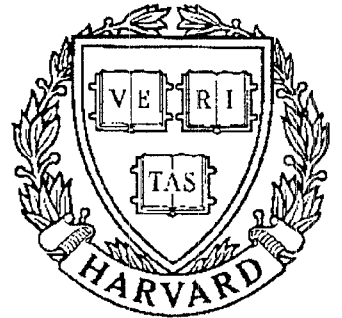


# TECHNICAL RESEARCH REPORT



S Y S T E M S  
R E S E A R C H  
C E N T E R



*Supported by the  
National Science Foundation  
Engineering Research Center  
Program (NSFD CD 8803012),  
the University of Maryland,  
Harvard University,  
and Industry*

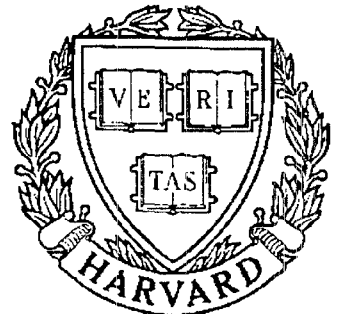
## Multi-Sensor Correlation and Quantization in Distributed Detection Systems

*by Y.A. Chau and E. Geraniotis*

# TECHNICAL RESEARCH REPORT



S Y S T E M S  
R E S E A R C H  
C E N T E R



*Supported by the  
National Science Foundation  
Engineering Research Center  
Program (NSFD CD 8803012),  
the University of Maryland,  
Harvard University,  
and Industry*

Research support for this  
report has been partially  
provided by the Office of  
Naval Research under  
Contract  
N00014-89-J-1375

## Multi-Sensor Correlation and Quantization in Distributed Detection Systems

*by Y.A. Chau and E. Geraniotis*



# MULTI-SENSOR CORRELATION AND QUANTIZATION IN DISTRIBUTED DETECTION SYSTEMS

Yawgeng A. Chau and Evaggelos Geraniotis

Copyright c 1991. Yawgeng A. Chau and Evaggelos Geraniotis. All Rights Reserved.



# MULTI-SENSOR CORRELATION AND QUANTIZATION IN DISTRIBUTED DETECTION SYSTEMS

Yawgeng A. Chau and Evaggelos Geraniotis

Department of Electrical Engineering  
and Systems Research Center  
University of Maryland  
College Park, MD 20742

## ABSTRACT

Quantization and fusion schemes are derived for multi-sensor correlation in distributed  $K$ -sensor systems that are used for the detection of weak signals or general signal discrimination from dependent observations. The dependence in the observations across time and sensors is modeled via stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing processes. The observation sequences of the various sensors have identical individual statistics and identical pairwise statistics (symmetric conditions). Each sensor observation is passed through a memoryless nonlinearity or quantizer (the same one for all sensors) to form the sensor test statistic; the decision statistics of the various sensors are then passed to the fusion center in an unquantized or binary quantized manner to form the final decision statistic of the fusion center. Based on a common large sample size for each sensor that is necessary for achieving high-quality performance, an asymptotic analysis is applied for the error probabilities of the fusion center. This provides design criteria for the optimal memoryless nonlinearity and quantizer. Optimization of these design criteria yields the optimal nonlinearity or quantizer as solutions to linear integral equations involving the first- and second-order pdfs of the sensor observations describing the individual and pairwise dependence. The analytical results obtained are valid for any number of sensors  $K$ . Numerical results based on the simulation of the performance of our schemes with different number of sensors are presented. The performance of the optimal nonlinearities and quantizers is shown to outperform that of nonlinearities or quantizers obtained by ignoring the dependence in sensor observations and to improve as the number of sensors increases.



## I. INTRODUCTION

In most practical multi-sensor detection systems with or without a fusion center, the observations are dependent. For each sensor, the dependence in observation across time is usually due to increased sampling rates or the observed phenomenon itself, whereas the dependence in observations across sensors is due to the geographic proximity of the locations of the sensors. Yet, in the earlier works on distributed detection and data fusion (see [1]-[5]), the sensor observations were assumed to be independent across time and sensors, primarily for the purpose of making the analysis tractable.

In our recent work of [6] and [7], we formulated and solved multi-sensor detection and fusion problems with correlated observations across time and/or sensors of general stationary  $m$ -dependent or mixing type. In particular, in [6] we designed and analyzed quantization and fusion schemes based on memoryless nonlinearities for the detection of weak signals in additive stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing noise; whereas in [7] we introduced and evaluated the performance of similar schemes for general signal discrimination (not of the signal plus noise type) from stationary  $m$ -dependent or mixing observations.

However, although the formulation and techniques of [6] and [7] are quite general and in principle applicable to multi-sensor systems, the specific designs presented and evaluated there were for detection systems consisting of two sensors and a fusion center. In particular, the case of dependence across both time and sensors was analytically solved in [6] and [7] only for two sensors with a fusion center.

In this paper, we extend the results of [6] and [7] to detection systems with  $K$  sensors ( $K \geq 2$ ) and a fusion center. We consider symmetric conditions, under which the sensors have identical individual statistics and identical pairwise joint statistics. Asymptotically optimal



memoryless quantization and fusion schemes are derived for weak-signal detection and general signal discrimination problems with dependence in the observations across-time and sensors. The results obtained are valid for an arbitrary number of sensors and allow us to compare the performance of multi-sensor systems, as the number of sensors increases and the correlation in the sensor observations across time and sensors varies. In deriving these new schemes, this paper extends the useful and popular methodologies of memoryless locally-optimal detection of weak-signals and memoryless nonlocal detection from the single-sensor case to the case of multiple sensors with a fusion center, and also advances earlier work in the area dealing with i.i.d. sensor observations.

The observations sequence of each sensor consists of (i) a weak signal in additive dependent stationary noise or (ii) arbitrary stationary dependent processes, under the two hypotheses. To model the dependence in the observations of each individual sensor and between observations of different sensors we use **stationary mixing processes**, such as stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing processes. The definitions of these mixing processes and the associated central limit theorems are detailed in [8] and [9]. These models of dependence have been successfully used in single-sensor weak-signal detection and discrimination problems (see [10]-[13]).

The sensors collect  $n$  observations each (with  $n$  denoting the common sample size) and form test statistics, which, after (possible) quantization, are transmitted to the fusion center, or directly transmit to the fusion center a quantized version of their observations. At the fusion center, the information sent by the sensors is processed further and a decision is reached as to which of two hypotheses,  $H_0$  or  $H_1$ , is true.

As a rule, the optimal detection schemes under the above models of dependence, for either single-sensor or multi-sensor/fusion-center configurations, involve high-order (larger than

two) probability densities of the observations, which are not only difficult to characterize, but also impose prohibitive requirements on processor memory and complexity for the storage and processing of the dependent data. To avoid these complications in our analysis, design, and implementation, we employ **suboptimal sensor decision statistics and sensor quantization schemes based on memoryless nonlinearities**; the fusion rules employed at the fusion center are basically likelihood ratio tests based on these suboptimal sensor decision statistics or quantization schemes. These are easier to implement and their derivation requires knowledge of only the univariate and bivariate probability densities of the observation sequences. Large sample sizes are necessary for the validity of the analysis and the usefulness of the resulting detectors; but these are well justified, because they are necessary for achieving high-quality performance in hypothesis testing problems involving the detection of a very weak signal or the discrimination between observation sequences with (nearly) identical first- and second-order moments.

We now describe the three schemes (termed Schemes 1 to 3) for quantization and fusion in  $K$ -sensor detection systems with dependent observations, whose analysis, optimal design, and performance evaluation (via simulation) constitute the subject of this paper. The same memoryless nonlinearity  $g(\cdot)$  is employed by all sensors. This is justified by the symmetric conditions that we assumed characterize our problem and the simplicity of the implementation. In general, the argument of  $g(\cdot)$  is a continuous-amplitude (real) variable and  $g$  can take all real values. (In reality,  $g$  is implemented through a discretized form with a large but finite number of amplitude levels) When quantization is employed,  $g(\cdot)$  takes the form of a quantizer and is characterized by a finite number of quantization levels and breakpoints.

For **Scheme 1** (illustrated in Fig. C.1), the  $k$ -th sensor employs the aforementioned  $g$ , forms

the test statistic  $T_{n,k}$  given by

$$T_{n,k} = \frac{1}{n} \sum_{l=1}^n g(X_l^{(k)}); \quad k = 1, 2, \dots, K \quad (1)$$

where  $n$  is the common (large) sample size and  $\{X_l^{(k)}\}_{l=1}^n$  a stationary observation sequence characterized by  $m$ -dependent or mixing type dependence, and transmits it directly to the fusion center, where a likelihood ratio test is performed based on  $T_{n,k}$ , for  $k = 1, 2, \dots, K$ . If one could reliably transmit a real number through a bandlimited channel (as are the channels between the sensors and the fusion center), this fusion scheme, which is optimal within the class of schemes employing memoryless nonlinearities in the sensor test statistics, would also be practical. However, in reality we can only transmit a finite number of bits of information through the aforementioned channels. Therefore, quantization of the test statistics or of the sensor observations themselves is of interest and this is considered in the following two schemes.

In **Scheme 2** (illustrated in Fig. C.2), the observation  $X_l^{(k)}$  of sensor  $k$  at time  $l$  is first quantized by a quantizer  $Q$  (common to all sensors), which is obtained from minimizing an appropriate design criterion, and then transmitted to the fusion center, where the test statistic

$$\bar{T}_{n,k} = \frac{1}{n} \sum_{l=1}^n Q(X_l^{(k)}) \quad (2)$$

is formed. Finally, the fusion center performs a likelihood ratio test based on  $\bar{T}_{n,k}$ , for  $k = 1, 2, \dots, K$ . The optimal quantizer is relatively simple to implement. Even simpler implementations, in which the breakpoints are uniformly spaced over the interval defining the support of the sensor observations are possible at the expense of some degradation in the performance. For a large number of quantization levels these schemes approximate Scheme 1.

In **Scheme 3** (illustrated in Fig. C.3), the test statistic  $T_{n,k}$  for each sensor is first formed as in Scheme 1 by the continuous-amplitude nonlinearity  $g$  and then quantized by a two-bit

quantizer or equivalently compared to a threshold, in order to decide which hypothesis is true. The single bit (0 or 1) of the sensor's decision (also termed a hard decision) is transmitted to the fusion center. The fusion center performs a likelihood ratio test based on the binary decisions of the  $K$  sensors.

As described above, the function of the fusion center is to collect (as in Schemes 1 and 3) or form (as in Schemes 2) the test statistics from the information transmitted by the sensors. In the design of optimal nonlinearities or quantizers for Schemes 1 and 2, we employ the Neyman-Pearson criterion on the error probabilities of the fusion center, whereas for Scheme 3 we employ the asymptotic rates of the error probabilities of the fusion center. It turns out that the final design criterion for each nonlinearity or quantizer has the form of **deflection** for weak-signal detection problems and of the **generalized signal-to-noise ratio** for discrimination problems; these are among the most popular second-order measures of quality for signal classification (see the tutorial in [14]).

The course of obtaining the optimal nonlinearities or quantizers involves solving **linear coupled integral equations**, which depend on the **univariate and bivariate probability density functions** (pdfs) of the sensor noise sequence. The optimal threshold of each sensor  $\eta_k$  is determined by the optimization of the individual design criterion for each case; it depends only on the means and the variances of the test statistics  $T_{n,k}$ , under the two hypotheses, and thus on the aforementioned univariate and bivariate pdfs.

The remainder of this paper is organized as follows. In Sections II, some preliminaries pertaining to the detection of weak signals, to general discrimination, and to all the quantization/fusion schemes considered in this paper are provided. In Section III, the three quantization/fusion schemes are analyzed and optimally designed for the case of weak-signal detec-

ion and dependence across time and sensors under symmetric conditions. The same quantization/fusion schemes are analyzed and optimally designed in Section IV for the case of general signal discrimination and similar models of dependence across time and sensors. In Section V, numerical results based on simulation are presented for the detection of weak signals in Cauchy  $\rho$ -mixing noise. Finally, in Section VI, conclusions are drawn.

## II. PRELIMINARIES

### II.1 Models of Dependent Observations

As mentioned in Section I, the model of dependence is one of stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing sequences. These models are described in detail in the tutorial of [9]. Here we review briefly two of these models of dependence, which are the most pertinent for the purposes of this paper.

The simplest model of dependence assumes that, under both hypotheses, the observations are stationary and **m-dependent**, meaning that the stationary data  $X_k$  and  $X_l$  are dependent with known joint pdf, for  $|k - l| \leq m_i$ , and independent, for  $|k - l| > m_i$ , under hypothesis  $H_i$  (see [8]).

