# Robust decentralized detection by asymptotically many sensors ⋆

Hakan Deliç, P. Papantoni-Kazakos [1]

*Department of Electrical Engineering, Thornton Hall, University of Virginia, Charlottesville, Virginia 22903-2442, USA*

Received 25 October 1991; revised 19 November 1992

**Abstract**

We consider a decentralized hypothesis testing structure with asymptotically many sensors, each collecting a single datum. The sensors deploy robust test functions that are designed for outlier classes of hypotheses. The sensor outputs are transmitted to the fusion center for the global decision. In this paper, we concentrate on sensor-level decision making, and study the asymptotic performance of the decentralized detection scheme described above. In particular, we utilize the asymptotic relative efficiency performance measure, defined as the ratio of the number of sensors needed by the decentralized structure over the number of data needed by the centralized one to attain the same performance level. Our results indicate the superiority and the necessity of using robust statistical techniques.

*Keywords:* Decentralized detection; robust test function; saddle point game; asymptotics.

⋆ *Signal Processing*, Vol. 33, No. 2, pp. 223-233, August 1993.

⋆⋆ This work was supported in part by the NSF grant NCR-9010302.

[1] Correspondence to: Dr. Hakan Deliç, Department of Neurological Surgery, Health Sciences Center Box 420, Univeristy of Virginia, Charlottesville, VA 22908, USA. Tel.: (804) 924-5014; fax (804) 982-3829; e-mail: mhd2u@virginia.edu

## 1 Introduction

Decentralized detection systems consist of a number of distributed sensors that transmit their decisions to a fusion center, where the global decision is declared based on these. Decentralized systems have several advantages over the centralized ones, such as reduced communication bandwidth requirements, improved computational efficiency, fault-tolerance and the ability to employ different sensing techniques.

Most of the research on distributed detection based on multisensor data is limited to parametrically defined hypotheses, where the statistics of the observations are assumed to be completely known a priori. In general, a Bayesian [4] or Neyman-Pearson [3] type performance criterion is used to find the optimal decision rules at the sensor level. For an overview of optimal decision fusion rules, see [11].

In realistic situations, however, only partial or incomplete knowledge of the observation statistics may be available to the system designer. Detection schemes that are designed for some specific observation model may suffer drastic performance degradation even for small deviations from the nominal assumption [12]. Based on this observation, robust statistics has emerged over the past twenty-five years, providing techniques with good performance under nominal conditions and acceptable performance for signal and noise processes that deviate from the nominal within classes of possible characteristics. The theory of robust statistics can be found in [6]; a survey on applications to signal processing and communications is presented in [8].

Regarding decentralized detection, in [2], robust data fusion rules are derived for general uncertainty classes described by 2-alternating Choquet capacities [7], using error probability as the performance criterion. In this paper, we consider a decentralized detection scheme with asymptotically many sensors, each collecting a single datum. We concentrate on sensor-level decision making, where the sensors are assumed to deploy robust test functions, and utilize the asymptotic relative efficiency as the performance measure to compare the centralized and robust decentralized systems.

## 2 Preliminaries

We consider a decentralized detection problem where sensors receive data that are nonparametrically defined; that is, each hypothesis is generated by a class of processes rather than a single known process. We assume that the classes of processes that generate the hypotheses are memoryless and stationary. Then, each hypothesis is equivalently determined by a class of first order density functions. We will denote by $F_k$, $k = 0, 1$, the class of densities corresponding to hypothesis $H_k$, $k = 0, 1$. Then, $f_1 \in F_1$ and $f_0 \in F_0$. In the case that the two hypotheses are distinguished by a shift in location parameter, it suffices to consider the class $F_0$ only, which will be then denoted by $F$, and $f_0 \in F_0$ will be denoted by $f \in F$.

Looking into the asymptotics of the decentralized structure and focusing on the asymptotic relative efficiency performance measure, we will discuss the case with asymptotically many sensors and finite data only. For this case, we define the asymptotic relative efficiency (ARE) as the ratio of the number of sensors needed by the decentralized structure over the number of data needed by the centralized structure, to attain identical performance levels. The asymptotically many data and finite sensors case is not considered in this paper, since the ARE is not a function of the distributions then [1]. Thus, throughout the paper, the term "decentralized structure" will imply large number of sensors, each collecting single datum.

Let us now return to the case where two classes, $F_0$ and $F_1$, are given, and all that is known is that $f_1 \in F_1$ and $f_0 \in F_0$. The issue then is the design of the decision rules, $\delta(N)$, per sensor.

Let false alarm probability be defined as the probability of deciding on $H_1$, given that $H_0$ is true; and, let power probability be defined as the probability of deciding on $H_1$, given that $H_1$ is true. Given false alarm rate $\alpha$ per sensor, we select as performance criterion the power probability, $P(\delta(N),\, f_1)$, attained by each sensor at time $n$ when the decision rule $\delta(N)$ is deployed and the data are generated by the density $f_1 \in F_1$. We then consider a saddle point game. That is, we search for a pair $(\delta^*(N),\, f_1^*)$ such that

$$P(\delta(N),\, f_1^*) \leq P(\delta^*(N),\, f_1^*) \leq P(\delta^*(N),\, f_1), \forall f_1 \in F_1, \forall \delta(N), \tag{1a}$$

subject to

$$P(\delta^*(N),\, f_0) \leq \alpha, \forall f_0 \in F_0. \tag{1b}$$

Next, we consider the asymptotics of the game in (1). That is, we assume $N \to \infty$, when the system attains its "best" performance. The asymptotic solution of the game in (1) determines the pair $(f_0^*,\, f_1^*)$, $f_0^* \in F_0$ and $f_1^* \in F_1$, of densities at which the decision rules $\delta^*(N)$ are designed, and it is known [9]. In particular, the test function, $T^*(N,\, x^N)$, of the decision rule $\delta^*(N)$ is as follows:

$$T^*(N,\, x^N) = \frac{1}{N} \sum_{i=1}^{N} \log \frac{f_1^*(x_i)}{f_0^*(x_i)}, \tag{2}$$

where $x^N = \{x_i;\, i = 1, \ldots,\, N\}$, and $f_1^*$and $f_0^*$ are those densities, in $F_1$ and

