Team Decision Problem with Communication Cost", *System Fault Diagnostics, Reliability and Related Knowledge-Based Approaches, Vol. 1*, 1987, pp. 99-130.

[PA88] Papastavrou, J.D., and M. Athans, "Optimum Configuration for Distributed Teams of Two Decision Makers", *Proceedings of the 1988 Symposium on Command and Control Research*, Monterey, Ca., 1988, pp. 162-168.

[PT88] Polychronopoulos, G., and J.N. Tsitsiklis, "Explicit Solutions for Some Simple Decentralized Detection Problems", TR-LIDS-P-1825, LIDS, M.I.T., Cambridge, Ma., 1988.

[R87] Reibman, A.R., "Performance and Fault-Tolerance of Distributed Detection Networks", Ph.D. dissertation, Dept. of Elec. Eng., Duke University, Durham, NC, 1987.

[RN87] Reibman, A.R., and L.W. Nolte, "Optimal Detection and Performance of Distributed Sensor Systems", *IEEE Transactions on Aerospace and Electronic Systems*, AES-23, 1987, pp. 24-30.

[RN87a] Reibman, A.R., and L.W. Nolte, "Design and Performance Comparison of Distributed Detection Networks", *IEEE Transactions on Aerospace and Electronic Systems*, AES-23, 1987, pp. 789-797.

[S86] Sadjadi, F., "Hypothesis Testing in a Distributed Environment", *IEEE Transactions on Aerospace and Electronic Systems*, AES-22, 1986, pp. 134-137.

[S86a] Srinivasan, R., "A Theory of Distributed Detection", *Signal Processing*, 11, 1986, pp. 319-327.

[T86] Tsitsiklis, J.N., "On Threshold Rules in Decentralized Detection", *Proceedings of the 25th Conference on Decision and Control*, Athens, Greece, 1986, pp. 232-236.

[T88] Tsitsiklis, J.N., "Decentralized Detection by a Large Number of Sensors", *Mathematics of Controls, Signals and Systems*, 1, 1988, pp. 167-182.

[T89] Tsitsiklis, J.N., "Decentralized Detection", to appear in *Advances in Statistical Signal Processing, vol. 2: Signal Detection.*

[T89a] Tsitsiklis, J.N., "Extremal Properties of Likelihood-Ratio Quantizers", TR-LIDS-P-1923, LIDS, M.I.T., Cambridge, Ma., 1989.

[TA85] Tsitsiklis, J.N., and M. Athans, "On the Complexity of Decentralized Decision Making and Detection Problems", *IEEE Transactions on Automatic Control*, AC-30, 5, 1985, pp. 440-446.

[TP89] Tang, Z.B., et al., "A Distributed Hypothesis-Testing Problem with Correlated Observations", preprint, 1989.

[TS81] Tenney, R.R. and N.R. Sandell, Jr., "Detection with Distributed Sensors", *IEEE Transactions on Aerospace and Electronic Systems*, AES-17, 4, 1981, pp. 501-510.

[TV87] Thomopoulos, S.C.A., et al., "Optimal Decision Fusion in Multiple Sensor Systems", *IEEE Transaction on Aerospace and Electronic Systems*, AES-23, 1987.

[V68] Van Trees, H.L., *Detection Estimation, and Modulation Theory*, Vol. I, J. Wiley, New York, 1968.

[VT88] Viswanathan, R., et al., "Optimal Serial Distributed Decision Fusion", *IEEE Transactions on Aerospace and Electronic Systems*, AES-24, 1988, pp. 366-375.

APPENDIX A. The ROC Curve and its Properties

Since in our analysis we often employ the ROC curve and its properties, we present a summary of the basic concepts relevant to our research from [V68]. For this define the *likelihood ratio* as:

$$\Lambda(y) \equiv \frac{P(y \mid H_1)}{P(y \mid H_0)} \tag{A.1}$$

The likelihood ratio is a random variable; denote the density of Λ when H is true by $P(\Lambda \mid H)$ and for some threshold η define the *probability of false alarm* as:

$$P_F(\eta) \equiv \int_{\eta}^{\infty} P\left(\Lambda \mid H_0\right) d\Lambda \tag{A.2}$$

and the *probability of detection* as:

$$P_D(\eta) \equiv \int_{\eta}^{\infty} P\left(\Lambda \mid H_1\right) d\Lambda \tag{A.3}$$