For each positive integer  $n$ , let  $A$  be an event from  $\mathcal{X}_1^k$  (which is the  $\sigma$ -algebra generated by the random variables  $\{X_1, X_2, \dots, X_k\}$ ) and  $B$  an event from  $\mathcal{X}_{k+n}^\infty$  (which is the  $\sigma$ -algebra generated by the random variables  $\{X_{k+n}, X_{k+n+1}, \dots\}$ ). Then, if

$$\text{cov}_i\{X, Y\} \leq \rho_{i,n} \quad (3)$$

for all real random variables which are measurable with respect to  $A$  and  $B$ , respectively, where  $\rho_{i,n}$  is a sequence of real numbers with  $\rho_{i,n} \rightarrow 0$ , as  $n \rightarrow \infty$ , then the sequence of observations

$\{X_1, X_2, \dots\}$  is termed  $\rho$ -**mixing**. Notice that (3) implies the weaker but more intuitive result

$$cov_i\{X_k, X_{k+n}\} \leq \rho_{i,n}. \quad (4)$$

This represents a good model for a time series of data that are asymptotically uncorrelated; since this model requires only asymptotic uncorrelatedness of the data, it is the most general one for our purposes and the numerical results of Section V will be dedicated exclusively to this model. The tutorial of [8] has detailed definitions of the  $\phi$ -**mixing** and  $\alpha$ -**mixing** (also termed strong mixing) models of dependence, which characterize sequences that are respectively asymptotically conditionally independent and asymptotically independent.

## II.2 Asymptotic Distribution of Test Statistics

In general, the true distribution of the sensor decision statistics  $T_{n,k}$  and  $\bar{T}_{n,k}$  ( $k = 1, 2, \dots, K$ ) of (1) and (2) is difficult to obtain. However, for large  $n$ , one can employ the central limit theorem for the aforementioned mixing processes (see [9]). Specifically, assuming that the statistics of the sensor observations are identical and that the pairwise joint statistics of any two sensors are identical (symmetric conditions), we define, under hypothesis  $H_i^{(k)}$  ( $i = 0, 1$ ),

$$\mu_i(g) = E_i[T_{n,k}]; \quad k = 1, 2, \dots, K \quad (5)$$

$$\sigma_i^2(g) = var_i[g(X_1^{(k)})] + 2 \sum_{j=1}^{m_i} cov_i[g(X_1^{(k)}), g(X_{j+1}^{(k)})] \quad (6)$$

where  $var_i[\cdot]$  and  $cov_i[\cdot]$  are the variance and the covariance operators, respectively, and

$$\rho_i(g) = \frac{E_i[(T_{n,k} - \mu_i) \cdot (T_{n,l} - \mu_i)]}{(\frac{1}{n} \sigma_i \sigma_i)} \quad (7)$$

where the numerator of (7) is equal to  $\frac{1}{n}[cov_i\{g(X_1^{(k)}), g(X_1^{(l)})\} + 2 \sum_{j=1}^{\tilde{m}_i} cov_i\{g(X_1^{(k)}), g(X_{j+1}^{(l)})\}]$  ( $k, l = 1, 2, \dots, K$ );  $m_i$  is the range of dependence under  $H_i^{(k)}$ , for the  $m$ -dependent observations of the  $k$ -th sensor; and  $\tilde{m}_i$  the corresponding parameter for the cross-dependence of the observation

sequences between any pair  $(k, l)$  of sensors. For mixing-type observations, we take  $m_i \rightarrow \infty$  and  $\tilde{m}_i \rightarrow \infty$ . Clearly, because of the stationarity, under hypothesis  $H_i^{(k)}$ ,  $\mu_i(g)$  depends on the univariate pdf of the observation sequence  $\{X_l^{(k)}\}_{l=1}^n$  (same for all  $k = 1, 2, \dots, K$ ) denoted by  $f_i(\cdot)$ ;  $\sigma_i^2(g)$  depends on  $f_i$  and the bivariate pdfs of  $(X_1^{(k)}, X_{j+1}^{(k)})$ , for  $j = 1, 2, \dots$ , denoted by  $f_i^{(j)}(\cdot, \cdot)$ ; and  $\rho_i(g)$  depends on  $f_i$  and the bivariate pdfs of  $(X_1^{(k)}, X_{j+1}^{(l)})$ , for  $j = 0, 1, 2, \dots$ , denoted by  $\tilde{f}_i^{(j)}(\cdot, \cdot)$  [same for all pairs  $(k, l)$ ].

According to the central limit theorem for mixing processes, under  $H_i^{(k)}$  ( $i = 0, 1$ ) and suitable conditions (see [8]-[9]), if  $\sigma_i^2$  converges (which is trivially satisfied for  $m$ -dependent processes but needs to be assumed for mixing-type processes) and  $\sigma_i^2 > 0$ ,  $(T_{n,k} - \mu_i)/(\frac{1}{\sqrt{n}}\sigma_i)$  is asymptotically Gaussian distributed with the standard  $\mathcal{N}(0, 1)$  pdf. The conditions for the validity of the central limit theorem for each individual type of mixing processes are described in [9] and are assumed to be satisfied in this paper. For example, for  $\rho$ -mixing processes (for which the numerical results in this paper are generated), besides the convergence of the infinite series above and the strict positivity of the variance, it is also required that the series  $\sum_{n=1}^{\infty} \rho_{i,2^n}$  and  $\sum_{n=1}^{\infty} \tilde{\rho}_{i,2^n}$  converge, where  $\rho_{i,n}$  denotes the sequence of  $\rho$ -mixing parameters [involved in (4)] for the observation sequence of sensor  $k$  (for any  $k = 1, 2, \dots, K$ ), under hypothesis  $H_i^{(k)}$ , and  $\tilde{\rho}_{i,n}$  denotes the  $\rho$ -mixing parameter for the jointly mixing observation sequences  $\{X_j^{(k)}\}_{j=1}^{\infty}$  and  $\{X_{n+j}^{(l)}\}_{j=1}^{\infty}$  of any pair of sensors  $(k, l)$ .

Henceforth we suppose that, under hypothesis  $H_i^{(k)}$  ( $i = 0, 1$ ),  $T_{n,k}$  of sensor  $k$  is asymptotically Gaussian distributed with mean  $\mu_i(g)$  and variance  $\sigma_i^2(g)/n$ ; moreover, that any pair  $(T_{n,k}, T_{n,l})$  of sensor decision statistics is asymptotically jointly Gaussian distributed with the correlation coefficient  $\rho_i(g)$ . In general  $-1 \leq \rho_i(g) \leq 1$ ; however, in Sections III-V below and in the analysis and simulation for the third scheme displayed by Fig. C.3, we assume that

$1 \geq \rho_i \geq 0$  (a model for many practical situations) to derive useful design criteria. We also assume that  $\mu_0 < \mu_1$ . Moreover, for the scheme of Fig. C.3 we assume that the threshold of the sensors  $\eta_s$  satisfies the consistency condition  $\mu_0 \leq \eta_s \leq \mu_1$ , under which every error probability of each sensors approaches to zero, as the sample size increases, according to the mean ergodic theorem (see [12]).

Define the vector of the decision statistics of the  $K$  sensors as

$$\boldsymbol{\xi}_{n,K} = [T_{n,1} \ T_{n,2} \cdots T_{n,K}]^T. \quad (8)$$

Then the mean vector of  $\boldsymbol{\xi}_{n,K}$  under hypothesis  $H_i$  ( $i = 0, 1$ ) is

$$\mathbf{b}_i = E_i[\boldsymbol{\xi}_{n,K}] = [\mu_i \ \mu_i \cdots \mu_i]^T \quad (9)$$

and the corresponding ( $K$  by  $K$ ) covariance matrix

$$E_i[(\boldsymbol{\xi}_{n,K} - \mathbf{b}_i)^T(\boldsymbol{\xi}_{n,K} - \mathbf{b}_i)] = \frac{1}{n} \boldsymbol{\Sigma}_i \quad (10)$$

where

$$\boldsymbol{\Sigma}_i = \sigma_i^2 \begin{bmatrix} 1 & \rho_i & \cdots & \rho_i & \rho_i \\ \rho_i & 1 & \cdots & \rho_i & \rho_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_i & \rho_i & \cdots & 1 & \rho_i \\ \rho_i & \rho_i & \cdots & \rho_i & 1 \end{bmatrix}, \quad (11)$$

In (11) we have suppressed the dependence of  $\mu_i$ ,  $\sigma_i^2$ , and  $\rho_i$  on the common (for all sensors) nonlinearity  $g$ . In the sequel, since we are interested in a large common sample size  $n$ ,  $[T_{n,1} T_{n,2} \cdots T_{n,K}]$  is assumed to be a Gaussian random vector with mean vector  $\mathbf{b}_i$  and covariance matrix  $\frac{1}{n} \boldsymbol{\Sigma}_i$ , under hypothesis  $H_i$  ( $i = 0, 1$ ).



### II.3 Asymptotic Rates for Gaussian Statistics

To characterize the asymptotic rates of decrease of the error probabilities for Gaussian statistics as the sample size  $n$  increases, we apply the large deviations principle given in [15]. We consider the  $d$ -dimensional Gaussian random vector  $\mathbf{Y} = [Y_1 \ Y_2 \ \cdots \ Y_d]^T$  where each of the  $Y_1, Y_2, \dots$  is generated as in (1) for large  $n$  via a Central Limit Theorem, and define the vectors  $\mathbf{s}^T = [s_1 \ s_2 \ \cdots \ s_d]$  and  $\mathbf{z} = [z_1 \ z_2 \ \cdots \ z_d]^T$ . Under  $H_i$  ( $i = 0, 1$ ), the scalar functions relevant to the application of the large deviations principle to this case are (see [15])

$$c_i(\mathbf{s}) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln E_i\{\exp(n\mathbf{s}^T \mathbf{Y})\} = \mathbf{s}^T E_i[\mathbf{Y}] + \frac{1}{2} \mathbf{s}^T \overline{\text{cov}}_i[\mathbf{Y}] \mathbf{s} \quad (12)$$

and

$$I_i(\mathbf{z}) = \sup_{\mathbf{s} \in \mathbb{R}^d} \{\mathbf{s}^T \mathbf{z} - c_i(\mathbf{s})\} = \frac{1}{2} (\mathbf{z} - E_i[\mathbf{Y}])^T \{\overline{\text{cov}}_i[\mathbf{Y}]\}^{-1} (\mathbf{z} - E_i[\mathbf{Y}]), \quad i = 0, 1 \quad (13)$$

where  $\frac{1}{n} \overline{\text{cov}}_i[\mathbf{Y}]$  is the covariance matrix of  $\mathbf{Y}$ ,  $c_i(\mathbf{s})$  are closed convex functions, and  $I_i$  are termed the Legendre-Fenchel transforms of  $c_i(\mathbf{s})$  (see [15]). Next, we define the entropy functions on any set  $G \in \mathbb{R}^d$  as

$$\bar{I}_i(G) = \inf\{I_i(\mathbf{z}) : \mathbf{z} \in G\} \quad (14)$$

and the induced probabilities

$$P_i(G) = \text{Pr}_i\{\mathbf{Y} : \mathbf{Y} \in G\} \quad (15)$$

under  $H_i$ , for  $i = 0, 1$ . Then, according to Hypothesis II.1 and Theorem II.2 of [15] and for any set  $G \in \mathbb{R}^d$ , the exponential rates of the probabilities  $P_i(G)$  are characterized by the formula

$$\frac{1}{n} \ln P_i(G) \rightarrow \bar{I}_i(G), \quad i = 0, 1. \quad (16)$$

as  $n \rightarrow \infty$ . In Sections III and IV, (16) is applied to derive closed-forms of the exponential rates of the error probabilities of the fusion center.

### III. DETECTION OF WEAK SIGNALS

We consider the following binary hypotheses testing model for the detection of a weak signal in additive stationary dependent noise

$$\begin{aligned} H_0^{(k)} &: X_i^{(k)} = N_i^{(k)} \\ H_1^{(k)} &: X_i^{(k)} = \theta + N_i^{(k)}, \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, K \end{aligned} \quad (17)$$

where  $\theta \rightarrow 0$  with  $\theta > 0$  is the nonrandom weak signal and  $\{N_i^{(k)}\}_{i=1}^n$  are stationary noise sequences of the  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing type. Under the symmetric conditions discussed in Section I, the univariate and bivariate densities of  $N_i^{(k)}$  are identical for different sensors and are denoted by  $f(\cdot)$  and  $f^{(j)}(\cdot, \cdot)$  ( $j = 1, 2, \dots, m$ ) for the observations  $N_i^{(k)}$  and  $(N_i^{(k)}, N_{i+j}^{(k)})$ , respectively; the cross-bivariate densities are also identical for any pairs of sensors  $(k, l)$  and are denoted by  $\tilde{f}^{(j)}(\cdot, \cdot)$  ( $j = 0, 1, \dots, \tilde{m}$ ) for the observations  $(N_i^{(k)}, N_{i+j}^{(l)})$ .