$F_0$, that are the "closest" to each other in the Kullback-Leibler (KL) sense. That is,

$$\mathrm{KL}(f_1^*,\, f_0^*) = \int f_1^*(x) \log \frac{f_1^*(x)}{f_0^*(x)} dx = \inf_{f_1 \in F_1} \inf_{f_0 \in F_0} \mathrm{KL}(f_1,\, f_0), \tag{3a}$$

$$\mathrm{KL}(f_0^*,\, f_1^*) = \inf_{f_1 \in F_1} \inf_{f_0 \in F_0} \mathrm{KL}(f_0,\, f_1). \tag{3b}$$

The pair $(f_1^*,\, f_0^*)$ determined by (3) is the least favorable in $F_1 \times F_0$. It attains the worst power performance at each time $N$, or, equivalently, it induces the slowest convergence rate to a prespecified power level. The specific form of the densities $(f_0^*,\, f_1^*)$ in (3) is determined by the specific definition of the classes $F_0$ and $F_1$. Two interesting classes arise when weak topologies [10] around two distinct nominal known densities are considered; these classes also represent outlier models. The weak topology classes are defined as follows [5]:

> Let $g_0(x)$ and $g_1(x)$ be two distinct and known densities, called nominal. Let $\varepsilon_0$ and $\varepsilon_1$ be two constants in $(0,\, 1)$. Then, for $H$ denoting the class of all first order density functions, define
>
> $$F_0 = \{f_0 :\, f_0 = (1 - \varepsilon_0) g_0 + \varepsilon_0 h;\; h \in H\}, \tag{4a}$$
>
> $$F_1 = \{f_1 :\, f_1 = (1 - \varepsilon_1) g_1 + \varepsilon_1 h;\; h \in H\}. \tag{4b}$$

For the classes in (4), the pair $(f_1^*,\, f_0^*)$ in (3) is found to be [5]

$$\log \frac{f_1^*(x)}{f_0^*(x)} = \begin{cases} \log \left[ c_1 \frac{1-\varepsilon_1}{1-\varepsilon_0} \right] & \text{if } \frac{g_1(x)}{g_0(x)} \le c_1 \\ \log \left[ \frac{1-\varepsilon_1}{1-\varepsilon_0} \frac{g_1(x)}{g_0(x)} \right] & \text{if } c_1 < \frac{g_1(x)}{g_0(x)} < c_0 \\ \log \left[ c_0 \frac{1-\varepsilon_1}{1-\varepsilon_0} \right] & \text{if } \frac{g_1(x)}{g_0(x)} \ge c_0, \end{cases} \tag{5}$$

where $c_0$ and $c_1$ are uniquely defined constants. The expression in (5) defines the robust test function deployed by each sensor when the weak topology classes, also called the outlier model classes, in (4) are present. The robust test function has powerful characteristics and protects the system from performance breakdowns when highly erroneous data (outliers) occur. In the interesting special case where $g_0(x)$ is Gaussian with zero mean and variance $\sigma^2$, and $g_1(x) = g_0(x - \theta)$, the expression in (2) reduces to the following form [9]:

$$T^*(N,\, x^N) = \frac{1}{N} \sum_{i=1}^{N} z(x_i)$$

where

$$z(x) = \begin{cases} d_1 \text{ if } x \le d_1 \\ x \;\; \text{if } d_1 < x < d_0 \\ d_0 \text{ if } x \ge d_0. \end{cases} \tag{6}$$

Let $\Phi(x)$ be the cumulative distribution function of zero-mean, unit-variance Gaussian random variable at point $x$. The constants $d_0$ and $d_1$ are solutions to

$$\Phi\left(\frac{d_0}{\sigma}\right) + \exp\left\{-\frac{\theta}{\sigma^2}d_0 + \frac{\theta^2}{2\sigma^2}\right\}\Phi\left(\frac{-d_0+\theta}{\sigma}\right) = \frac{1}{1-\varepsilon_0}$$

and

$$\Phi\left(\frac{-d_1+\theta}{\sigma}\right) + \exp\left\{\frac{\theta}{\sigma^2}d_1 - \frac{\theta^2}{2\sigma^2}\right\}\Phi\left(\frac{d_1}{\sigma}\right) = \frac{1}{1-\varepsilon_1},$$

respectively, and they are unique if $\varepsilon_0 < 0.5$ and $\varepsilon_1 < 0.5$. It is worth observing that the test function in (6) performs data truncations; a phenomenon that results in system protection against outlier data. When $\varepsilon_0 = \varepsilon_1 = \varepsilon < 0.5$, we have $-d_1 = d_0 = d > 0$, where $d$ is such that

$$\Phi\left(\frac{d}{\sigma}\right) + \exp\left\{-\frac{\theta}{\sigma^2}d + \frac{\theta^2}{2\sigma^2}\right\}\Phi\left(\frac{-d+\theta}{\sigma}\right) = \frac{1}{1-\varepsilon}. \tag{7}$$

Using the results summarized above, in the next section we study the performance of the robust test functions when they are deployed by the decentralized detection structure defined earlier in this section.