The ROC curve is the plot of P_D versus P_F with η as the varying parameter. It is usually defined in an *open-form* expression by the two parametric equations as the following set:

$$\text{ROC} \equiv \{(P_F, P_D) \mid P_F = P_F(\eta), P_D = P_D(\eta), 0 \leq \eta < \infty\}$$

or equivalently in a *close-form* expression, if one exists, as:

$$P_D = f(P_F)$$

The ROC curve is a concave (except for some special esoteric cases [T89]) curve which goes through (0,0) and (1,1). One very important property of the ROC curve, which we explore in our analysis, is that the slope of the tangent to an ROC curve at a particular point is equal to the value of the threshold η required to achieve the P_F and P_D of that point.

The ROC curve is a convenient tool for describing DMs and it is also a convenient tool for ranking DMs. The higher the ROC curve the better the DM since a higher probability of detection corresponds to the same level of probability of false alarm. Thus, we can give the following definition:

DEFINITION A.1. Consider two DMs a and b with associated ROC curves f_a and f_b respectively. We say that DM a is ***better than*** *DM b if:*

$$f_a(P_F) \geq f_b(P_F) \quad \text{for every } P_F : 0 \leq P_F \leq 1$$

and:

$$f_a(\mathrm{P_F}) > f_b(P_F) \quad \text{for some } P_F : 0 \leq P_F \leq 1.$$

Note that, in the cases where the ROC curves of two DMs intersect, the DMs cannot be unequivocally ranked; which DM is better depends on factors which are external to the team, namely the threshold η. We remark that since it is possible to construct the ROC curves for human DM based on empirical data, this allows us to rank human DMs with respect to the specific decision of binary hypothesis testing. Moreover, if a DM desires to minimize the probability of error it must operate at the point of its ROC curve which corresponds to the threshold:

$$\eta \equiv \frac{P(H_0)\,[J(1, H_0) - J(0, H_0)]}{P(H_1)\,[J(0, H_1) - J(1, H_1)]} \tag{A.4}$$

APPENDIX B. Optimal Decision Rules for Two DMs in Tandem

PROBLEM B. A team consisting of two DMs in tandem performs binary hypothesis testing (Figure 1(a) *). The costs $J(u_t, H)$ which are incured by the team when the team decision is u_t and H is the true hypothesis as well as the prior probabilities ($P(H_i) > 0$ for $i = 0, 1$) are assumed to be known. Each DM receives a conditional independent observation. The consulting DM makes a decision based on its observation and transmits a binary message ($u_c = 0$ or $u_c = 1$) to the primary DM. The primary has the responsibility of the final team decision ($u_p = 0$ or $u_p = 1$) which will be based on its own observation and the communication from the consulting DM. The decision rules for the two DMs which minimize the expected cost are to be determined.*

The optimal decision rules have to satisfy the following *necessary* conditions [ET82]. For the primary DM:

$$\text{If } u_c = 0: \qquad \Lambda(y_p) \underset{u_p = 0}{\overset{u_p = 1}{\gtrless}} \frac{1 - P_F^c}{1 - P_D^c}\,\eta \;\equiv \eta_0 \tag{B.1}$$

$$\text{If } u_c = 1: \qquad \Lambda(y_p) \underset{u_p=0}{\overset{u_p=1}{\gtrless}} \frac{P_F^c}{P_D^c}\,\eta \;\equiv \eta_1 \tag{B.2}$$

For the consulting DM:

$$\Lambda(y_c) \underset{u_c=0}{\overset{u_c=1}{\gtrless}} \frac{P_F^1 - P_F^0}{P_D^1 - P_D^0}\,\eta \;\equiv \eta_c \tag{B.3}$$

where η was defined in (A.4), with P_D^i and P_F^i respectively the probability of detection and probability of false alarm for the primary DM when $u_c = i$ was send by the consulting DM ($i = 0$, 1) and, P_D^c and P_F^c respectively the probability of detection and probability of false alarm for the consulting DM.