In this weak signal detection problem, we adopt the notation  $\mu_\theta = \mu_1$ ,  $\sigma_\theta^2 = \sigma_1^2$ , and  $\rho_\theta = \rho_1$  to express the functional dependence on  $\theta$  and assume the following regularity conditions:

- (a)  $\frac{\partial}{\partial \theta} [\int g(x) f(x - \theta) dx] = \int \frac{\partial}{\partial \theta} [g(x) f(x - \theta)] dx$
- (b)  $\frac{\mu_\theta(g) - \mu_0(g)}{\theta} \rightarrow \frac{\partial \mu_\theta(g)}{\partial \theta} |_{\theta=0} > 0$
- (c)  $\sigma_\theta^2(g) \rightarrow \sigma_0^2(g)$
- (d)  $\rho_\theta(g) \rightarrow \rho_0(g)$

as  $\theta \rightarrow 0$ . We use  $\sigma_0(g)$  and  $\rho_0(g)$  to represent the (normalized) variance and correlation coefficient of the sensor decision statistics under either hypothesis in this section.

#### III.1 Fusion of Unquantized Test Statistics

We here consider the scheme described by Fig. C.1. For jointly Gaussian distributed  $T_{n,k}$  ( $k = 1, 2, \dots, K$ ), the (normalized) log-likelihood ratio function for the above weak signal detec-

tion problem becomes

$$\begin{aligned}\frac{1}{n} \ln L_n &= \frac{1}{n} \ln \frac{f_\theta(T_{n,1}, T_{n,2}, \dots, T_{n,K})}{f_0(T_{n,1}, T_{n,2}, \dots, T_{n,K})} \\ &= \frac{1}{2}(\boldsymbol{\xi}_{n,K} - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\xi}_{n,K} - \mathbf{b}_0) - \frac{1}{2}(\boldsymbol{\xi}_{n,K} - \mathbf{b}_\theta)^T \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\xi}_{n,K} - \mathbf{b}_\theta)\end{aligned}\quad (18)$$

$$= (\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\xi}_{n,K} + \frac{1}{2} \mathbf{b}_0^T \boldsymbol{\Sigma}_0^{-1} \mathbf{b}_0 - \frac{1}{2} \mathbf{b}_\theta^T \boldsymbol{\Sigma}_0^{-1} \mathbf{b}_\theta \quad (19)$$

where the inverse of the (normalized) covariance matrix has the form

$$\begin{aligned}\boldsymbol{\Sigma}_0^{-1} &= \frac{1}{\sigma_0^2(1-\rho_0)} \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & & 1 \end{bmatrix} - \frac{\rho_0}{\sigma_0^2(1-\rho_0)(1+(K-1)\rho_0)} \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & & 1 \end{bmatrix} \\ &= \frac{1}{\sigma_0^2(1-\rho_0)} \mathbf{I} - \frac{\rho_0}{\sigma_0^2(1-\rho_0)(1+(K-1)\rho_0)} \mathbf{1}\end{aligned}\quad (20)$$

in which  $\mathbf{I}$  is the  $K$  by  $K$  identity matrix and  $\mathbf{1}$  denotes the  $K$  by  $K$  matrix having all its elements equal to 1. We notice that the above  $\frac{1}{n} \ln L_n$  is Gaussian distributed, since it is a linear function of the Gaussian statistics  $\boldsymbol{\xi}_{n,K}$ .

The asymptotic means of  $\ln L_n$  under the two hypotheses take the forms

$$\varphi_0 = E_0 \left[ \frac{1}{n} \ln L_n \right] = -\frac{1}{2}(\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{b}_\theta - \mathbf{b}_0) \quad (21)$$

and

$$\varphi_\theta = E_\theta \left[ \frac{1}{n} \ln L_n \right] = \frac{1}{2}(\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{b}_\theta - \mathbf{b}_0) \quad (22)$$

and its asymptotic variance under either hypothesis is

$$\text{var}_0 \left[ \frac{1}{n} \ln L_n \right] = \frac{1}{n} \overline{\text{var}}_0 \left[ \frac{1}{n} \ln L_n \right] = \frac{1}{n}(\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{b}_\theta - \mathbf{b}_0). \quad (23)$$

For the threshold of the fusion center  $\eta$ , we impose the consistency condition  $\varphi_0 \leq \eta \leq \varphi_1$  so that the error probability of the fusion center  $P_{K,\eta}$  converges to zero, as the sensor sample size increases to infinity, which results from an application of the ergodic mean theorem. The error probability of the fusion center is

$$P_{K,\eta} = pP_0(\ln L_n > n\eta) + (1-p)P_\theta(\ln L_n \leq n\eta) = pP_0(1) + (1-p)P_\theta(0) \quad (24)$$

where we used the notation  $P_0(1) = P_0(\ln L_n > n\eta)$  and  $P_\theta(0) = P_\theta(\ln L_n \leq n\eta)$ . Let  $P_{max} = \max\{P_0(\ln L_n > n\eta), P_\theta(\ln L_n \leq n\eta)\}$ . Applying the result of large deviations principle given by (16) to the one-dimensional case ( $d = 1$ ) we obtain

$$\begin{aligned} -\frac{1}{n} \ln P_{K,\eta} &= -\frac{1}{n} \ln P_{max} - \frac{1}{n} \left[ \ln \left( \frac{pP_0(1)}{P_{max}} + \frac{(1-p)P_\theta(0)}{P_{max}} \right) \right] \\ &\rightarrow -\frac{1}{n} \ln P_{max} = \min \left\{ -\frac{1}{n} \ln P_0 \left( \frac{1}{n} \ln L_n > \eta \right), -\frac{1}{n} \ln P_\theta \left( \frac{1}{n} \ln L_n \leq \eta \right) \right\} \\ &= \min \{ \bar{I}_0(\{z > \eta\}), \bar{I}_\theta(\{z \leq \eta\}) \} \end{aligned} \quad (25)$$

where, from (13), (14) and (21)-(23) and under the consistency condition  $\varphi_0 \leq \eta \leq \varphi_1$ , we obtain

$$\begin{aligned} \bar{I}_0(\{z > \eta\}) &= \inf_{z > \eta} \frac{1}{2} \left( z - E_0 \left[ \frac{1}{n} \ln L_n \right] \right)^2 \left\{ \overline{var}_0 \left[ \frac{1}{n} \ln L_n \right] \right\}^{-1} \\ &= \frac{(\eta - \varphi_0)^2}{2(\mathbf{b}_\theta - \mathbf{b}_0)^T \Sigma_0^{-1} (\mathbf{b}_\theta - \mathbf{b}_0)} \end{aligned} \quad (26)$$

and

$$\begin{aligned} \bar{I}_\theta(\{z \leq \eta\}) &= \inf_{z \leq \eta} \frac{1}{2} \left( z - E_\theta \left[ \frac{1}{n} \ln L_n \right] \right)^2 \left\{ \overline{var}_\theta \left[ \frac{1}{n} \ln L_n \right] \right\}^{-1} \\ &= \frac{(\varphi_\theta - \eta)^2}{2(\mathbf{b}_\theta - \mathbf{b}_0)^T \Sigma_0^{-1} (\mathbf{b}_\theta - \mathbf{b}_0)}. \end{aligned} \quad (27)$$

Then the optimal threshold is determined as

$$\eta_{opt} = \arg \max_{\varphi_0 \leq \eta \leq \varphi_1} \left\{ -\frac{1}{n} \ln P_{K,\eta} \right\}$$

$$\begin{aligned}
&= \arg \max_{\varphi_0 \leq \eta \leq \varphi_1} \min\{\bar{I}_0(\{z > \eta\}), \bar{I}_\theta(\{z \leq \eta\})\} \\
&= \frac{\varphi_0 + \varphi_1}{2} = 0.
\end{aligned} \tag{28}$$

Therefore

$$\begin{aligned}
-\frac{1}{n} \ln P_{K, \eta_{opt}} &\rightarrow \min\{\bar{I}_0(\{z > 0\}), \bar{I}_\theta(\{z \leq 0\})\} \\
&= \frac{\varphi_0^2}{2(\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{b}_\theta - \mathbf{b}_0)} \\
&= \frac{1}{8} \cdot (\mathbf{b}_\theta - \mathbf{b}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{b}_\theta - \mathbf{b}_0) \\
&= \frac{1}{8} \cdot \frac{K[\mu_\theta(g) - \mu_0(g)]^2}{[1 + (K-1)\rho_0(g)]\sigma_0^2(g)} \\
&\rightarrow \frac{\theta^2}{8} \cdot \frac{K[\mu'_0(g)]^2}{[1 + (K-1)\rho_0(g)]\sigma_0^2(g)}
\end{aligned} \tag{29}$$

where we used (20) for  $\boldsymbol{\Sigma}_0^{-1}$  and the regularity condition (b) to obtain that  $\mu_\theta(g) - \mu_0(g) = \mu'_0(g)\theta$  for the derivative  $\mu'_0(g)$  at  $\theta = 0$  and small  $\theta$  ( $\theta \rightarrow 0$ ).

The optimal nonlinearity  $g_K^*$  of a  $K$ -sensor system is obtained as

$$g_K^* = \arg \max_g \frac{[\mu'_0(g)]^2}{[1 + (K-1)\rho_0(g)]\sigma_0^2(g)}. \tag{30}$$

The maximization in (30) is similar to those performed in [6] and the optimal nonlinearity satisfies the linear integral equation

$$-f'(x) - \int K_K(x, y) g_K^*(y) dy = g_K^*(x) f(x) \tag{31}$$

where the integration kernel is

$$\begin{aligned}
K_K(x, y) &= 2 \sum_{j=1}^m [f^{(j)}(x, y) - f(x)f(y)] - f(x)f(y) \\
&\quad + (K-1) \left\{ [\tilde{f}^{(0)}(x, y) - f(x)f(y)] + 2 \sum_{j=1}^{\tilde{m}} [\tilde{f}^{(j)}(x, y) - f(x)f(y)] \right\}
\end{aligned} \tag{32}$$

The above linear integral equation is a necessary condition for the optimality of  $g_K^*$ , to this we should add the sufficient condition

$$\rho_0(g) > -\frac{1}{K-1} \quad (33)$$

for all functions  $g$  satisfying conditions (a)-(d) for which  $\sigma_0^2(g)$  and  $\rho_0(g)$  converge and  $\sigma_0^2(g) > 0$ .

We can actually show that if  $g_K^*$  and  $g_{K-1}^*$  are the optimal nonlinearities for  $K$ -sensor and  $(K-1)$ -sensor systems, respectively, then

$$\begin{aligned} \frac{K[\mu'_0(g_K^*)]^2}{[1 + (K-1)\rho_0(g_K^*)]\sigma_0^2(g_K^*)} &\geq \frac{K[\mu'_0(g_{K-1}^*)]^2}{[1 + (K-1)\rho_0(g_{K-1}^*)]\sigma_0^2(g_{K-1}^*)} \\ &\geq \frac{(K-1)[\mu'_0(g_{K-1}^*)]^2}{[1 + (K-2)\rho_0(g_{K-1}^*)]\sigma_0^2(g_{K-1}^*)} \end{aligned} \quad (34)$$

where the first inequality results from the optimality of  $g_K^*$  and the second inequality is due to the fact that  $K/[1 + (K-1)\rho] \geq (K-1)/[1 + (K-2)\rho]$  for  $1 \geq \rho$ . The inequality in (34) implies that increasing the number of sensors increases the asymptotic rate of the error probabilities of the fusion center, i.e.,  $\lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K, \eta_{opt}} \right] \geq \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K-1, \eta_{opt}} \right]$ , and thus, since  $P_{K, \eta_{opt}} \leq P_{K-1, \eta_{opt}}$  for large  $n$ , it improves the performance of the multi-sensor system.

### III.2 Fusion of Quantized Observations

Considering quantizers as special (discontinuous) forms of nonlinearities we may modify the formulation of the previous section to derive the design criterion for the optimal quantizer used in the scheme given by Fig. C.2.