## 3 Robust test functions for asymptotically many sensors

Let us consider the decentralized structure in the presence of hypotheses generated by stationary and memoryless processes. The sensors in the structure are assumed to be identical, their number, $N$, is large, and each collects a single datum. The main assumption here is that each sensor deploys a robust test function. We will specifically consider the adoption of a particular such test function, whose powerful properties are known and have been thoroughly studied [6, 9]. In particular, let the output of the $j$th sensor be as follows:

$$u_j = \begin{cases} 1 \text{ if } z_d(x_j) \ge \lambda \\ 0 \text{ otherwise,} \end{cases} \tag{8}$$

for $j = 1, \ldots, M$, where, for some $d > 0$, $z_d(x)$ is

$$z_d(x) = \begin{cases} d & \text{if } x \geq d \\ x & \text{if } -d < x < d \\ -d & \text{if } x \leq -d. \end{cases} \tag{9}$$

As we saw, the test function in (9) evolves from the saddle point solution of the game in (1), when the classes $F_0$ and $F_1$ of densities are as in (4) with $\varepsilon_1 = \varepsilon_0 = \varepsilon < 0.5$ and the nominal densities, $g_0$ and $g_1$, are Gaussian and are such that $g_1(x) = g_0(x - \theta) = \sigma^{-1}\phi[(x - \theta)/\sigma]$, where $\phi(x)$ is the density function of the zero-mean, unit-variance Gaussian random variable at point $x$. The test function in (9) has strong merits in terms of performance, and it can be deployed regardless of its origin and the truncation constant $d$ can be then selected ad hoc. The effects of the selection of $d$ on a given system can be quantified by pertinent performance studies.

Given the decentralized structure, we will study the performance of the test function $z_d(x)$ in (9) in terms of asymptotic relative efficiencies. We will do that for the case where the two hypotheses are generated by Gaussian processes and a location parameter shift, where, for asymptotics, the shift is assumed very small. Our result is stated and proven in Lemma 1 below.

**Lemma 1** *Consider the decentralized structure with large number of sensors, where each sensor deploys the test function $z_d(x)$ in (9). Consider also a centralized, single-element structure whose test function, upon the collection of $N$ data, is $N^{-1}\sum_{i=1}^{N} z_d(x_i)$. For both structures, let the acting hypotheses be generated by stationary and memoryless Gaussian processes. In particular, let the density function per datum be $\sigma^{-1}\phi(x/\sigma)$, under the $H_0$ hypothesis, and $\sigma^{-1}\phi[(x - \theta)/\sigma]$, under the $H_1$ hypothesis. Let $\sigma$ be arbitrary and let $\theta < \varepsilon$, where $0 < \varepsilon \ll 1$. The asymptotic relative efficiency of the decentralized structure at the Gaussian process, when the test function $z_d(x)$ is deployed, is defined as the ratio of the number of sensors needed by this structure over the number of data needed by the centralized structure to attain the same false alarm and power rates, when the acting hypotheses are generated by the Gaussian processes described above. This asymptotic relative efficiency is denoted by $\mathrm{ARE}(\infty, z_d, G)$, and it is given below.*

$$\mathrm{ARE}(\infty, z_d, G) = \frac{\pi}{2}\eta\left(\frac{d}{\sigma}\right) \tag{10}$$

*where*

$$\eta(x) = \frac{[2\Phi(x) - 1]^2}{2(1 - x^2)\Phi(x) + 2x^2 - 1 - 2x\phi(x)}. \qquad \square \tag{11}$$

**Proof.** (a) For $\theta \ll 1$, using first order approximations in Taylor series expansions, we will use

$$\phi\left(\frac{x-\theta}{\sigma}\right) \approx \phi\left(\frac{x}{\sigma}\right) + \frac{\theta x}{\sigma^2}\phi\left(\frac{x}{\sigma}\right) \tag{12a}$$

and

$$\Phi\left(\frac{x-\theta}{\sigma}\right) \approx \Phi\left(\frac{x}{\sigma}\right) - \frac{\theta}{\sigma}\phi\left(\frac{x}{\sigma}\right). \tag{12b}$$

For both structures, let the false alarm and power rates, $\alpha$ versus $\beta$, be given. (b) Consider the centralized structure first. The decision at time $N$ is

$$u(N) = \begin{cases} 1 \text{ if } \frac{1}{N}\sum_{i=1}^{N} z_d(x_i) \geq \lambda(N) \\ 0 \text{ if } \frac{1}{N}\sum_{i=1}^{N} z_d(x_i) < \lambda(N), \end{cases} \tag{13}$$

where

$$\frac{1}{n}\sum_{i=1}^{n} z_d(x_i) \longrightarrow_{N\to\infty} \begin{cases} G(m_0,\, \rho_0/\sqrt{N}), \text{ under hypothesis } H_0 \\ G(m_1,\, \rho_1/\sqrt{N}), \text{ under hypothesis } H_1, \end{cases} \tag{14}$$

and where

$$m_0 \triangleq \int \frac{1}{\sigma}\phi\left(\frac{x}{\sigma}\right) z_d(x)dx = 0,$$

$$\begin{aligned} m_1 &\triangleq \int \frac{1}{\sigma}\phi\left(\frac{x-\theta}{\sigma}\right) z_d(x)dx \\ &\approx \int \frac{1}{\sigma}\left[\phi\left(\frac{x}{\sigma}\right) + \frac{\theta x}{\sigma^2}\phi\left(\frac{x}{\sigma}\right)\right] z_d(x)dx \\ &= \theta\left[2\Phi\left(\frac{d}{\sigma}\right) - 1\right], \end{aligned}$$

$$\begin{aligned} \rho_0^2 &\triangleq \int \frac{1}{\sigma}\phi\left(\frac{x}{\sigma}\right) z_d^2(x)dx \\ &= \sigma^2\left\{2\left[1-\left(\frac{d}{\sigma}\right)^2\right]\Phi\left(\frac{d}{\sigma}\right) + 2\left(\frac{d}{\sigma}\right)^2 - 1 - 2\frac{d}{\sigma}\phi\left(\frac{d}{\sigma}\right)\right\}, \end{aligned} \tag{15}$$

$$\begin{aligned} \rho_1^2 &\triangleq \int \frac{1}{\sigma}\phi\left(\frac{x-\theta}{\sigma}\right) z_d^2(x)dx - m_1^2 \\ &\approx \int \frac{1}{\sigma}\left[1 + \frac{\theta x}{\sigma^2}\right]\phi\left(\frac{x}{\sigma}\right) z_d^2(x)dx - m_1^2 \end{aligned}$$

$$= \rho_0^2 - \theta^2 \left[ 2\Phi\left(\frac{d}{\sigma}\right) - 1 \right]^2 \approx \rho_0^2.$$

Unless otherwise specified, the integrals are taken over the entire real line.