REMARK 1. It should be clear by now that the decision thresholds of the two DMs are given by a set of *coupled* equations. For example:

$$P_F^0 = P\left(\Lambda(y_p) \geq \frac{1-P_F^c}{1-P_D^c}\,\eta \;\middle|\; H_0\right) \tag{B.4}$$

REMARK 2. The two messages assigned to the consulting DM do not have to be denoted 0 and 1. For that matter they can be denoted m_1 and m_2. Without loss of generality assume that:

$$\frac{P(u_c = m_1 \mid H_0)}{P(u_c = m_1 \mid H_1)} \geq \frac{P(u_c = m_2 \mid H_0)}{P(u_c = m_2 \mid H_1)} \tag{B.5}$$

Then it can be shown that when the primary DM receives $u_c = m_1$, it will always be more likely to decide $u_p = 0$ and when he receives $u_c = m_2$, it will always be more likely to decide $u_p = 1$. Hence the interpretation of m_1 as 0 and of m_2 as 1.

REMARK 3. The ROC curve of the team as a whole can be computed and is given by the following two parametric equations:

$$P_F^t(\eta) = \left[1 - P_F^c(\eta)\right] P_F^0(\eta) + P_F^c(\eta)\, P_F^1(\eta) \tag{B.6}$$

$$P_D^t(\eta) = \left[1 - P_D^c(\eta)\right] P_D^0(\eta) + P_D^c(\eta)\, P_D^1(\eta) \tag{B.7}$$

Note that the team ROC curve depends not only upon the characteristics ("expertise") of the individual DMs, but also on the particular way that they have been constrained to interact (the team or organization architecture).

APPENDIX C. The Analysis of a Two DM Tandem Team

The discussion below may seem confusing, but in fact it is simple and straightforward. We compare the performance of four teams for a particular value of the decision threshold; Figure C.1 should provide a brief summary.

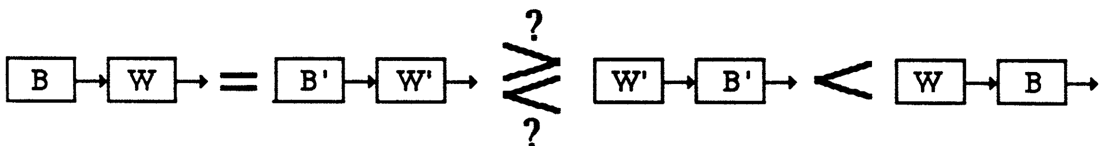


Figure C.1. Summary of the Analysis

Consider two DMs, a DM B better than a DM W in the ROC curve sense and suppose that DM B has been designated as the consulting DM. The solution to the problem for some η^* can be described by the three operating points (P_F^0, P_D^0) and (P_F^1, P_D^1) of the primary DM W and (P_F^c, P_D^c) of the consulting DM B, which have been defined in appendix B above. Consider also a DM W' whose ROC curve consists of the points (0,0), (P_F^0, P_D^0), (P_F^1, P_D^1) and (1,1) as well as the line segments joining them. Finally consider a DM B' whose ROC curve consists of the points (0,0), (P_F^c, P_D^c) and (1,1), the line segments joining them and whose ROC curve lies **above** the ROC curve of DM W'. There are six different cases for the shape of the ROC curve of DM B' depending on where (P_F^c, P_D^c) lies with respect to (P_F^0, P_D^0) and (P_F^1, P_D^1) as can be seen in Figure C.2.

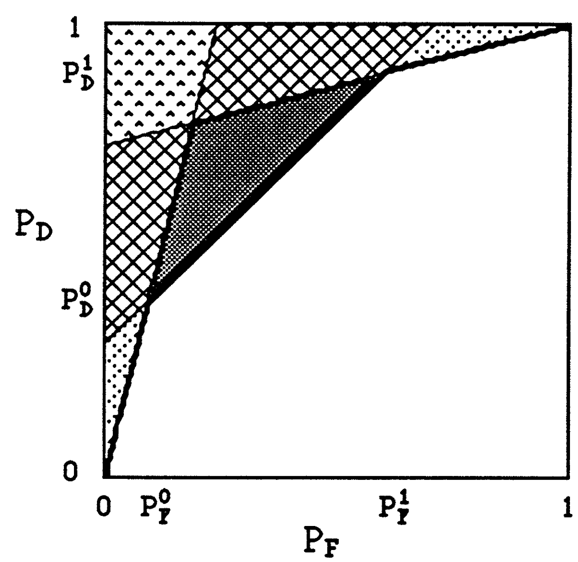


Figure C.2. The Six Cases for the Consulting DM

Consider the 'restricted' team with DM W' as the primary DM and DM B' as the consulting DM. Note that the performance of this team will never be better than the performance of the original team since DM W' is worse than DM W and DM B' is worse than DM B by construction; but for the particular η^* both teams achieve the same performance since the three optimal operating points are attainable by the respective DMs of both teams.