Let  $Q(X_i^{(k)})$  ( $i = 1, 2, \dots, n$ ) be the output of the quantizer with  $M$  quantization levels common for all sensors. Denote by  $\mathbf{t} = [t_0, t_1, \dots, t_M]$  its breakpoints and by  $\mathbf{u} = [u_1, u_2, \dots, u_M]$  its quantization levels. Then, from (6) and (7),

$$\bar{\sigma}_0^2(Q) = \sum_{l=1}^M (u_l)^2 \int_{t_{l-1}}^{t_l} f(x) dx + 2 \sum_{j=1}^m \sum_{r=1}^M \sum_{l=1}^M u_r u_l \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} f^{(j)}(x, y) dx dy$$

$$-(2m+1) \left[ \sum_{l=1}^M u_l \int_{t_{l-1}}^{t_l} f(x) dx \right]^2. \quad (35)$$

and

$$\begin{aligned} \tilde{\rho}_0(Q) \tilde{\sigma}_0^2(Q) &= E_0[Q(X_1^{(1)})Q(X_1^{(2)})] + 2 \sum_{j=1}^{\tilde{m}} E_0[Q(X_1^{(1)})Q(X_{j+1}^{(2)})] - (2\tilde{m}+1)(E_0[Q(X_1^{(1)})])^2 \\ &= \sum_{r=1}^M \sum_{l=1}^M u_r u_l \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}^{(0)}(x, y) dx dy + 2 \sum_{j=1}^{\tilde{m}} \sum_{r=1}^M \sum_{l=1}^M u_r u_l \cdot \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}^{(j)}(x, y) dx dy \\ &\quad - (2\tilde{m}+1) \left[ \sum_{l=1}^M u_l \int_{t_{l-1}}^{t_l} f(x) dx \right]^2. \end{aligned} \quad (36)$$

Moreover, the partial derivative of the mean takes the form

$$\tilde{\mu}'_0(Q) = \frac{\partial \tilde{\mu}_\theta(Q)}{\partial \theta} \Big|_{\theta=0} = - \int Q(x) f'(x) dx = \sum_{l=1}^M u_l [f(t_{l-1}) - f(t_l)]. \quad (37)$$

Define the vector

$$\Delta \mathbf{f} = [-\Delta f_1, -\Delta f_2, \dots, -\Delta f_M] \quad (38)$$

where  $\Delta f_l = f(t_l) - f(t_{l-1})$  and  $\mathbf{F} = \text{diag}\{\int_{t_0}^{t_1} f(x) dx, \dots, \int_{t_{M-1}}^{t_M} f(x) dx\}$ ,  $\mathbf{P} = [P_{r,l}]$ ,  $\tilde{\mathbf{P}} = [\tilde{P}_{r,l}]$ ,  $\mathbf{R} = [R_{r,l}]$ , and  $\tilde{\mathbf{R}} = [\tilde{R}_{r,l}]$ , where

$$P_{r,l} = 2 \sum_{j=1}^m \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} f^{(j)}(x, y) dx dy \quad (39)$$

$$\tilde{P}_{r,l} = \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}^{(0)}(x, y) dx dy + 2 \sum_{j=1}^{\tilde{m}} \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}^{(j)}(x, y) dx dy \quad (40)$$

$$R_{r,l} = (2m+1) \cdot \int_{t_{r-1}}^{t_r} f(x) dx \cdot \int_{t_{l-1}}^{t_l} f(x) dx \quad (41)$$

and

$$\tilde{R}_{r,l} = (2\tilde{m}+1) \cdot \int_{t_{r-1}}^{t_r} f(x) dx \cdot \int_{t_{l-1}}^{t_l} f(x) dx \quad (42)$$

for  $r, l = 1, 2, \dots, M$ . Then from (30), the design criterion for the optimal quantizer (levels and breakpoints) is

$$\frac{[\tilde{\mu}'_0(Q)]^2}{[1 + (K-1)\tilde{\rho}_0(Q)]\tilde{\sigma}_0^2(Q)} = \frac{(\mathbf{u} \Delta \mathbf{f}^T)^2}{\mathbf{u}(\mathbf{F} + \mathbf{P} + \tilde{\mathbf{P}} - \mathbf{R} - \tilde{\mathbf{R}})\mathbf{u}^T}. \quad (43)$$

If we assume that  $t_l$ , for  $l = 0, 1, \dots, M$ , are chosen such that the matrix  $F + P + \tilde{P} - R - \tilde{R}$  is positive definite, then for fixed breakpoints we can obtain the optimal quantization levels as

$$\mathbf{u}^T = (F + P + \tilde{P} - R - \tilde{R})^{-1} \Delta \mathbf{f}^T \quad (44)$$

Substituting for these optimal quantization levels we obtain the design criterion as a function of the breakpoints  $\mathbf{t}$  in the form

$$\Delta \mathbf{f} (F + P + \tilde{P} - R - \tilde{R})^{-1} \Delta \mathbf{f}^T \quad (45)$$

This expression can be optimized with respect to  $\mathbf{t}$  to obtain the optimal breakpoints by means of numerical optimization techniques such as the gradient method.

Similar arguments to those used in deriving (34) establish that the performance of this scheme can be improved by increasing the number of sensors.

### III.3 Fusion of Binary Sensor Decisions

Consider the scheme given by Fig. C.3, in which the  $k$  sensor ( $k = 1, 2, \dots, K$ ) performs a threshold test based on the test statistic  $\xi_{n,k}$  with the same nonlinearity  $g$  and threshold  $\eta_s$  and transmits its binary decision (0 or 1) to the fusion center, where a likelihood ratio test is carried out on the basis of the sensor decisions. Let  $d_{n,k}$  be the decision of sensor  $k$  based on its  $n$  observations. Then the log-likelihood ratio function of the fusion center is  $\ln L_n = \ln P_1(d_{n,1}, d_{n,2}, \dots, d_{n,K}) / P_0(d_{n,1}, d_{n,2}, \dots, d_{n,K})$  and the asymptotic log-likelihood ratio function under a large sample size is characterized by

$$\frac{1}{n} \ln L_n = -\frac{1}{n} \ln P_0(d_{n,1}, d_{n,2}, \dots, d_{n,K}) - \left[ -\frac{1}{n} \ln P_1(d_{n,1}, d_{n,2}, \dots, d_{n,K}) \right] \quad (46)$$

The fusion center executes the test  $\ln L_n \stackrel{H_1}{\underset{H_0}{>}} n\eta$ . In this scheme, the threshold of the fusion center  $\eta$  is selected to be zero since the case of interest involves equiprobable a priori probabilities for the two hypotheses.



In Appendix A, we derive by means of the large deviations principle the asymptotic rate of the joint error probability of the  $K$  sensors, when there are  $i$  0s transmitted from them. To compute the likelihood ratio test of the fusion center, we give here a proposition, which is proved in Appendix B.

**Proposition 1:** Under the above framework, and for the problem of weak signal detection, the likelihood ratio test of the fusion center is equivalent to a majority fusion rule if and only if the threshold  $\eta_s^*$  used by all sensors is

$$\eta_s^* = \frac{\mu_\theta + \mu_0}{2} \quad (47)$$

where  $\mu_\theta$  and  $\mu_0$  are functionals of the nonlinearity  $g_K$ . In this case, the asymptotic rate of the error probability of the fusion center becomes

$$-\frac{1}{n} \ln P_{K, \eta_s^*} \rightarrow \frac{[K/2][\mu_\theta(g) - \mu_0(g)]^2}{8[1 + ([K/2] - 1)\rho_0(g)]\sigma_0^2(g)} = \frac{\theta^2}{8} \cdot \frac{[K/2][\mu'_0(g)]^2}{[1 + ([K/2] - 1)\rho_0(g)]\sigma_0^2(g)}. \quad (48)$$

where  $[K/2]$  is equal to  $(K + 1)/2$  if  $K$  is odd and to  $K/2$  if  $K$  is even.

The above asymptotic rate is used as the design criterion for deriving the optimal nonlinearity  $g_K$  for this scheme. In particular the optimal nonlinearity is obtained as

$$g_K^* = \begin{cases} \max_g \frac{(K+1)[\mu'_0(g)]^2}{[2+(K-1)\rho_0(g)]\sigma_0^2(g)}, & \text{if } K \text{ is odd} \\ \max_g \frac{K[\mu'_0(g)]^2}{[2+(K-2)\rho_0(g)]\sigma_0^2(g)}, & \text{if } K \text{ is even} \end{cases} \quad (49)$$

By means of the calculus of variations method described in [10] (see also [6]-[7]), we obtain the following integral equations as the necessary condition for the optimization problem characterized by (49)

$$-f'(x) - \int H_K(x, y)g_K^*(y)dy = g_K^*(x)f(x) \quad (50)$$

with the integration kernel

$$H_K(x, y) = 2 \sum_{j=1}^m \left[ f^{(j)}(x, y) - f(x)f(y) \right] - f(x)f(y)$$

$$+\frac{K-1}{2} \left\{ \left[ \tilde{f}^{(0)}(x, y) - f(x)f(y) \right] + 2 \sum_{j=1}^{\tilde{m}} \left[ \tilde{f}^{(j)}(x, y) - f(x)f(y) \right] \right\} \quad (51)$$

if  $K$  is odd, or

$$\begin{aligned} H_K(x, y) &= 2 \sum_{j=1}^m \left[ f^{(j)}(x, y) - f(x)f(y) \right] - f(x)f(y) \\ &+ \frac{K-2}{2} \left\{ \left[ \tilde{f}^{(0)}(x, y) - f(x)f(y) \right] + 2 \sum_{j=1}^{\tilde{m}} \left[ \tilde{f}^{(j)}(x, y) - f(x)f(y) \right] \right\} \end{aligned} \quad (52)$$

if  $K$  is even. The sufficient condition for achieving the maximum in (49) is

$$\rho_0(g) > -\frac{1}{\lceil K/2 \rceil - 1} \quad (53)$$

for all functions  $g$  satisfying conditions (a)-(d) for which  $\sigma_0^2(g)$  and  $\rho_0(g)$  converge and  $\sigma_0^2(g) > 0$ .

Suppose that  $g_{K-1}^*$  and  $g_K^*$  denote the optimal nonlinearities of  $(K-1)$ -sensor and  $K$ -sensor systems employing the scheme of Fig. C.3, respectively. Using arguments similar to those used in the proof of (34) we can show that for odd  $K$  (and thus even  $K-1$ )

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K, \eta_s^*} \right] &= \frac{(K+1)[\mu'_0(g_K^*)]^2}{[2 + (K-1)\rho_0(g_K^*)]\sigma_0^2(g_K^*)} \\ &\geq \frac{(K+1)[\mu'_0(g_{K-1}^*)]^2}{[2 + (K-1)\rho_0(g_{K-1}^*)]\sigma_0^2(g_{K-1}^*)} \\ &\geq \frac{(K-1)[\mu'_0(g_{K-1}^*)]^2}{[2 + (K-3)\rho_0(g_{K-1}^*)]\sigma_0^2(g_{K-1}^*)} = \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K-1, \eta_s^*} \right] \end{aligned} \quad (54)$$

and for even  $K$  (and thus odd  $K-1$ )

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K, \eta_s^*} \right] &= \frac{K[\mu'_0(g_K^*)]^2}{(2 + (K-2)\rho_0(g_K^*))\sigma_0^2(g_K^*)} \\ &\geq \frac{K[\mu'_0(g_{K-1}^*)]^2}{(2 + (K-2)\rho_0(g_{K-1}^*))\sigma_0^2(g_{K-1}^*)} = \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K-1, \eta_s^*} \right]. \end{aligned} \quad (55)$$

Thus, under a large sample size and for any number of sensors  $K$  (odd or even), the error probability of the fusion center decreases as the number of sensors increases.

#### IV. DISCRIMINATION

We consider the problem of distributed discrimination from general (not necessarily signal in additive noise) sensor observations characterized by stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing processes. This discrimination problem can be modeled by the following binary hypotheses testing problem

$$\begin{aligned} H_0^{(k)} &: X_l^{(k)} \sim [f_0(\cdot), f_0^{(j)}(\cdot, \cdot)] \\ H_1^{(k)} &: X_l^{(k)} \sim [f_1(\cdot), f_1^{(j)}(\cdot, \cdot)], \quad l = 1, \dots, n; \quad k = 1, 2, \dots, K \end{aligned} \quad (56)$$

where  $f_i$  ( $i = 0, 1$ ) are the univariate pdfs of the sensor observation processes under  $H_i$  (identical for all sensors),  $f_i^{(j)}$  are the bivariate pdfs of  $(X_l^{(k)}, X_{l+j}^{(k)})$  (also identical for all  $k$ ), and  $\tilde{f}_i^{(j)}$  ( $i = 0, 1$ ) are the joint pdfs of  $(X_l^{(r)}, X_{l+j}^{(k)})$  (also identical for all pairs  $(r, k)$  of sensors with  $r \neq k$ ). Clearly, the above multi-sensor testing hypothesis problem describes several situations that are not characterized by signals in additive noise and thus is a suitable model for target discrimination problems.

For a large common sample size  $n$ , the statistics  $(T_{n,1}, T_{n,2}, \dots, T_{n,K})$  obtained via (1) and (2) are jointly asymptotic Gaussian distributed when conditioned on  $H_i$  ( $i = 0, 1$ ), as in the previous section.