Then, the threshold, $\lambda(N)$, in (13) and the power probability, $\beta(N)$, attained by the centralized structure are as follows:

$$\lambda(N): \ \alpha = 1 - \Phi\left(\frac{\lambda(N) - m_0}{\rho_0}\sqrt{N}\right) \Longrightarrow \lambda(N) = \frac{\rho_0}{\sqrt{N}}\Phi^{-1}(1-\alpha), \quad (16)$$

$$\beta(N) = 1 - \Phi\left(\frac{\lambda(N) - m_1}{\rho_1}\sqrt{N}\right)$$
$$= \Phi\left\{\frac{\sqrt{N}}{\rho_1}\theta\left[2\Phi\left(\frac{d}{\sigma}\right) - 1\right] + \Phi^{-1}(\alpha)\right\}, \quad (17)$$

where $\Phi^{-1}(x) = y$ implies $x = \Phi(y)$. Requiring that $\beta(N) = \beta$, and solving expression (17) with respect to $N$, we find for $\eta(x)$ as in (11) that the sample size needed to attain $\alpha$ and $\beta$ is

$$N_c(\alpha, \beta) = [\Phi^{-1}(\beta) - \Phi^{-1}(\alpha)]^2 \frac{\rho_1^2}{\theta^2 \eta(d/\sigma)}. \quad (18)$$

(c) Consider now the decentralized structure, with the number of sensors $M \to \infty$ and single datum per sensor. The decision of each sensor is here as follows.

$$u(x) = \begin{cases} 1 \text{ if } z_d(x) \geq \lambda \\ 0 \text{ otherwise} \end{cases}$$

where

$$p_{\lambda,k} \triangleq P[z_d(x) \geq \lambda | H_k] = \begin{cases} 1 & \text{if } \lambda \leq -d \\ 0 & \text{if } \lambda > d \\ P(x \geq \lambda | H_k) & \text{if } -d < \lambda \leq d, \end{cases}$$

for $k = 0, 1$. We thus select $\lambda$ such that $-d < \lambda \leq d$, and then,

$$p_{\lambda,0} = P[z_d(x) \geq \lambda | H_0] = P[x \geq \lambda | H_0] = 1 - \Phi\left(\frac{\lambda}{\sigma}\right), \quad (19a)$$

$$p_{\lambda,1} = P[z_d(x) \geq \lambda | H_1] = P[x \geq \lambda | H_1] = 1 - \Phi\left(\frac{\lambda - \theta}{\sigma}\right)$$
$$\approx 1 - \Phi\left(\frac{\lambda}{\sigma}\right) + \frac{\theta}{\sigma}\phi\left(\frac{\lambda}{\sigma}\right). \quad (19b)$$

Given $\alpha$ and $\beta$, directly from Proposition 1 in Appendix A (setting $\beta(N) = \beta$ in (A.1)), we now have that the number of sensors needed is as follows:

$$\begin{aligned} M &= M_\lambda(\alpha, \beta) \\ &= \left\{ \frac{[p_{\lambda,1}(1-p_{\lambda,1})]^{1/2}}{p_{\lambda,1}-p_{\lambda,0}} \Phi^{-1}(\beta) - \frac{[p_{\lambda,0}(1-p_{\lambda,0})]^{1/2}}{p_{\lambda,1}-p_{\lambda,0}} \Phi^{-1}(\alpha) \right\}^2. \end{aligned} \tag{20}$$

Substituting the quantities $p_{\lambda,0}$ and $p_{\lambda,1}$ in (19) in expression (20), and using the approximation (for $\theta \ll 1$)

$$\begin{aligned} &\left\{ \Phi\left(\frac{\lambda}{\sigma}\right) \Phi\left(-\frac{\lambda}{\sigma}\right) + \frac{\theta}{\sigma} \phi\left(\frac{\lambda}{\sigma}\right) \left[2\Phi\left(\frac{\lambda}{\sigma}\right) - 1\right] \right\}^{1/2} \\ &\approx \left[\Phi\left(\frac{\lambda}{\sigma}\right) \Phi\left(-\frac{\lambda}{\sigma}\right)\right]^{1/2} + \frac{\theta}{2\sigma} \phi\left(\frac{\lambda}{\sigma}\right) \left[2\Phi\left(\frac{\lambda}{\sigma}\right) - 1\right] \left[\Phi\left(\frac{\lambda}{\sigma}\right) \Phi\left(-\frac{\lambda}{\sigma}\right)\right]^{-1/2}, \end{aligned}$$

we obtain, for $\Gamma(x) = \Phi(x)\Phi(-x)/\phi^2(x)$,

$$\begin{aligned} [M_\lambda(\alpha, \beta)]^{1/2} \approx & [\Phi^{-1}(\beta) - \Phi^{-1}(\alpha)] \left[\Gamma\left(\frac{\lambda}{\sigma}\right)\right]^{1/2} \\ & + \frac{1}{2} \left[2\Phi\left(\frac{\lambda}{\sigma}\right) - 1\right] \left[\Phi\left(\frac{\lambda}{\sigma}\right) \Phi\left(-\frac{\lambda}{\sigma}\right)\right]^{-1/2} \Phi^{-1}(\beta). \end{aligned} \tag{21}$$