We would now like to compare the performance of the 'restricted' team with the performance of the 'reverse' restricted team which consists of DM B' as the primary DM and DM W' as the consulting DM. Note the the performance of the 'reverse' team is obviously never better than the performance of the team which consists of DM B as the primary DM and DM W as the consulting DM.

REMARK. Henceforth, for the sake of convenience, we will refer to each team by its primary DM since it is clear which DM is its "complementary" consulting DM.

So we compare the 'restricted' team with DM W' as the primary DM to the 'reverse' team with DM B' as the primary DM for all six cases of Figure C.2. If the 'reverse' team performed better than the 'restricted' team in all six cases, we would conclude that the team with DM B as its primary DM would perform better than the team with DM W as its primary DM; hence the conjecture would have been proven true. But, there exist certain cases where the 'restricted' performs better than the 'reverse' team; hence, Example 1 is obtained.

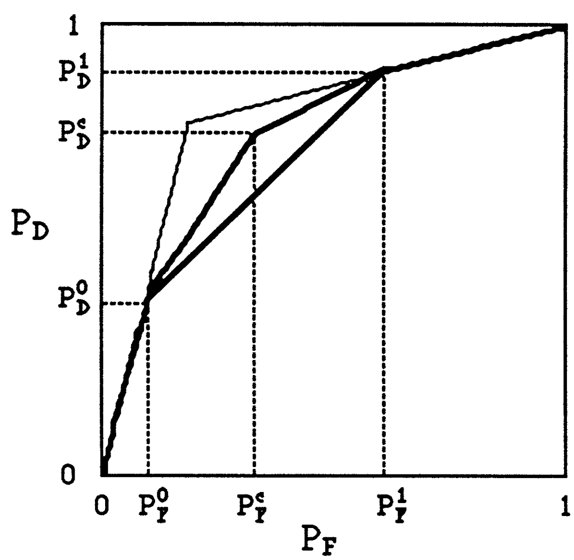


Figure C.3. The 'Restricted' DMs of the Analysis

We need only analyze the one case of Figure C.2, which is presented in more detail in Figure C.3. This is done to demonstrate the method for the analysis of such problems. From Figure C.3 we see that in this case the slopes of the operating points have to satisfy the following conditions:

$$\frac{P_D^0}{P_F^0} \geq \frac{P_D^c}{P_F^c} \geq \frac{P_D^1}{P_F^1} \geq \frac{P_D^1 - P_D^0}{P_F^1 - P_F^0} \geq \frac{1 - P_D^0}{1 - P_F^0} \geq \frac{1 - P_D^c}{1 - P_F^c} \geq \frac{1 - P_D^1}{1 - P_F^1} \tag{C.1}$$

and:

$$\frac{P_D^c}{P_F^c} \geq \frac{P_D^c - P_D^0}{P_F^c - P_F^0} \geq \frac{P_D^1 - P_D^0}{P_F^1 - P_F^0} \geq \frac{P_D^1 - P_D^c}{P_F^1 - P_F^c} \geq \frac{1 - P_D^c}{1 - P_F^c} \tag{C.2}$$

Moreover recalling that the operating points are by construction optimal for the 'restricted' team for η^*, the necessary optimality conditions for the primary DM can be written as:

$$\frac{P_D^0}{P_F^0} \geq \frac{1 - P_F^c}{1 - P_D^c}\eta^* \geq \frac{P_D^1 - P_D^0}{P_F^1 - P_F^0} \tag{C.3}$$

and:

$$\frac{P_D^1 - P_D^0}{P_F^1 - P_F^0} \geq \frac{P_F^c}{P_D^0}\eta^* \geq \frac{1 - P_D^1}{1 - P_F^1} \tag{C.4}$$

and for the consulting DM as:

$$\frac{P_D^c - P_D^0}{P_F^c - P_F^0} \geq \frac{P_F^1 - P_F^0}{P_D^1 - P_D^0}\eta^* \geq \frac{P_D^1 - P_D^c}{P_F^1 - P_F^c} \tag{C.5}$$