##### IV.1 Fusion of Unquantized Test Statistics

For the scheme depicted in Fig. C.1, and for jointly Gaussian distributed  $T_{n,k}$  ( $k = 1, 2, \dots, K$ ), the log-likelihood ratio function of the fusion center takes the form

$$\begin{aligned} \frac{1}{n} \ln L_n &= \frac{1}{n} \ln \frac{f_1(T_{n,1}, T_{n,2}, \dots, T_{n,K})}{f_0(T_{n,1}, T_{n,2}, \dots, T_{n,K})} \\ &= \frac{1}{n} \ln \frac{|\Sigma_0|}{|\Sigma_1|} + \frac{1}{2} (\xi_{n,K} - b_0)^T \Sigma_0^{-1} (\xi_{n,K} - b_0) - \frac{1}{2} (\xi_{n,K} - b_1)^T \Sigma_1^{-1} (\xi_{n,K} - b_1) \end{aligned} \quad (57)$$

where  $\frac{1}{n}\Sigma_i$  denotes the covariance matrix of  $\xi_{n,K}$  and was defined in (10)-(11) for  $i = 0, 1$ . Its inverse is given by

$$\Sigma_i^{-1} = \frac{1}{\sigma_i^2(1 - \rho_i)} \mathbf{I} - \frac{\rho_i}{\sigma_i^2(1 - \rho_i)(1 + (K - 1)\rho_i)} \mathbf{1} \quad (58)$$

where  $\mathbf{I}$  and  $\mathbf{1}$  were defined in (20).

The mean of  $\frac{1}{n} \ln L_n$  under  $H_0$  is

$$E_0 \left[ \frac{1}{n} \ln L_n \right] = \frac{1}{n} \ln \frac{|\Sigma_0|}{|\Sigma_1|} + \frac{1}{2} E_0 [(\xi_{n,K} - \mathbf{b}_0)^T \Sigma_0^{-1} (\xi_{n,K} - \mathbf{b}_0)] - \frac{1}{2} E_0 [(\xi_{n,K} - \mathbf{b}_1)^T \Sigma_1^{-1} (\xi_{n,K} - \mathbf{b}_1)] \quad (59)$$

where

$$\frac{1}{2} E_0 [(\xi_{n,K} - \mathbf{b}_0)^T \Sigma_0^{-1} (\xi_{n,K} - \mathbf{b}_0)] = \frac{K}{2n} \quad (60)$$

and

$$\begin{aligned} & \frac{1}{2} E_0 [(\xi_{n,K} - \mathbf{b}_1)^T \Sigma_1^{-1} (\xi_{n,K} - \mathbf{b}_1)] \\ &= \frac{1}{2} E_0 [(\xi_{n,K} - \mathbf{b}_0)^T \Sigma_1^{-1} (\xi_{n,K} - \mathbf{b}_0)] + \frac{1}{2} (\mathbf{b}_1 - \mathbf{b}_0)^T \Sigma_1^{-1} (\mathbf{b}_1 - \mathbf{b}_0) \\ &= \frac{1}{2n} \left[ \frac{\sigma_0^2 K (1 + (K - 2)\rho_1)}{2\sigma_1^2 (1 - \rho_1)(1 + (K - 1)\rho_1)} - \frac{\rho_0 \rho_1 \sigma_0^2 K (K - 1)}{2\sigma_1^2 (1 - \rho_1)(1 + (K - 1)\rho_1)} \right] + \frac{1}{2} (\mathbf{b}_1 - \mathbf{b}_0)^T \Sigma_1^{-1} (\mathbf{b}_1 - \mathbf{b}_0) \end{aligned} \quad (61)$$

Thus the asymptotic mean of  $\frac{1}{n} \ln L_n$  under  $H_0$  as  $n \rightarrow \infty$  becomes

$$\varphi_0 = \lim_{n \rightarrow \infty} E_0 \left[ \frac{1}{n} \ln L_n \right] = -\frac{1}{2} (\mathbf{b}_1 - \mathbf{b}_0)^T \Sigma_1^{-1} (\mathbf{b}_1 - \mathbf{b}_0) = -\frac{K[\mu_1(g) - \mu_0(g)]^2}{2[1 + (K - 1)\rho_1(g)]\sigma_1^2(g)} \quad (62)$$

Similarly, we can obtain the asymptotic mean of  $\frac{1}{n} \ln L_n$  under  $H_1$  as

$$\varphi_1 = \lim_{n \rightarrow \infty} E_1 \left[ \frac{1}{n} \ln L_n \right] = \frac{1}{2} (\mathbf{b}_1 - \mathbf{b}_0)^T \Sigma_0^{-1} (\mathbf{b}_1 - \mathbf{b}_0) = \frac{K[\mu_1(g) - \mu_0(g)]^2}{2[1 + (K - 1)\rho_0(g)]\sigma_0^2(g)}. \quad (63)$$

Again we assume the consistency condition  $\varphi_0 \leq \eta \leq \varphi_1$  for the threshold of the fusion center.

In the discrimination case the asymptotic rate of the error probabilities of the fusion center can not be obtained in closed form in the way done for the detection of weak signals (Section III). Instead, we use the lower bounds on the asymptotic rate obtained in Section II of [7] (see Lemma 1, Theorem 2, and Appendix A), where the large deviations principle was applied to yield

$$\lim_{n \rightarrow \infty} \left\{ -\frac{1}{n} \ln P_0 \left( \frac{1}{n} \ln L_n > \eta \right) \right\} \geq (\eta - \varphi_0) \tilde{s} \quad (64)$$

and

$$\lim_{n \rightarrow \infty} \left\{ -\frac{1}{n} \ln P_1 \left( \frac{1}{n} \ln L_n \leq \eta \right) \right\} \geq (\varphi_1 - \eta) \tilde{s} \quad (65)$$

for a small constant  $\tilde{s} > 0$ . Therefore, we use  $\min\{(\eta - \varphi_0)\tilde{s}, (\varphi_1 - \eta)\tilde{s}\}$  as a lower bound on the asymptotic rate. We require that the threshold of the fusion center  $\eta$  maximizes the above lower bound; this leads to an optimal threshold  $\eta^* = (\varphi_0 + \varphi_1)/2$  and a maximum value of the lower bound to the asymptotic rate  $\tilde{s}(\varphi_1 - \varphi_0)/2$ . Therefore, a suitable design criterion for the discrimination case is

$$J(\varphi_0, \varphi_1) = \frac{\varphi_1 - \varphi_0}{2} = \frac{K[\mu_1(g) - \mu_0(g)]^2}{4[1 + (K-1)\rho_1(g)]\sigma_1^2(g)} + \frac{K[\mu_1(g) - \mu_0(g)]^2}{4[1 + (K-1)\rho_0(g)]\sigma_0^2(g)}. \quad (66)$$

The optimal nonlinearity  $g_K$  which maximizes this design criterion satisfies a nonlinear integral equation (a necessary condition) but we can not derive sufficient conditions that guarantee that the solutions to this equation achieve the maximum. Instead of pursuing this path, we consider the following design criterion

$$\frac{K[\mu_1(g) - \mu_0(g)]^2}{[1 + (K-1)\rho_0(g)]\sigma_0^2(g) + [1 + (K-1)\rho_1(g)]\sigma_1^2(g)} = \bar{J}_K(g) \leq J(\varphi_0, \varphi_1). \quad (67)$$

which is lower bound on (66) as attested by the inequality (for all  $x, y > 0$ )

$$\frac{1}{4x} + \frac{1}{4y} = \frac{x+y}{4xy} \geq \frac{x+y}{(x+y)^2} = \frac{1}{x+y}$$

This lower bound becomes tight as  $\rho_0 \rightarrow \rho_1$  and  $\sigma_0^2 \rightarrow \sigma_1^2$  (e.g., in the case of detection of a weak signal): Thus the optimal nonlinearity is determined as

$$g_K^* = \max_{g_K} \bar{J}_K(g_K) = \max_{g_K} \frac{[\mu_1(g_K) - \mu_0(g_K)]^2}{[1 + (K-1)\rho_0(g_K)]\sigma_0^2(g_K) + [1 + (K-1)\rho_1(g_K)]\sigma_1^2(g_K)}. \quad (68)$$

The above maximization leads to the necessary condition characterized by the linear integral equation

$$[f_1(x) - f_0(x)]/[f_1(x) + f_0(x)] - \int \bar{K}_K(x, y) g_K^*(y) dy = g_K^*(x) \quad (69)$$

where the integration kernel takes the form

$$\begin{aligned} \bar{K}_K(x, y) = & \left\{ 2 \sum_{j=1}^{m_1} [f_1^{(j)}(x, y) - f_1(x)f_1(y)] + 2(K-1) \sum_{j=1}^{\tilde{m}_1} [\tilde{f}_1^{(j)}(x, y) - f_1(x)f_1(y)] \right. \\ & + 2 \sum_{j=1}^{m_0} [f_0^{(j)}(x, y) - f_0(x)f_0(y)] + 2(K-1) \sum_{j=1}^{\tilde{m}_0} [\tilde{f}_0^{(j)}(x, y) - f_0(x)f_0(y)] \\ & + (K-1) [\tilde{f}_1^{(0)}(x, y) - f_1(x)f_1(y)] + (K-1) [\tilde{f}_0^{(0)}(x, y) - f_0(x)f_0(y)] \\ & \left. - f_1(x)f_1(y) - f_0(x)f_0(y) \right\} / \{f_1(x) + f_0(x)\}. \end{aligned} \quad (70)$$

Sufficient conditions for the maximization of  $\bar{J}$  by the nonlinearity  $g_K^*$  are

$$\rho_i(g) > -\frac{1}{K-1} \quad i = 0, 1 \quad (71)$$

for all functions  $g$  for which  $\sigma_i^2(g)$  and  $\rho_i(g)$  converge and  $\sigma_i^2(g) > 0$  ( $i = 0, 1$ ).

Following arguments similar to the ones used for the proof of (54)-(55) we can show that

$$\bar{J}_K(g_K^*) \geq \bar{J}_{K-1}(g_{K-1}^*) \quad (72)$$

which implies that on the basis of the lower bound on the asymptotic rate for the error probabilities of the fusion center, the performance of the multi-sensor system improves (the error probability decreases) as the number of sensors increases.

## IV.2 Fusion of Quantized Observations

We consider the scheme displayed by Fig. C.2 for the general signal discrimination problem.

Let  $Q(X_l^{(k)})$  ( $l = 1, 2, \dots, n$ ),  $\mathbf{t} = [t_0 \ t_1 \ \dots \ t_M]$  and  $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_M]$  be defined as in Subsection III.2. Under  $H_i$  ( $i = 0, 1$ ), we have equations similar to the ones given by (35)-(37) namely

$$\begin{aligned} \tilde{\sigma}_i^2(Q) &= \sum_{l=1}^M (u_l)^2 \int_{t_{l-1}}^{t_l} f_i(x) dx + 2 \sum_{j=1}^{m_i} \sum_{r=1}^M \sum_{l=1}^M u_r u_l \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} f_i^{(j)}(x, y) dx dy \\ &\quad - (2m_i + 1) \left[ \sum_{l=1}^M u_l \int_{t_{l-1}}^{t_l} f_i(x) dx \right]^2, \end{aligned} \quad (73)$$

$$\begin{aligned} \tilde{\rho}_i(Q) \tilde{\sigma}_i^2(Q) &= E_i[Q(X_1^{(1)})Q(X_1^{(2)})] + 2 \sum_{j=1}^{\tilde{m}_i} E_i[Q(X_1^{(1)})Q(X_{j+1}^{(2)})] - (2\tilde{m}_i + 1)(E_i[Q(X_1^{(1)})])^2 \\ &= \sum_{r=1}^M \sum_{l=1}^M u_r u_l \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}_i^{(0)}(x, y) dx dy + 2 \sum_{j=1}^{\tilde{m}_i} \sum_{r=1}^M \sum_{l=1}^M u_r u_l \cdot \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}_i^{(j)}(x, y) dx dy \\ &\quad - (2\tilde{m}_i + 1) \left[ \sum_{l=1}^M u_l \int_{t_{l-1}}^{t_l} f_i(x) dx \right]^2 \end{aligned} \quad (74)$$

and

$$\tilde{\mu}_i(Q) = - \int Q(x) f_i(x) dx = \sum_{l=1}^M u_l \int_{t_{l-1}}^{t_l} f_i(x) dx. \quad (75)$$