Due to Proposition 2 in Appendix A, the first term in (21) is minimized at $\lambda = 0$. It is also trivially concluded that the second term in (21) is minimized at $\lambda = 0$ as well. Thus, the infimum of $[M_\lambda(\alpha, \beta)]^{1/2}$ is attained at $\lambda = 0$, and we easily find that

$$M_0(\alpha, \beta) = \inf_\lambda M_\lambda(\alpha, \beta) = \left[\frac{\sigma}{2\theta\phi(0)}\right]^2 [\Phi^{-1}(\beta) - \Phi^{-1}(\alpha)]^2. \tag{22}$$

(d) Dividing $M_0(\alpha, \beta)$ in (22) by $N_c(\alpha, \beta)$ in (18), we find the result in the Lemma. $\square$

**Remark 1** (i) Comparing the result in (10) with (A.6) of Proposition 3 in Appendix A, we observe that when the two hypotheses are generated by stationary and memoryless Gaussian processes with location parameter shift, the ARE of the decentralized structure behaves as follows: When the structure deploys the robust test function in (9), its ARE is $\eta(d/\sigma)$ times the ARE attained when the structure deploys the optimal-at-Gaussian test function. It is thus interesting to study the behavior of the function $\eta(d/\sigma)$, when the truncation threshold $d$ in (9) varies. (ii) It is not hard to conclude that the function $\eta(x)$ in (11) has the following properties: It is monotonically increasing with $x$, for $x \geq 0$, it is smaller than one for all $x$ in $(0, \infty)$, and $\lim_{x\to\infty} \eta(x) = 1$. We thus conclude that, for the Gaussian hypotheses considered, the decentralized structure has better asymptotic performance when the robust test function in (9) is deployed, as compared to the performance when the optimal-at-Gaussian model test function is adopted. This is so because, when the robust test function is deployed, the centralized structure requires more data to attain given performance, while the decentralized structure remains basically unaffected. This latter characteristic makes the deployment of the robust test function by decentralized structure a necessity. Indeed, the system performance is then not affected, while powerful resistance to highly erroneous data (outliers) is simultaneously accomplished. (iii) When the truncation constant $d$ in (9) approaches asymptotically large values, data truncation is basically eliminated and the parametric and robust test functions become identical to each other. (iv) Let us consider the case where hypothesis $H_0$ is generated by a first order density function $f(x)$, where $f$ is analytic, symmetric around zero, unimodal and $f(0) \geq f(x)$, $\forall x$. Let hypothesis $H_1$ be generated by a shift in location parameter of the density $f(x)$. Consider the decentralized and centralized structures, both deploying the test function $z_d(x)$ in (9). Then, from the steps in the proof of Lemma 1, we can easily conclude that, asymptotically, the number of data needed by the decentralized structure to attain a given $(\alpha, \beta)$ performance is identical to the number needed when the test that is optimal at the density function $f(x)$ is deployed, instead. On the other hand, the number of data needed by the centralized structure to attain the $(\alpha, \beta)$ performance increases, as compared to the number needed when the optimal-at-the density $f(x)$ test function is deployed. Thus, the deployment of the test function $z_d(x)$ in (9) by the decentralized structure presents a powerful advantage for the whole class of analytic, symmetric and unimodal density functions $f(x)$, given the location parameter case.

From the above, the benefits of the test function in (9), as far as decentralized detection with large number of sensors is concerned, are clear. These benefits generalize to a certain class of test functions. This is stated and proven in Theorem 1 below.

**Theorem 1** *Let $C$ denote the class of real, scalar functions such that they are odd and monotonically nondecreasing. Let $F$ be the class of density functions that are symmetric around the zero point, unimodal with maximum at zero and analytic, and converging to zero at infinity. Let hypothesis $H_0$ be generated by a memoryless and stationary process whose first order density, $f$, belongs to $F$. Let hypothesis $H_1$ be generated by a shift in location parameter of hypothesis $H_0$. Consider the decentralized and centralized structures in the presence of the $H_0$ versus $H_1$ hypotheses defined above. Let both structures deploy as test function some function in $C$. Then, given $z \in C$ and $f \in F$, the asymptotic relative efficiency of the decentralized structure as compared to the centralized* *structure at the pair* $(z,\, f)$ *is as follows:*

$$\mathrm{ARE}(\infty,\, z,\, f) = \frac{1}{4f^2(0)} \frac{\left[\int z(x) f'(x) dx\right]^2}{\int z^2(x) f(x) dx}, \tag{23}$$

*where* $f'(x) = \partial f(x)/\partial x$. *The asymptotic relative efficiency in (23) is generally smaller than that attained when the two detection structures deploy the optimal-at-$f$ test function, and it attains its supremum in the latter case. This supremum is given by expression (A.5) in Appendix A.* □

**Proof.** (a) Let $\theta$ denote the location parameter when hypothesis $H_1$ is acting. For $\theta \ll 1$, we have

$$f(x-\theta) \approx f(x) - \theta f'(x). \tag{24}$$

(b) Let us define

$$p_{\lambda,0} \triangleq P[z(x) \geq \lambda | f] = \int_{\{x:\, z(x) \geq \lambda\}} f(x) dx,$$

$$p_{\lambda,1} \triangleq P[z(x) \geq \lambda | f, H_1] = \int_{\{x:\, z(x) \geq \lambda\}} f(x-\theta) dx.$$