(C.3)–(C.5) can be summarized with the help of (C.1) and (C.2) as:

$$min\left\{\frac{1 - P_D^c}{1 - P_F^c}\frac{P_D^0}{P_F^0}, \frac{P_D^c - P_D^0}{P_F^c - P_F^0}\frac{P_D^1 - P_D^0}{P_F^1 - P_F^0}\right\} \geq \eta^* \tag{C.6a}$$

$$\eta^* \geq max\left\{\frac{P_D^c}{P_F^c}\frac{1 - P_D^1}{1 - P_F^1}, \frac{P_D^1 - P_D^c}{P_F^1 - P_F^c}\frac{P_D^1 - P_D^0}{P_F^1 - P_F^0}\right\} \tag{C.6b}$$

We should note here that the probabilities of false alarm and detection for the 'restricted' team are given by:

$$P_F^t = (1 - P_F^c)P_F^0 + P_F^c P_F^1 \tag{C.7}$$

and:

$$P_D^t = (1 - P_D^c)P_D^0 + P_D^c P_D^1 \tag{C.8}$$

We now consider the 'reverse' team. The two possibly optimal operating points for the new consulting DM W' are (P_F^0, P_D^0) and (P_F^1, P_D^1).

CASE I. Operating Point of the 'Reversed' Consulting DM: (P_F^0, P_D^0)

That is assume from (C.3) that:

$$\frac{P_D^0}{P_F^0} \geq \eta_c \geq \frac{P_D^1 - P_D^0}{P_F^1 - P_F^0} \tag{C.9}$$

Then from (C.1):

$$\eta_0 = \frac{1 - P_F^0}{1 - P_D^0}\,\eta^* \tag{C.10}$$

and from (C.2):

$$\eta_1 = \frac{P_F^0}{P_D^0}\,\eta^* \tag{C.11}$$

Combining (C.1), (C.6) and (C.10) we obtain that:

$$\eta_0 \geq \frac{P_D^1 - P_D^c}{P_F^1 - P_F^c} \tag{C.12}$$

Thus the optimal operating point for the new primary DM B' when $u_c = 0$ is received is either (P_F^0, P_D^0) or (P_F^c, P_D^c).

Similarly combining (C.6) and (C.11) we obtain that:

$$\frac{1 - P_D^c}{1 - P_F^c} \geq \eta_1 \tag{C.13}$$

Thus the optimal operating point for the new primary DM B' when $u_c = 1$ is received is either (P_F^1, P_D^1) or (1,1). We have four subcases for the operating points of the primary DM B' which we examine separately.

SUBCASE I.1. Operating Point of the 'Reversed' Primary DM when $u_c = 0$: (P_F^c, P_D^c)

Operating Point of the 'Reversed' Primary DM when $u_c = 1$: (P_F^1, P_D^1)

Then the optimality conditions can be written as:

$$\frac{P_D^c - P_D^0}{P_F^c - P_F^0} \geq \eta_0 \geq \frac{P_D^1 - P_D^c}{P_F^1 - P_F^c}$$

$$\Rightarrow \quad \frac{1-P_D^0}{1-P_F^0}\,\frac{P_D^c-P_D^0}{P_F^c-P_F^0} \;\geq\; \eta^* \;\geq\; \frac{1-P_D^0}{1-P_F^0}\,\frac{P_D^1-P_D^c}{P_F^1-P_F^c} \tag{C.14}$$

and:

$$\frac{P_D^1-P_D^c}{P_F^1-P_F^c} \;\geq\; \eta_1 \;\geq\; \frac{1-P_D^1}{1-P_F^1}$$

$$\Rightarrow \quad \frac{P_D^0}{P_F^0}\,\frac{P_D^1-P_D^c}{P_F^1-P_F^c} \;\geq\; \eta^* \;\geq\; \frac{P_D^0}{P_F^0}\,\frac{1-P_D^1}{1-P_F^1} \tag{C.15}$$

Also from (C.3):

$$\eta_c = \frac{P_F^1-P_F^c}{P_D^1-P_D^c}\,\eta^* \tag{C.16}$$

In view of this we check our assumption of (C.9):

$$\frac{P_D^0}{P_F^0} \;\geq\; \frac{P_F^1-P_F^c}{P_D^1-P_D^c}\,\eta^* \;\geq\; \frac{P_D^1-P_D^0}{P_F^1-P_F^0} \tag{C.17}$$

This is indeed true from (C.2) and (C.6).