For  $r, l = 1, 2, \dots, M$  and under  $H_i^{(k)}$  ( $i = 0, 1$ ), define the vector  $\mathbf{f}_i = [f_{1,i} \ f_{2,i} \ \dots \ f_{M,i}]$ , where  $f_{l,i} = \int_{t_{l-1}}^{t_l} f_i(x) dx$ , and the matrices  $\mathbf{F}_i = \text{diag}\{\int_{t_0}^{t_1} f_i(x) dx, \dots, \int_{t_{M-1}}^{t_M} f_i(x) dx\}$ ,  $\mathbf{P}_i = [P_{r,l,i}]$ ,  $\tilde{\mathbf{P}}_i = [\tilde{P}_{r,l,i}]$ ,  $\mathbf{R}_i = [R_{r,l,i}]$ , and  $\tilde{\mathbf{R}}_i = [\tilde{R}_{r,l,i}]$  where

$$P_{r,l,i} = 2 \sum_{j=1}^{m_i} \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} f_i^{(j)}(x, y) dx dy \quad (76)$$

$$R_{r,l,i} = (2m_i + 1) \int_{t_{r-1}}^{t_r} f_i(x) dx \cdot \int_{t_{l-1}}^{t_l} f_i(x) dx \quad (77)$$

$$\tilde{P}_{r,l,i} = \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}_i^{(0)}(x, y) dx dy + 2 \sum_{j=1}^{\tilde{m}_i} \int_{t_{r-1}}^{t_r} \int_{t_{l-1}}^{t_l} \tilde{f}_i^{(j)}(x, y) dx dy \quad (78)$$

and

$$\tilde{R}_{r,l,i} = (2\tilde{m}_i + 1) \cdot \int_{t_{r-1}}^{t_r} f_i(x)dx \cdot \int_{t_{l-1}}^{t_l} f_i(x)dx. \quad (79)$$

For the quantizer  $Q$ , the objective function  $\bar{J}_K(g)$  given by (67) becomes

$$\bar{J}_K(Q) = \frac{[u(f_1 - f_0)^T]^2}{u[\sum_{i=0}^1(F_i + P_i + \tilde{P}_i - R_i - \tilde{R}_i)]u^T}. \quad (80)$$

Assuming that  $t$  are such that the matrix  $\sum_{i=0}^1(F_i + P_i + \tilde{P}_i - R_i - \tilde{R}_i)$  is positive definite, we again fix the breakpoints first and derive the optimal quantization levels for any sensor as

$$u^T = \left[ \sum_{i=0}^1 (F_i + P_i + P_i - R_i - \tilde{R}_i) \right]^{-1} (f_1 - f_0)^T \quad (81)$$

Then, upon substitution for the above optimal levels into  $\bar{J}_K$ , we obtain the performance measure as a function of the breakpoints  $t$  as follows

$$\bar{J}_K(Q) = (f_1 - f_0) \left[ \sum_{i=0}^1 (F_i + P_i + P_i - R_i - R_i) \right]^{-1} (f_1 - f_0)^T. \quad (82)$$

The maximization of this expression with respect to the breakpoints can be accomplished via numerical optimization techniques.

It not difficult to show that, if  $Q_K^*$  denotes the optimal quantizer (breakpoints and quantization levels) for the  $K$ -sensor system, then  $\bar{J}_K(Q_K^*) \geq \bar{J}_{K-1}(Q_{K-1}^*)$  and thus as the number of sensors increases the performance of the multi-sensor system improves.

### IV.3 Fusion of Binary Sensor Decisions

In this case of general signal discrimination via the scheme illustrated in Fig. C.3, the majority vote fusion rule discussed in Subsection III.3 is employed. Following arguments similar to the ones in Subsection III.3 we can derive the asymptotic rate of the error probability of the fusion center as [cf. to (A.21) of Appendix B]

$$-\frac{1}{n} \ln P_{K,\eta_s} \rightarrow \min \left\{ -\frac{1}{n} \ln P_0 \{ (K-i) \text{ "0"s are transmitted} \}, \right.$$



$$\begin{aligned}
& -\frac{1}{n} \ln P_1 \{ i \text{ "0"s are transmitted } \}; \text{ for } i = \lceil K/2 \rceil, \lceil K/2 \rceil + 1, \dots, K \} \\
& = \min \left\{ \frac{\lceil K/2 \rceil (\eta_s - \mu_0)^2}{[1 + (\lceil K/2 \rceil - 1) \rho_0] \sigma_0^2}, \frac{\lceil K/2 \rceil (\mu_1 - \eta_s)^2}{[1 + (\lceil K/2 \rceil - 1) \rho_1] \sigma_1^2} \right\}
\end{aligned} \tag{83}$$

where the optimal threshold employed by all sensors is determined as

$$\begin{aligned}
\eta_s^* &= \max_{\mu_0 \leq \eta_s \leq \mu_1} \left\{ -\frac{1}{n} \ln P_{K, \eta_s} \right\} \\
&= \max_{\mu_0 \leq \eta_s \leq \mu_1} \min \left\{ \frac{\lceil K/2 \rceil (\eta_s - \mu_0)^2}{[1 + (\lceil K/2 \rceil - 1) \rho_0] \sigma_0^2}, \frac{\lceil K/2 \rceil (\mu_1 - \eta_s)^2}{[1 + (\lceil K/2 \rceil - 1) \rho_1] \sigma_1^2} \right\} \\
&= \frac{\mu_0 \sqrt{1 + (\lceil K/2 \rceil - 1) \rho_0} \sigma_1 + \mu_1 \sqrt{1 + (\lceil K/2 \rceil - 1) \rho_1} \sigma_0}{\sqrt{1 + (\lceil K/2 \rceil - 1) \rho_0} \sigma_0 + \sqrt{1 + (\lceil K/2 \rceil - 1) \rho_1} \sigma_1}.
\end{aligned} \tag{84}$$

for any  $K$ . Upon substituting of this threshold into (83) we obtain the asymptotic rate of the error probability of the fusion center as

$$-\frac{1}{n} \ln P_{K, \eta_s^*} \rightarrow \frac{[\mu_1(g) - \mu_0(g)]^2}{\sqrt{1 + (\lceil K/2 \rceil - 1) \rho_0(g)} \sigma_0(g) + \sqrt{1 + (\lceil K/2 \rceil - 1) \rho_1(g)} \sigma_1(g)} \tag{85}$$

Maximizing the above asymptotic rate directly (via calculus of variations techniques) results in a nonlinear integral equation that the optimal nonlinearity must satisfy and in very complex sufficient conditions. Therefore, in order to avoid these complications we consider the following lower bound on the asymptotic rate of (85)

$$\frac{1}{2} \cdot \frac{[\mu_1(g) - \mu_0(g)]^2}{[1 + (\lceil K/2 \rceil - 1) \rho_0(g)] \sigma_0^2(g) + [1 + (\lceil K/2 \rceil - 1) \rho_1(g)] \sigma_1^2(g)}. \tag{86}$$

[obtained from an application of the inequality  $\frac{1}{(x+y)^2} \geq \frac{1}{2(x^2+y^2)}$ ]. This lower bound becomes tight as  $\sigma_0^2 \rightarrow \sigma_1^2$  and  $\rho_0 \rightarrow \rho_1$ ; this corresponds to the weak-signal detection case. Finally, the optimal nonlinearity is obtained as

$$g_K^* = \begin{cases} \max_g \frac{[\mu_1(g) - \mu_0(g)]^2}{[2 + (K-1) \rho_0(g)] \sigma_0^2(g) + [2 + (K-1) \rho_1(g)] \sigma_1^2(g)}, & \text{if } K \text{ is odd} \\ \max_g \frac{[\mu_1(g) - \mu_0(g)]^2}{[2 + (K-2) \rho_0(g)] \sigma_0^2(g) + [2 + (K-2) \rho_1(g)] \sigma_1^2(g)}, & \text{if } K \text{ is even.} \end{cases} \tag{87}$$

The above maximization leads to the linear integral equation

$$[f_1(x) - f_0(x)]/[f_1(x) + f_0(x)] - \int \bar{H}_K(x, y) g_K^*(y) dy = g_K^*(x) \tag{88}$$

where

$$\begin{aligned}
\bar{H}_K(x, y) = & \left\{ 2 \sum_{j=1}^{m_1} [f_1^{(j)}(x, y) - f_1(x)f_1(y)] + (K-1) \sum_{j=1}^{\bar{m}_1} [\tilde{f}_1^{(j)}(x, y) - f_1(x)f_1(y)] \right. \\
& + 2 \sum_{j=1}^{m_0} [f_0^{(j)}(x, y) - f_0(x)f_0(y)] + (K-1) \sum_{j=1}^{\bar{m}_0} [\tilde{f}_0^{(j)}(x, y) - f_0(x)f_0(y)] \\
& + \frac{K-1}{2} [\tilde{f}_1^{(0)}(x, y) - f_1(x)f_1(y)] + \frac{K-1}{2} [\tilde{f}_0^{(0)}(x, y) - f_0(x)f_0(y)] \\
& \left. - f_1(x)f_1(y) - f_0(x)f_0(y) \right\} / \{f_1(x) + f_0(x)\}. \tag{89}
\end{aligned}$$

if  $K$  is odd, and

$$\begin{aligned}
\bar{H}_K(x, y) = & \left\{ 2 \sum_{j=1}^{m_1} [f_1^{(j)}(x, y) - f_1(x)f_1(y)] + (K-2) \sum_{j=1}^{\bar{m}_1} [\tilde{f}_1^{(j)}(x, y) - f_1(x)f_1(y)] \right. \\
& + 2 \sum_{j=1}^{m_0} [f_0^{(j)}(x, y) - f_0(x)f_0(y)] + (K-2) \sum_{j=1}^{\bar{m}_0} [\tilde{f}_0^{(j)}(x, y) - f_0(x)f_0(y)] \\
& + \frac{K-2}{2} [\tilde{f}_1^{(0)}(x, y) - f_1(x)f_1(y)] + \frac{K-2}{2} [\tilde{f}_0^{(0)}(x, y) - f_0(x)f_0(y)] \\
& \left. - f_1(x)f_1(y) - f_0(x)f_0(y) \right\} / \{f_1(x) + f_0(x)\}. \tag{90}
\end{aligned}$$

if  $K$  is even. The sufficient conditions for the maximization in question are

$$\rho_i(g) > -\frac{1}{\lceil K/2 \rceil - 1} \quad ; \text{ for } i = 0, 1 \tag{91}$$

for all functions  $g$  for which  $\sigma_i^2(g)$  and  $\rho_i(g)$  converge and  $\sigma_i^2(g) > 0$  ( $i = 0, 1$ ).

Finally, following arguments similar to those used for the proof of (54)-(55) we can show that

$$\lim_{n \rightarrow \infty} \left\{ -\frac{1}{n} \hat{P}_{K, \eta_s^*} \right\} \geq \lim_{n \rightarrow \infty} \left\{ -\frac{1}{n} \hat{P}_{K-1, \eta_s^*} \right\} \tag{92}$$

for any  $K$  (odd or even), where  $\hat{P}$  denotes the upper bound on the error probability (lower bound on the asymptotic rate) of the fusion center obtained in this section. Thus, under a large sample size, the upper bound on the error probability of the fusion center decreases as the number of sensors increases.

## V. PERFORMANCE EVALUATION

To establish the usefulness of the data fusion schemes derived in the previous sections for signal detection problems, we evaluate the performance of the three schemes via computer simulation and compare it to that of the corresponding schemes which neglect dependence in sensor observations.

The simulation results presented in this section pertain to the detection of a weak signal in additive Cauchy  $\rho$ -mixing noise.

The two hypotheses are assumed to have equal a priori probabilities, which results in selecting  $p = 0.5$ . The two types of error probabilities, the probabilities of missing ( $\bar{P}_1(0)$ ) and of false alarm ( $\bar{P}_0(1)$ ) for the fusion center, are used as the performance measure. In all examples, 1000 simulations are run for generating all results in this section and receiver operating characteristics (ROCs) are provided for comparison. To obtain good performance ( $10^{-2}$  order of error probabilities), 5000 samples are collected for weak signal detection (with SNR=-72.5 dB, as seen in the following). In our simulations, the magnitude of noise (in the weak-signal detection formulation) is forced to be between  $X_{min}$  and  $X_{max}$ . In this way, very small or very large samples (in the tails of the distributions) caused by deficiencies in the random-number generators are discarded and good quality simulated data are fed into our schemes. The appropriate values of  $X_{min}$  and  $X_{max}$  are derived for the specific noise or observation pdfs by requiring that  $P_{k,i}\{X \leq X_{min}\} \leq \epsilon$  and  $P_{k,i}\{X \geq X_{max}\} \leq \epsilon$ , for  $i = 0, 1$  and  $k = 1, 2, \dots, K$ , where  $\epsilon = 10^{-6}$ .