Due to the monotonicity of $z(x)$, $z(x) \geq \lambda$ implies $x \geq g(\lambda)$ for some $g(\lambda)$. Due to that, in conjunction with (24), we obtain

$$p_{\lambda,0} = \int_{\{x:\, x \geq g(\lambda)\}} f(x) dx, \tag{25}$$

$$\begin{aligned} p_{\lambda,1} &= \int_{\{x:\, x \geq g(\lambda)\}} f(x-\theta) dx \\ &\approx \int_{\{x:\, x \geq g(\lambda)\}} f(x) dx - \theta \int_{\{x:\, x \geq g(\lambda)\}} f'(x) dx \\ &= p_{\lambda,0} + \theta f(g(\lambda)), \end{aligned}$$

$$p_{\lambda,1} - p_{\lambda,0} \approx \theta f(g(\lambda)),$$

$$p_{\lambda,1}[1-p_{\lambda,1}] \approx p_{\lambda,0}[1-p_{\lambda,0}] + \theta f(g(\lambda))[1-2p_{\lambda,0}],$$

$$\{p_{\lambda,1}[1-p_{\lambda,1}]\}^{1/2} \approx \{p_{\lambda,0}[1-p_{\lambda,0}]\}^{1/2} \left\{1 + \frac{\theta}{2} \frac{f(g(\lambda))[1-2p_{\lambda,0}]}{p_{\lambda,0}(1-p_{\lambda,0})}\right\}. \tag{26}$$

Given the false alarm rate $\alpha$ and the power $\beta$, the number of sensors needed by the decentralized structure to attain the $(\alpha,\, \beta)$ performance is as in (20) in the proof of Lemma 1. Substituting in the latter expression the expressions in (26), we conclude that the number of sensors in the decentralized structure, for given $\lambda$, $\alpha$ and $\beta$, is as follows:

$$M_{\lambda,\alpha,\beta}(z,\, f) \approx \frac{p_{\lambda,0}[1-p_{\lambda,0}]}{\theta^2 f^2(g(\lambda))} \times \left\{\Phi^{-1}(\beta) - \Phi^{-1}(\alpha) + \frac{\theta}{2}\frac{f(g(\lambda))[1-2p_{\lambda,0}]}{[p_{\lambda,0}(1-p_{\lambda,0})]}\Phi^{-1}(\beta)\right\}^2, \quad (27)$$

where $p_{\lambda,0}$ is as in (25). Due to the symmetry and unimodality of $f$, and since $g(0)=0$, we easily conclude that

$$\inf_{\lambda} M_{\lambda,\alpha,\beta}(z,\, f) = M_{0,\alpha,\beta}(z,\, f) \approx \frac{[\Phi^{-1}(\beta)-\Phi^{-1}(\alpha)]^2}{[2\theta f(0)]^2}. \quad (28)$$

It can be easily shown that $M_{0,\alpha,\beta}(z,\, f)$ is identical to the number of sensors needed when the optimal-at-$f$ test function is deployed.
(c) Consider the centralized structure. Given $\alpha$, $\beta$, $z$ and $f$, the number of data needed by this structure is

$$N_{c,\alpha,\beta}(z,\, f) = \frac{[\rho_1\Phi^{-1}(\beta) - \rho_0\Phi^{-1}(\alpha)]^2}{[m_1-m_0]^2}, \quad (29)$$

where due to (24) and the properties of $f$ and $z$, we have

$$m_0 \triangleq \int z(x)f(x)dx = 0, \quad (30a)$$

$$m_1 \triangleq \int z(x)f(x-\theta)dx \approx -\theta\int z(x)f'(x)dx, \quad (30b)$$

$$\rho_0^2 \triangleq \int z^2(x)f(x)dx - m_0^2 = \int z^2(x)f(x)dx, \quad (30c)$$

$$\begin{aligned}\rho_1^2 &\triangleq \int z^2(x)f(x-\theta)dx - m_1^2\\ &\approx \int z^2(x)f(x)dx - \theta\int z^2(x)f'(x)dx - \theta^2\left[\int z(x)f'(x)dx\right]^2\\ &\approx \int z^2(x)f(x)dx - \theta\int z^2(x)f'(x)dx\\ &= \int z^2(x)f(x)dx. \end{aligned} \quad (30d)$$

Substituting expressions (30) in (29), we obtain

$$N_{c,\alpha,\beta}(z, f) \approx \frac{[\Phi^{-1}(\beta)-\Phi^{-1}(\alpha)]^2\int z^2(x)f(x)dx}{[\theta\int z(x)f'(x)dx]^2}. \quad (31)$$

We note that the number in (31) is larger than that obtained when $z$ is the optimal-at-$f$ test function.

(d) Dividing (28) by (31), we obtain (23). Applying Schwartz inequality, we then get

$$\left[\int z(x)f'(x)dx\right]^2 = \left[\int [z(x)\sqrt{f(x)}]\frac{f'(x)}{\sqrt{f(x)}}dx\right]^2$$
$$\leq \left[\int z^2(x)f(x)dx\right]\left[\int \frac{[f'(x)]^2}{f(x)}dx\right], \qquad (32)$$

with equality if and only if $z(x) = z_f(x) \triangleq \pm f'(x)/f(x)$, almost everywhere (a.e.) Applying (32) to (23), we find

$$\mathrm{ARE}(\infty,\, z,\, f) \leq \mathrm{ARE}(\infty,\, z_f,\, f), \qquad (33)$$

with equality if and only if $z(x) = z_f(x)$ a.e., where $z_f(x)$ is the optimal-at-$f$ test function and $\mathrm{ARE}(\infty,\, z_f,\, f)$ is defined in Appendix A. □

We will conclude this section with an intuitively appealing game formalization and its solution. Specifically, let us consider the hypotheses, the detector structures and the $C$ and $F$ classes defined in Theorem 1. Given $z \in C$ and $f \in F$, let us consider a saddle point game with the pay-off function being the asymptotic relative efficiency in (23). Then, we search for a pair $(z^*,\, f^*)$ such that

$$\mathrm{ARE}(\infty,\, z,\, f^*) \leq \mathrm{ARE}(\infty,\, z^*,\, f^*) \leq \mathrm{ARE}(\infty,\, z^*,\, f),$$
$$\forall f \in F,\, \forall z \in C,\, f^* \in F,\, z^* \in C. \qquad (34)$$

In the case where $f(0)$ is fixed, for every $f \in F$, the pair $(z^*,\, f^*)$ is easily found from the results in Theorem 1, and is stated in a corollary.