In this subcase the probabilities of false alarm and detection for the 'reversed' team are given by:

$$P_F^v = \left(1-P_F^0\right)P_F^c + P_F^0 P_F^1 \tag{C.18}$$

and:

$$P_D^v = \left(1-P_D^0\right)P_D^c + P_D^0 P_D^1 \tag{C.19}$$

Comparing the probabilities of false alarm for the 'restricted' and the 'reversed' teams:

$$P_F^t = \left(1-P_F^c\right)P_F^0 + P_F^c P_F^1 \;\leq\; \left(1-P_F^0\right)P_F^c + P_F^0 P_F^1 = P_F^v \tag{C.20}$$

For the 'reversed' architecture to be better:

$$\frac{P_D^v-P_D^t}{P_F^v-P_F^t} = \frac{P_D^c-P_D^0}{P_F^c-P_F^0}\,\frac{1-P_D^1}{1-P_F^1} \;\geq\; \eta^* \;\geq\; \frac{1-P_D^v}{1-P_F^v} \tag{C.21}$$

The left inequality of (C.21) does **not** hold, as it is in contradiction with the right inequality of (C.15). Therefore, the 'restricted' team results in superior performance in this subcase.

SUBCASE I.2. *Operating Point of the 'Reversed' Primary DM when* $u_c = 0$: (P_F^c, P_D^c)

Operating Point of the 'Reversed' Primary DM when $u_c = 1$: $(1,1)$

Then the optimality conditions can be written as (C.16) again and:

$$\frac{1-P_D^1}{1-P_F^1} \geq \eta_1 \quad \Rightarrow \quad \frac{P_D^0}{P_F^0}\frac{1-P_D^1}{1-P_F^1} \geq \eta^* \tag{C.22}$$

Also from (C.3):

$$\eta_c = \frac{1-P_F^c}{1-P_D^c}\,\eta^* \tag{C.23}$$

In view of this we check our assumption of (C.9):

$$\frac{P_D^0}{P_F^0} \geq \frac{1-P_F^c}{1-P_D^c}\,\eta^* \geq \frac{P_D^1-P_D^0}{P_F^1-P_F^0} \tag{C.24}$$

This is indeed true from (C.2) and (C.6).

In this subcase the probabilities of false alarm and detection for the 'reversed' team are given by:

$$P_F^v = \left(1-P_F^0\right)P_F^c + P_F^0\,1 \tag{C.25}$$

and:

$$P_D^v = \left(1-P_D^0\right)P_D^c + P_D^0\,1 \tag{C.26}$$

Comparing the probabilities of false alarm for the 'restricted' and the 'reversed' teams:

$$P_F^t = \left(1-P_F^c\right)P_F^0 + P_F^c P_F^1 \leq \left(1-P_F^0\right)P_F^c + P_F^0 = P_F^v \tag{C.27}$$

For the 'reversed' architecture to be better:

$$\frac{P_D^v-P_D^t}{P_F^v-P_F^v} = \frac{P_D^c}{P_F^c}\frac{1-P_D^1}{1-P_F^1} \geq \eta^* \geq \frac{1-P_D^v}{1-P_F^v} \tag{C.28}$$

The left inequality of (C.28) does **not** hold, as it is in contradiction with the right inequality of (C.6b). Therefore, the 'restricted' team results in superior performance in this subcase.

SUBCASE I.3. *Operating Point of the 'Reversed' Primary DM when $u_c = 0$:* (P_F^0, P_D^0)
Operating Point of the 'Reversed' Primary DM when $u_c = 1$: (P_F^1, P_D^1)

We can immediately state that the 'restricted' team results in superior performance in this subcase. Otherwise the optimality of the operating points (P_F^0, P_D^0) and (P_F^1, P_D^1) of DM W' and (P_F^c, P_D^c) of DM B' would be contradicted, because the operating points (P_F^0, P_D^0) and (P_F^1, P_D^1) can be employed by DM W' and (P_F^0, P_D^0) can be employed by DM B'.