Without loss of generality, the magnitude of the weak signal  $\theta$  is taken to be 0.2, which implies a signal power of  $-20dB$ . The univariate density of the Cauchy noise is given by

$$f_k(x^{(k)}) = \frac{1}{\pi(1 + (x^{(k)})^2)}, \quad k = 1, 2, \dots, K$$

Although the second-order joint density of Cauchy noise is difficult to derive directly, it

can be evaluated from the second-order joint density of a Gaussian process by a nonlinear transformation. Let

$$f_G(y) = \frac{\exp[-(y)^2/2]}{\sqrt{2\pi}}$$

be the univariate density of the underlying Gaussian process, Then the aforementioned nonlinear transformation has the form (see [11])

$$z(y) = \tan[\pi \cdot \text{erf}(y/\sqrt{2})/2].$$

The range of the Gaussian random variable  $x$  is taken to be  $[-5, 5]$  and the pair  $X_{min}$  and  $X_{max}$  mentioned above is obtained from  $z(-5)$  and  $z(5)$ . We can compute the noise power for the purpose of our simulations by limiting the noise in the range  $[X_{min}, X_{max}]$  as  $\int_{X_{min}}^{X_{max}} z^2 f_k(z) dz = \int_{-5}^5 [z(x)]^2 f_G(x) dx$ , where  $z(x)$  was defined above. From this noise power calculation, we obtain a SNR of  $-72.5dB$ .

We generate the underlying Gaussian process for the Cauchy noise process of each sensor by the following autoregressive model with correlation coefficients  $-1 \leq \rho_{G,k} \leq 1$ , for  $k = 1, 2, \dots, K$ ,

$$\begin{aligned} \tilde{N}_1^{(k)} &= V_1^{(k)} \\ \tilde{N}_i^{(k)} &= \rho_{G,k} \tilde{N}_{i-1}^{(k)} + \sqrt{1 - \rho_k^2} V_i^{(k)}; i > 1. \end{aligned}$$

In the above recursion formula,  $V_i^{(k)}$ , for  $i = 1, 2, \dots, n$ , are sequences of i.i.d. random variables with standard Gaussian ( $\mathcal{N}(0, 1)$ ) densities and are generated in the manner described by

$$\begin{aligned} V_i^{(2)} &= \rho_c V_i^{(1)} + \sqrt{1 - \rho_c^2} W_i^{(2)} \\ V_i^{(3)} &= \frac{\rho_c}{1 + \rho_c} V_i^{(1)} + \frac{\rho_c}{1 + \rho_c} V_i^{(2)} + \frac{\sqrt{1 + 2\rho_c - \rho_c^2 - \rho_c^3}}{1 + \rho_c} W_i^{(3)} \\ &\vdots \\ V_i^{(k)} &= a_i^{(k)} V_i^{(1)} + a_i^{(k)} V_i^{(2)} + \dots + a_i^{(k)} V_i^{(k-1)} + b_i^{(k)} W_i^{(k)} \end{aligned}$$

where

$$\begin{aligned} a_i^{(k)} &= \frac{\rho_c}{1 + (k-2)\rho_c} \\ b_i^{(k)} &= \frac{1}{1 + (k-2)\rho_c} \left[ 1 + 2(k-2)\rho_c + (k^2 - 5k + 5)\rho_c^2 - \frac{(k-1)(k-2)}{2}\rho_c^3 \right], \end{aligned}$$

$W_i^{(k)}$  is another i.i.d. standard Gaussian ( $\mathcal{N}(0,1)$ ) process, for each  $k$  ( $k = 2, 3, \dots, K$ ), and  $-1 \leq \rho_{G,c} \leq 1$  is the correlation coefficient characterizing the dependence across sensors;  $V_i^{(1)}$  and  $W_i^{(k)}$  are mutually independent. Under the nonlinear transformation  $N_i^{(k)} = z(\tilde{N}_i^{(k)})$ ,  $\{N_i^{(k)}\}_{n=1}^n$ , for each,  $k$  is a  $\rho$ -mixing Cauchy noise sequence.

In Example 2, we employ the 16-level quantizers that are symmetric with respect to zero for symmetric Cauchy noise. Therefore, in the following tables for quantization levels and breakpoints, only the half of positive breakpoints is given.

**Example 1:** This example pertains to Scheme 1. Fig. 1.1 gives the optimal nonlinearities  $g_{opt}$  for different  $K$  and the one obtained by ignoring dependence across time and sensors,  $g_{iid}$ . Figs. 1.2 and 1.3 present the comparison of ROCs for  $K = 6$  and  $K = 3$ , where  $P_{01}$  and  $P_{10}$  are the probabilities of false alarm and miss, respectively, and  $g_{ias}$  is the optimal nonlinearity obtained by ignoring the dependence across sensors. Then in Fig. 1.4, the comparison of ROCs for  $K = 1, 2, 3, 6$  by using the optimal nonlinearity is given.

**Example 2:** This example pertains to Scheme 2. Tables 1 and 2 give the optimal 16-level quantization levels and breakpoints for optimal quantizers  $q_{opt}$  and  $K = 1, 2, 3, 6$ . Table 3 gives the ones for the optimal nonlinearity obtained under i.i.d. observations. In Figs. 2.1 and 2.2, the comparisons of ROCs for  $K = 6$  and  $K = 3$  are presented for schemes using  $q_{opt}$ ,  $q_{ias}$  (the optimal nonlinearity obtained by ignoring the dependence across sensors) and  $q_{iid}$ . Fig. 2.3 presents the comparison of ROCs for  $K = 1, 2, 3, 6$  by using corresponding  $q_{opt}$ .

**Example 3:** This example pertains to Scheme 3. Fig. 3.1 gives the optimal nonlinearities  $g_{opt}$

for  $K = 1, 2, 3, 6$  and  $g_{iid}$ . Figs. 3.2 and 3.3 present the comparison of ROCs for  $K = 6$  and  $K = 3$  by using  $g_{opt}$ ,  $g_{ias}$  and  $g_{iid}$ , respectively. Then, Fig. 3.4 gives the comparison of ROCs for  $K = 1, 2, 3, 6$  by using corresponding  $g_{opt}$ .

Finally, in Figs. 4.1 and 4.2, the comparisons of ROCs for  $K = 6$  and  $K = 3$  are given for the three different schemes.

## VI. CONCLUSIONS

We have designed and analyzed the performance of multi-sensor fusion and quantization schemes with  $K$  identical sensors for the detection of weak signals in dependent stationary noise and signal discrimination under symmetric conditions. These schemes employ memoryless nonlinearities and take advantage of dependence in the observations across time and/or sensors for a better performance.

The dependence of the sensor observations is characterized by stationary  $m$ -dependent,  $\phi$ -mixing, or  $\rho$ -mixing sequences. The performance is measured by the two types of error probabilities of the fusion center. The optimal nonlinearities and/or quantizers are obtained by solving uncoupled or coupled linear integral equations that involve the univariate and bivariate pdfs of sensor observations.

As seen from Figs. 1.4, 2.3 and 3.4, performance improves as the number of sensors  $K$  increases. As expected, Figs 4.1 and 4.2 show that fusing the unquantized test statistics is superior to all schemes. Another interesting result from Figs. 4.1 and 4.2 is that Scheme 3 outperforms Scheme 2 in the context of ROCs. Furthermore, schemes which consider dependence across time and sensors outperform those that ignore dependence across sensors and also significantly outperform those that ignore dependence across time and sensors.

## Appendix A: Asymptotic Rates of Joint Error Probabilities for Scheme 3

In the following, we use the large deviations principle cited in Subsection II.2 to characterize the asymptotic rate of the error probabilities of the fusion center for the scheme displayed by Fig. C.3.

Since  $\xi$  is multi-dimensional Gaussian distributed, the corresponding  $I_0(z)$  given by (13) with dimension  $d = K$  is

$$\begin{aligned}
I_0(z) &= \frac{1}{2}(z - E_0[\xi])^T \overline{\text{cov}}_0[\xi]^{-1}(z - E_0[\xi]) \\
&= \frac{1}{2}(z - b_0)^T \Sigma_0^{-1}(z - b_0) \\
&= \frac{1}{2} \left[ (1 + (K-1)\rho_0) \sum_{j=1}^K \frac{(z_j - \mu_0)^2}{\sigma_0^2(1 - \rho_0)(1 + (K-1)\rho_0)} - \rho_0 \sum_{j=1}^K \sum_{l=1, l \neq j}^K \frac{(z_j - \mu_0)(z_l - \mu_0)}{\sigma_0^2(1 - \rho_0)(1 + (K-1)\rho_0)} \right] \\
&= \frac{1}{2} \left[ (1 + (K-1)\rho_0) \sum_{j=1}^K u_j^2 - \rho_0 \sum_{j=1}^K \sum_{l=1, l \neq j}^K u_j u_l \right] = \tilde{I}_0(u) \tag{A.1}
\end{aligned}$$

where we denote the normalized  $z_j$  by  $u_j = (z_j - \mu_0) / \{\sigma_0 \sqrt{(1 - \rho_0)[1 + (K-1)\rho_0]}\}$ , for  $j = 1, 2, \dots, K$ .

If there are  $i$  0 s (or  $K - i$  1 s) in the sensor decision vector  $(d_{n,1}, d_{n,2}, \dots, d_{n,K})$  which is transmitted to the fusion center, then the asymptotic rate in (40) is

$$\begin{aligned}
& -\frac{1}{n} \ln P_0(d_{n,1}, d_{n,2}, \dots, d_{n,K}) = -\frac{1}{n} \ln P_0\{i \text{ 0s are transmitted}\} \\
& = -\frac{1}{n} \ln P_0(T_{n,1} \leq \eta_s, T_{n,2} \leq \eta_s, \dots, T_{n,i} \leq \eta_s, T_{n,i+1} > \eta_s, \dots, T_{n,K} > \eta_s) \tag{A.2}
\end{aligned}$$

where  $T_{n,k}$ ; for  $k = 1, 2, \dots, K$ , are Gaussian distributed with identical individual means, variances, and pair-wise correlation coefficients. Let

$$\begin{aligned}
A &= \{z \in \mathbf{R}^K : z_1 \leq \eta_s, z_2 \leq \eta_s, \dots, z_i \leq \eta_s, z_{i+1} > \eta_s, \dots, z_K > \eta_s\} \\
&= \{u \in \mathbf{R}^K : u_1 \leq \eta_{s,0}, u_2 \leq \eta_{s,0}, \dots, u_i \leq \eta_{s,0}, u_{i+1} > \eta_{s,0}, \dots, u_K > \eta_{s,0}\} \tag{A.3}
\end{aligned}$$

where  $\eta_{s,0} = (\eta_s - \mu_0) / \{\sigma_0 \sqrt{(1 - \rho_0)[1 + (K - 1)\rho_0]}\}$  is the normalized threshold. Then, from the large deviations principle given by (10), we have

$$\begin{aligned}
& -\frac{1}{n} \ln P_0(T_{n,1} \leq \eta_s, T_{n,2} \leq \eta_s, \dots, T_{n,i} \leq \eta_s, T_{n,i+1} > \eta_s, \dots, T_{n,K} > \eta_s) = \\
& = -\frac{1}{n} \ln P_0(\xi \in A) = \inf_{\mathbf{z} \in A} I_0(\mathbf{z}) = \inf_{\mathbf{u} \in A} \tilde{I}_0(\mathbf{u}) \\
& = \frac{1}{2} \inf_{\mathbf{u} \in A} \left\{ (1 + (K - 1)\rho_0) \sum_{j=1}^K u_j^2 - \rho_0 \sum_{j=1}^K \sum_{l=1, l \neq j}^K u_j u_l \right\} \\
& = \frac{1}{2} \inf_{\mathbf{u} \in A} \left\{ (1 + (K - 1)\rho_0) \frac{1 + (K - i - 1)\rho_0}{1 + (K - i)\rho_0} \sum_{j=1}^i \left[ u_j - \frac{\rho_0}{1 + (K - j - 1)\rho_0} \sum_{l=j+1}^K u_l \right]^2 \right. \\
& \quad \left. + \frac{1 + (K - 1)\rho_0}{1 + (K - i - 1)\rho_0} \left[ (1 + (K - i - 2)\rho_0) \sum_{j=i+1}^K u_j^2 - \rho_0 \sum_{j=i+1}^K \sum_{l=i+1, l \neq j}^K u_j u_l \right] \right\} \tag{A.4}
\end{aligned}$$