**Corollary 1** *Let $f(0) = c$, $\forall f \in F$, where c is a constant, and let $I(f)$ denote the Fisher information of the density $f$; that is,*

$$I(f) = \int \frac{[f'(x)]^2}{f(x)}dx.$$

*Then, the pair $(z^*,\, f^*)$ in the game in (34) is as follows:*

$$f^* : \; I(f^*) = \inf_{f \in F} I(f),$$

$$z^*(x) = \pm\frac{f^{*'}(x)}{f^*(x)},$$

*if $\pm f^{*'}(x)/f^*(x)$ belongs to Class $C$.* □

**Proof.** Given $f \in F$, let $z_f$ be such that, $z_f(x) = \pm f'(x)/f(x)$. Directly from Theorem 1, we then have

$$\mathrm{ARE}(\infty,\, z,\, f^*) \leq \mathrm{ARE}(\infty,\, z_{f^*},\, f^*) = \mathrm{ARE}(\infty,\, z^*,\, f^*),\, \forall z \in C,$$

and this proves the left part of the inequality in (34). We now note that $\mathrm{ARE}(\infty, z, f)$ is strictly convex in $f$, and so is $I(f)$. Given $z$, let $f_z$ denote the density that attains the infimum of $\mathrm{ARE}(\infty, z, f)$. Applying calculus of variations, we find that

$$\mathrm{ARE}(\infty,\, z^*,\, f^*) \leq \mathrm{ARE}(\infty,\, z^*,\, f_{z^*}) \leq \mathrm{ARE}(\infty,\, z^*,\, f),\, \forall f \in F.$$

This proves the right part of the inequality in (34), and the proof of the Corollary is now complete. □

**Remark 2** From Theorem 1 and the Corollary we conclude that if a class of densities as in the Theorem is given and they have fixed values at the point zero, and if there exists then some density, $f^*$, in the class that attains infimum and finite Fisher information and is such that $\pm {f^*}'(x)/f^*(x) = z^*(x)$ belongs to Class $C$ in Theorem 1, then $z^*(x)$ is a candidate for a test function in the decentralized structure. If the decentralized structure deploys $z^*(x)$, then its best performance will be induced when the hypotheses are generated by the density $f^*(x)$ and its location parameter shift. When $z^*(x)$ is deployed and the hypotheses are generated by some other density, $f$, in the class, the system performance is better than that when the optimal-at-$f$ test function is deployed. Thus, the $z^*(x)$ choice generally guarantees better than the worst possible performance, for every density in Class $F$.

## 4 Conclusions

In this paper, we studied decentralized detection by a large number of sensors, where each sensor deployed a robust test function. In particular, we considered nonparametrically defined hypotheses, with special emphasis on the outlier nonparametric model. We studied the asymptotic performance of the robust decentralized structures and argued their superiority in terms of asymptotic relative efficiency (ARE) in conjunction with resistance to outliers.

When the acting hypotheses are Gaussian, the decentralized structure has better ARE characteristics if the robust test function is used, instead of the optimal-at-Gaussian test function. Asymptotic performance results, in terms of the ARE, are also obtained for the class of distributions that are symmetric, unimodal, analytic and converging to zero at infinity (Theorem 1).

## 5 Appendix A

In this section, we present some results that are used in Section 3.

**Proposition 1** *Consider a decentralized detection system with asymptotically many sensors, each collecting a single datum. Given the overall system false alarm rate alpha, the power of the system with $M$ sensors is*

$$\beta(M) = \Phi\left\{\frac{p_{\lambda,1} - p_{\lambda,0}}{[p_{\lambda,1}(1-p_{\lambda,1})]^{1/2}}\sqrt{M} - \left(\frac{p_{\lambda,0}(1-p_{\lambda,0})}{p_{\lambda,1}(1-p_{\lambda,1})}\right)^{1/2}\Phi^{-1}(1-\alpha)\right\}, \quad \text{(A.1)}$$

*where $p_{\lambda,0}$ and $p_{\lambda,1}$ are defined in (19).* □

**Proof.** The test function used by the fusion center is $M^{-1}\sum_{j=1}^{M} u_j$, where $u_j$ is the local decision of the $j$th sensor, $j = 1, \ldots, M$. This test function is asymptotically ($M \to \infty$) Gaussian under both hypotheses by the central limit theorem. In particular, we have

$$\frac{1}{M}\sum_{j=1}^{M} u_j \longrightarrow_{M\to\infty} G(m_k,\, \sigma_k/\sqrt{M}),$$

under hypothesis $H_k$, $k = 0,\, 1$, where

$$m_k \triangleq E\{u_j | H_k\} = p_{\lambda,k},$$

$$\sigma_k^2 \triangleq E\{[u_j - m_k]^2 | H_k\} = p_{\lambda,k}(1 - p_{\lambda,k}).$$

The decision $u$ by the fusion center is as follows:

$$u = \begin{cases} 1 \text{ if } \frac{1}{M}\sum_{j=1}^{M} u_j \geq \mu \\ 0 \text{ if } \frac{1}{M}\sum_{j=1}^{M} u_j < \mu, \end{cases}$$

where, for $M \to \infty$ and for a given false alarm rate $\alpha$, we have

$$\alpha = 1 - \Phi\left(\frac{\mu - m_0}{\sigma_0}\sqrt{M}\right) = 1 - \Phi\left(\frac{\mu - p_{\lambda,0}}{[p_{\lambda,0}(1-p_{\lambda,0})]^{1/2}}\sqrt{M}\right). \quad \text{(A.2)}$$