SUBCASE I.4. *Operating Point of the "Reversed' Primary DM when* $u_c = 0$: (P_F^0, P_D^0)
Operating Point of the 'Reversed' Primary DM when $u_c = 1$: $(1,1)$

We can state that the 'restricted' team results in superior performance in this subcase for the same reasons as in Subcase I.3 above.

CASE II. Operating Point of the 'Reversed' Consulting DM: (P_F^1, P_D^1)

The analysis is omitted since it is exactly analogous to the analysis of Case I above; again we conclude that the 'restricted' team results in superior performance.

Hence we can construct two DMs as the ones in Figure C.3 knowing that the tandem team with the better DM as the consulting DM will achieve superior performance.

APPENDIX D. The Optimal Decision Rules for the Two DM Parallel Team

PROBLEM D. A team consisting of two DMs, DM a and DM b, in parallel performs binary hypothesis testing (Figure 1(b)). *The costs* $J(u_t, H)$ *which are incured by the team when the decision of the team is* u_t *and H is the true hypothesis as well as the prior probabilities* ($P(H_i)$ *for i* = 0, 1) *are assumed to be known. Each DM receives a conditional independent observation, makes a decision based on its observation and transmits a binary message to a fusion center. The fusion center employs a maximum a posteriori probability (MAP) decision rule. The decision rules for the two DMs which minimize the expected cost are to be determined.*

The fusion center has two alternative choises for its decision rule; either employ the AND decision rule (i.e., decide $u_t = 1$ if $u_a = 1$ *and* $u_b = 1$) or employ the OR decision rule (i.e., decide $u_t = 1$ if $u_a = 1$ *or* $u_b = 1$). If the fusion center is employing the AND decision rule, the optimal decision rules of the DMs have to satisfy the following *necessary* conditions [E84]. For DM a:

$$\Lambda(y_a) \underset{u_a=0}{\overset{u_a=1}{\gtrless}} \frac{P_F^b}{P_D^b}\,\eta \;\equiv \eta_a \tag{D.1}$$

and for DM b:

$$\Lambda(y_b) \underset{u_b=0}{\overset{u_b=1}{\gtrless}} \frac{P_F^a}{P_D^a}\,\eta \;\equiv \eta_b \tag{D.2}$$

If the fusion center is employing the OR decision rule, the optimal decision rules of the DMs have to satisfy the following *necessary* conditions [E84]. For DM a:

$$\Lambda(y_a) \underset{u_a=0}{\overset{u_a=1}{\gtrless}} \frac{1-P_F^b}{1-P_D^b}\eta \equiv \eta_a \tag{D.3}$$

and for DM b:

$$\Lambda(y_b) \underset{u_b=0}{\overset{u_b=1}{\gtrless}} \frac{1-P_F^a}{1-P_D^a}\eta \equiv \eta_b \tag{D.4}$$

and η was defined in (A.4) with P_D^n and P_F^n respectively the probability of detection and of false alarm of DM n, for $n = a, b$.

REMARK 1. The decision thresholds of the two DMs are given by a set of *coupled* equations. For example, if the AND decision rule is employed by the fusion center:

$$P_D^b = P\left(\Lambda(y_b) \geq \frac{P_F^a}{P_D^a}\eta \,\middle|\, H_1\right) \tag{D.5}$$

REMARK 2. The ROC curve of the team as a whole can be computed; for example, if the AND decision rule is employed by the fusion center. the ROC curve of the team is given by the following two parametric equations:

$$P_F^t(\eta) = P_F^a(\eta)\, P_F^b(\eta) \tag{D.6}$$

$$P_D^t(\eta) = P_D^a(\eta)\, P_D^b(\eta) \tag{D.7}$$

APPENDIX E. Obtaining the Bounds for the Two DM Parallel Team

(i). Absolute Boumd

Suppose that (P_F^a, P_D^a) is given; then the problem of maximizing the least upper bound of (32a)-(34a) can be formulated as follows:

PROBLEM E.1.
$$\max_{(P_F^b, P_D^b) \in A} \Delta P_1^e = \frac{\eta^*}{\eta^*+1}\left[-P_F^a(P_F^b - P_F^a)\right] + \frac{1}{\eta^*+1}\left[P_D^a(P_D^b - P_D^a)\right] \tag{E.1}$$

where A is the set that is defined by all the pairs (P_F^b, P_D^b) satisfying the following four constraints:

$$(P_D^b)^2 = \frac{P_D^a}{P_F^a}\frac{1-P_D^a}{1-P_F^a}(P_F^b)^2 + \frac{P_D^a(P_D^a - P_F^a)}{1-P_D^a} \tag{E.2}$$

$$\frac{P_D^b}{P_F^b} \leq \frac{P_D^a}{P_F^a} \tag{E.3}$$

$$\frac{1-P_D^b}{1-P_F^b} \leq \frac{1-P_D^a}{1-P_F^a} \tag{E.4}$$

$$0 \leq P_F^b \leq P_D^b \leq 1 \tag{E.5}$$

with:

$$\eta^* = \frac{P_D^a}{P_F^a}\frac{1-P_D^a}{1-P_F^a} \tag{E.6}$$

(E.3)-(E.5) are constraints dictated by the concavity of the ROC curve. We solve (E.2) for $P_D^b \geq 0$ and substitute into (E.1); we then take the derivative of ΔP_1^e with respect to P_F^b:

$$\frac{d\Delta P_1^e}{dP_F^b} = \frac{\eta^*}{\eta^*+1}\left[-P_F^a + P_D^a P_F^b\left[\eta^*(P_F^b)^2 + \frac{P_D^a(P_D^a-P_F^a)}{1-P_F^a}\right]^{-0.5}\right] \tag{E.7}$$

The derivative in (E.7) is zero only if $P_F^b = P_F^a$, which results in a *minimum* ($\Delta P_1^e = 0$). Differentiating one more time both sides of (E.7), we obtain:

$$\frac{d^2\Delta P_1^e}{d(P_F^b)^2} = \frac{\eta^*}{\eta^*+1}P_D^a\left[\eta^*(P_F^b)^2 + \frac{P_D^a(P_D^a-P_F^a)}{1-P_F^a}\right]^{-1.5}\frac{P_D^a(P_D^a-P_F^a)}{1-P_F^a} > 0 \tag{E.8}$$

Thus, for all P_F^b, such that $P_F^b \geq P_F^a$, the derivative of ΔP_1^e with respect to P_F^b is positive and increasing; therefore, it is maximized with respect to (P_F^b, P_D^b) at the boundary point given by (40).

(ii). Relative Boumd

Suppose again that (P_F^a, P_D^a) is given; then the problem of maximizing the least upper bound of (45)-(47) can be formulated as follows:

PROBLEM E.2.

$$\max_{(P_F^b, P_D^b)\in A} \quad \omega_1^e = \frac{\eta^*(P_F^a)^2 + 1 - (P_D^a)^2}{\eta^* P_F^a P_F^b + 1 - P_D^a P_D^b} - 1 \tag{E.9}$$

where A is the same feasible set as in Problem E.1.

Proceeding in a similar manner as in *(i)* above, we show that the relative bound is maximized with respect to (P_F^b, P_D^b) again at the boundary point given by (40). To see this note that the maximization in (E.9) above is obviously equivalent to:

$$\max_{(P_F^b, P_D^b) \in A} \quad \psi(P_F^b, P_D^b) = \frac{\eta^* (P_F^a)^2 + 1 - (P_D^a)^2}{\eta^* P_F^a P_F^b + 1 - P_D^a P_D^b} \tag{E.10}$$

The function ψ is well defined and non-zero for all the pairs (P_F^b, P_D^b), since both the numerator and the denominator are always strictly positive. Then, the maximization of (E.10) is equivalent to:

$$\min_{(P_F^b, P_D^b) \in A} \quad \xi(P_F^b, P_D^b) = \frac{\eta^* P_F^a P_F^b + 1 - P_D^a P_D^b}{\eta^* (P_F^a)^2 + 1 - (P_D^a)^2} \tag{E.11}$$

This in turn is equivalent to:

$$\max_{(P_F^b, P_D^b) \in A} \quad -\xi(P_F^b, P_D^b) = \frac{-\eta^* P_F^a P_F^b - 1 + P_D^a P_D^b}{\eta^* (P_F^a)^2 + 1 - (P_D^a)^2} \tag{E.12}$$

But barring constant terms, the maximization of (E.12) is the exact same maximixation of (E.1). Since (E.1) is maximized for (P_F^b, P_D^b) given by (40), both (E.12) and consequently (E.9) are also maximized for (P_F^b, P_D^b) given by (40).