Thus, from (A.2),

$$\mu = p_{\lambda,0} + \left[\frac{p_{\lambda,0}(1-p_{\lambda,0})}{M}\right]^{1/2}\Phi^{-1}(1-\alpha). \quad \text{(A.3)}$$

The power probability $\beta(M)$ attained when the number of sensors is $M$, where $M$ is asymptotically large, is given by the following expression:

$$\beta(M) = 1 - \Phi\left(\frac{\mu - m_1}{\sigma_1}\sqrt{M}\right) = \Phi\left(\frac{p_{\lambda,1} - \mu}{[p_{\lambda,1}(1 - p_{\lambda,1})]^{1/2}}\sqrt{M}\right). \qquad \text{(A.4)}$$

Finally, substituting (A.3) in (A.4) gives the desired result in (A.1). $\square$

**Proposition 2** *Let $f(x)$ be a density function defined on the real line that is also analytic a.e. Let $f'(x) = \partial f/\partial x$, and define*

$$F(x) \triangleq \int_{-\infty}^{x} f(y)dy$$

*and*

$$h(x) \triangleq \frac{F(x)}{f(x)}.$$

*Assume further that $f(x)$ is symmetric around zero and such that (i) $f(\infty) = f(-\infty) = 0$, (ii) $-f'(x)/f(x)$ is monotonically increasing with $x$, and (iii)*

$$-\frac{f'(x)}{f(x)} > \frac{h(x) - h(-x)}{2h(x)h(-x)}, \; \forall x > 0.$$

*Define*

$$\Upsilon(x) = \frac{F(x)F(-x)}{f^2(x)} = \frac{F(x)}{f(x)}\frac{F(-x)}{f(-x)}.$$

*Then, $\Upsilon(x)$ attains its minimum at $x = 0$, where*

$$\Upsilon(0) = \min_x \Upsilon(x) = \frac{1}{4f^2(0)}. \qquad \square$$

**Proof.** Under the conditions stated in the Proposition, $\Upsilon(x) = \Upsilon(-x)$ and $\Upsilon'(x) > 0$, $\forall x > 0$. Due to the symmetry of $\Upsilon(x)$, $\Upsilon'(x) < 0$, $\forall x < 0$. Thus, the function $\Upsilon(x)$ has a minimum attained at $x = 0$. $\square$

**Proposition 3** *Consider the decentralized system in Proposition 1 and the location parameter case. That is, let the hypotheses be $f_0(x) = f(x)$ and $f_1(x) = f(x - \theta)$, where $\theta \ll 1$. Suppose further that $f(x)$ satisfies all the conditions in Proposition 2 and has finite variance. Let $\mathrm{ARE}(\infty, z_f, f)$ denote the asymptotic relative efficiency when the centralized and the decentralized detection systems deploy the optimal-at-$f$ test function $z_f$. Then,*

$$\mathrm{ARE}(\infty, z_f, f) = \frac{1}{4f^2(0)} \int [f'(x)]^2 f(x)dx, \qquad \text{(A.5)}$$

*where the integral in (A.5) is the Fisher information of the density $f$. In particular, if $f$ is Gaussian, then* $\mathrm{ARE}(\infty, z_f, f)$ *becomes*

$$\mathrm{ARE}(\infty, z_G, G) = \frac{\pi}{2}. \qquad \square \tag{A.6}$$

**Proof.** See [1]. □

## References


[1] H. Deliç, *Robustness and Asymptotic Performance Studies in Distributed Detection*, Ph.D. Dissertation, University of Virginia, Charlottesville, Virginia, May 1992.

[2] E. Geraniotis and Y. A. Chau, "Robust data fusion for multisensor detection systems", *IEEE Transactions on Information Theory*, Vol. 36, No. 6, pp. 1265-1279, November 1990.

[3] I. Y. Hoballah and P. K. Varshney, "Neyman-Pearson detection with distributed sensors", *Proceedings of the 25th IEEE Conference on Decision and Control*, Athens, Greece, December 1996, pp. 237-241.

[4] I. Y. Hoballah and P. K. Varshney, "Distributed Bayesian signal detection", *IEEE Transactions on Information Theory*, Vol. 35, No. 5, pp. 995-1000, September 1989.

[5] P. J. Huber, "A robust version of the probability ratio test", *Annals of Mathematical Statistics*, Vol. 36, pp. 1753-1758, December 1965.

[6] P. J. Huber, *Robust Statistics*, Wiley, New York, NY, 1981, Chapter 10, pp. 264-273.

[7] P. J. Huber and V. Strassen, "Minimax tests and the Neyman-Pearson lemma for capacities", *Annals of Statistics*, Vol. 1, No. 2, pp. 251-263, March 1973.

[8] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey", *Proceedings of the IEEE*, Vol. 73, No. 3, pp. 433-481, March 1985.

[9] D. Kazakos and P. Papantoni-Kazakos, *Detection and Estimation*, Computer Science Press, New York, NY, 1990, Chapter 6, pp. 166-195.

[10] H. L. Royden, *Real Analysis*, Macmillan, New York, NY, 1988, 3rd Edition, Chapter 10, pp. 236-238.

[11] S. C. A. Thomopoulos, "Theories in distributed decision fusion", *Proceedings of the IFAC International Symposium on Distributed Intelligence Systems*, Arlington, Virginia, August 1991, pp. 223-228.

[12] J. W. Tukey, "A survey of sampling from contaminated distributions", in: I. Olkin et al., eds., *Contributions to Probability and Statistics, Essays in Honor of Harold Hotelling*, Stanford University Press, Stanford, California, 1960, pp. 448-485.