Define the vector  $\mathbf{u}^{(i)} = [u_{i+1} \ u_{i+2} \ \dots u_K]^T$ , which belongs to  $\mathbf{R}^{(K-i)}$  and the set  $A_i = \{(u_{i+1}, u_{i+2}, \dots, u_K) \in \mathbf{R}^{(K-i)} : u_{i+1} \geq \eta_{s,0}, u_{i+2} \geq \eta_{s,0}, \dots, u_K \geq \eta_{s,0}\}$ . Apparently,

$$\inf_{\mathbf{u} \in A} \tilde{I}_0(\mathbf{u}) \geq \inf_{\mathbf{u}_i \in A_i} \tilde{I}_0^{(i)}(\mathbf{u}_i) \tag{A.5}$$

where

$$\tilde{I}_0^{(i)}(\mathbf{u}_i) = \frac{1 + (K - 1)\rho_0}{2(1 + (K - i - 1)\rho_0)} \left[ (1 + (K - i - 2)\rho_0) \sum_{j=i+1}^K u_j^2 - \rho_0 \sum_{j=i+1}^K \sum_{l=i+1, l \neq j}^K u_j u_l \right]. \tag{A.6}$$

We notice that the set  $A_i$  is symmetric with respect to  $u_j$ , for  $j = i + 1, i + 2, \dots, K$ . Thus it is easy to verify that

$$\mathbf{u}_i^* = [u_{i+1}^* \ u_{i+2}^* \ \dots u_K^*]^T = \arg \inf_{\mathbf{u}_i \in A_i} \tilde{I}_0^{(i)}(\mathbf{u}_i) \tag{A.7}$$

satisfies  $u_{i+1}^* = u_{i+2}^* = \dots = u_K^*$ . Therefore,

$$\begin{aligned}
\inf_{\mathbf{u}_i \in A_i} \tilde{I}_0^{(i)}(\mathbf{u}_i) & = \inf_{\mathbf{u}_i \in A_i} \left\{ \frac{1 + (K - 1)\rho_0}{2(1 + (K - i - 1)\rho_0)} \left[ (1 + (K - i - 2)\rho_0) \sum_{j=i+1}^K u_j^2 - \rho_0 \sum_{j=i+1}^K \sum_{l=i+1, l \neq j}^K u_j u_l \right] \right\} \\
& = \frac{1}{2} \inf_{u_{i+1} > \eta_{s,0}} \left\{ \frac{(K - i)(1 - \rho_0)(1 + (K - 1)\rho_0)}{1 + (K - i - 1)\rho_0} u_{i+1}^2 \right\}
\end{aligned}$$



$$\begin{aligned}
&= \frac{1}{2} \frac{(K-i)(1-\rho_0)(1+(K-1)\rho_0)}{1+(K-i-1)\rho_0} \eta_{s,0}^2 \\
&= \frac{1}{2} \frac{(K-i)(1-\rho_0)(1+(K-1)\rho_0)}{1+(K-i-1)\rho_0} \left[ \frac{\eta_s - \mu_0}{\sigma_0 \sqrt{(1-\rho_0)(1+(K-1)\rho_0)}} \right]^2
\end{aligned} \tag{A.8}$$

where we assumed that  $\rho_0 \geq -\frac{1}{K-1}$  such that  $(K-i)(1-\rho_0)(1+(K-1)\rho_0)/[1+(K-i-1)\rho_0]$  is positive for all  $i = 1, 2, \dots, K$ .

From (A.8), we obtain

$$u_j^* = \eta_{s,0} \quad \text{for } j = i+1, i+2, \dots, K, \tag{A.9}$$

with which we can easily check that if we define

$$u_j^* = \frac{\rho_0}{1+(K-j-1)\rho_0} \sum_{l=j+1}^K u_l^* = \frac{(K-i)\rho_0}{1+(K-i-1)\rho_0} \eta_{s,0} \quad \text{for } j = 1, 2, \dots, i, \tag{A.10}$$

(where for the second equality the computation is performed in reverse index order, that is, for  $j = i, i-1, \dots, 1$ ), then

$$\tilde{I}_0(\mathbf{u}^*) = \tilde{I}_0^{(i)}(\mathbf{u}_i^*). \tag{A.11}$$

Moreover, since  $u_j^* < \eta_{s,0}$  for  $j = 1, 2, \dots, i$ , the vector  $\mathbf{u}^*$  with  $u_j^*$  given by (A.9) and (A.10) belongs to the set  $A$  defined in (A.3). Consequently,

$$\inf_{\mathbf{u} \in A} \tilde{I}_0(\mathbf{u}) = \tilde{I}_0(\mathbf{u}^*) = \frac{(K-i)(1-\rho_0)(1+(K-1)\rho_0)}{2[1+(K-i-1)\rho_0]} \eta_{s,0}^2. \tag{A.12}$$

We therefore conclude that as  $n \rightarrow \infty$

$$\begin{aligned}
& -\frac{1}{n} \ln P_0\{i \text{ 0 s are transmitted}\} \\
&= -\frac{1}{n} \ln P_0(T_{n,1} \leq \eta_s, T_{n,2} \leq \eta_s, \dots, T_{n,i} \leq \eta_s, T_{n,i+1} > \eta_s, \dots, T_{n,K} > \eta_s) \\
&\rightarrow \inf_{\mathbf{u} \in A} \tilde{I}_0(\mathbf{u}) = \frac{(K-i)(1-\rho_0)(1+(K-1)\rho_0)}{2[1+(K-i-1)\rho_0]} \eta_{s,0}^2 \\
&= \frac{(K-i)}{1+(K-i-1)\rho_0} \cdot \frac{(\eta_s - \mu_0)^2}{2\sigma_0^2}.
\end{aligned} \tag{A.13}$$

Following similar arguments to the above ones we can verify that as  $n \rightarrow \infty$

$$-\frac{1}{n}P_1\{i \text{ 0s are transmitted}\} \rightarrow \frac{i}{1+(i-1)\rho_1} \cdot \frac{(\mu_1 - \eta_s)^2}{2\sigma_1^2}. \quad (\text{A.14})$$

## Appendix B: Proof of Proposition 1

If the optimal threshold is given by (47), then from (A.13) and (A.14) in Appendix A,

$$\lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_0\{i \text{ 0s are transmitted}\} \right] = \frac{(K-i)}{1+(K-i-1)\rho_0} \cdot \frac{(\mu_\theta - \mu_0)^2}{4\sigma_0^2}. \quad (\text{A.15})$$

and

$$\lim_{n \rightarrow \infty} \left[ -\frac{1}{n} P_1\{i \text{ 0s transmitted}\} \right] = \frac{i}{1+(i-1)\rho_1} \cdot \frac{(\mu_\theta - \mu_0)^2}{4\sigma_0^2}. \quad (\text{A.16})$$

Therefore, when there are  $i$  0s transmitted by the sensors, we have

$$\begin{aligned} -\frac{1}{n} \ln L_n &= \left[ -\frac{1}{n} \ln P_0\{i \text{ 0s are transmitted}\} \right] - \left[ -\frac{1}{n} P_1\{i \text{ 0s transmitted}\} \right] \\ &\rightarrow \frac{(K-i-i)(1-\rho_0)}{(1+(K-i-1)\rho_0)(1+(i-1)\rho_0)} \cdot \frac{(\mu_\theta - \mu_0)^2}{4\sigma_0^2}, \end{aligned} \quad (\text{A.17})$$

thus the asymptotic likelihood test of the fusion center  $\frac{1}{n} \ln L_n \stackrel{H_0}{\underset{H_1}{<}} 0$  is equivalent to the test  $K-i \stackrel{H_0}{\underset{H_1}{<}} i$ , which is the majority rule

$$\text{no. of 1s} \stackrel{H_1}{\underset{H_0}{>}} \text{no. of 0s} \quad (\text{A.18})$$

On the other hand, if the likelihood ratio test is equivalent to a majority rule of the fusion center, the error probability of the fusion center is

$$P_{K,\eta_s} = p \sum_{i=\lceil K/2 \rceil}^K P_0\{(K-i) \text{ 0s are transmitted}\} + (1-p) \sum_{i=\lceil K/2 \rceil}^K P_1\{i \text{ 0s are transmitted}\} \quad (\text{A.19})$$

for odd  $K$ , where  $\lceil K/2 \rceil = (K+1)/2$ , and

$$\begin{aligned} P_{K,\eta_s} &= p \sum_{i=K/2+1}^K P_0\{(K-i) \text{ 0s are transmitted}\} + (1-p) \sum_{i=K/2+1}^K P_1\{i \text{ 0s are transmitted}\} \\ &\quad + \frac{p}{2} P_0\{K/2 \text{ 0s are transmitted}\} + \frac{1-p}{2} P_1\{K/2 \text{ 0s are transmitted}\} \end{aligned} \quad (\text{A.20})$$

for even  $K$ , where we adopt the fusion rule of flipping a fair coin (and thus incurring an error probability of  $1/2$ ) when the number of 1s and the number of 0s are equal.

From (A.13) and (A.14) and for the special case of weak-signal detection, for which  $\mu_1 = \mu_\theta$ ,  $\sigma_1 = \sigma_0$ , and  $\rho_0 = \rho_1$ , we can obtain the asymptotic rates for each of the terms in the sums in the error probability expressions in (A.19) and (A.20). Since  $j/(1+(j-1)\rho_0) \leq k/(1+(k-1)\rho_0)$  for all  $j \leq k$ , the overall asymptotic rate of the total error probability of the fusion center is given by

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[ -\frac{1}{n} \ln P_{K, \eta_s} \right] &\rightarrow \min \left\{ -\frac{1}{n} \ln P_0 \{ (K-i) \text{ 0s are transmitted} \}, \right. \\ &\quad \left. -\frac{1}{n} \ln P_1 \{ i \text{ 0s are transmitted} \}; \text{ for } i = \lceil K/2 \rceil, \lceil K/2 \rceil + 1, \dots, K \right\} \\ &= \min \left\{ \frac{\lceil K/2 \rceil (\eta_s - \mu_0)^2}{[1 + (\lceil K/2 \rceil - 1)\rho_0]\sigma_0^2}, \frac{\lceil K/2 \rceil (\mu_\theta - \eta_s)^2}{[1 + (\lceil K/2 \rceil - 1)\rho_0]\sigma_0^2} \right\}, \end{aligned} \quad (\text{A.21})$$

for any  $K$ . Consequently, the optimal threshold  $\eta_s^*$  is obtained as

$$\begin{aligned} \eta_s^* &= \max_{\mu_0 \leq \eta_s \leq \mu_\theta} \left\{ -\frac{1}{n} \ln P_{K, \eta_s} \right\} \\ &= \max_{\mu_0 \leq \eta_s \leq \mu_\theta} \min \left\{ \frac{\lceil K/2 \rceil (\eta_s - \mu_0)^2}{[1 + (\lceil K/2 \rceil - 1)\rho_0]\sigma_0^2}, \frac{\lceil K/2 \rceil (\mu_\theta - \eta_s)^2}{[1 + (\lceil K/2 \rceil - 1)\rho_0]\sigma_0^2} \right\} \\ &= \frac{\mu_0 + \mu_\theta}{2} \end{aligned} \quad (\text{A.22})$$

for any  $K$ .

## References

- [1] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-17, pp. 501-510, July 1981.
- [2] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-22, pp. 98-101, January 1986.
- [3] A. R. Reibman and L. W. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-23, pp. 24-30, January 1987.
- [4] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol.2, 1988.
- [5] E. Geraniotis and Y. A. Chau, "Robust Data Fusion for Multi-Sensor Detection Systems," to appear in *IEEE Transactions on Information Theory*.
- [6] Y. A. Chau and E. Geraniotis, "Memoryless Quantization and Fusion in Multi-Sensor Systems for the Detection of Weak Signals in Dependent Noise," submitted to *IEEE Transactions on Information Theory*, 1989.
- [7] Y. A. Chau and E. Geraniotis, "Quantization and Fusion for Multi-Sensor Discrimination from Dependent Observations," submitted to *IEEE Transactions on Acoustics, Speech, Signal Processing*, 1990.
- [8] R. C. Bradley, "Basic properties of strong mixing conditions," *Dependence in Probability and Statistics*, Birkhauser, 1985.

- [9] M. Peligrad, "Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (a survey)," *Dependence in Probability and Statistics*, Birkhauser, 1985.
- [10] S. A. Kassam, "Optimum quantization for signal detection," *IEEE Trans. Commun.*, vol. COM-25, pp. 479-484, May 1977.
- [11] H. V. Poor and J. B. Thomas, "Memoryless quantizer-detectors for constant signals in  $m$ -dependent noise", *IEEE Trans. Inform. Theory* , vol. IT-26, pp. 423-432, July 1980.
- [12] J. S. Sadowsky and J. Bucklew, "A nonlocal approach for asymptotic memoryless detection theory," *IEEE Trans. Inform. Theory*, vol. IT-32, pp.115-120, Jan. 1986.
- [13] D. Sauder and E. Geraniotis, "Optimal and robust memoryless discrimination from dependent observations", submitted for publication to *IEEE Trans. Inform. Theory*, 1989.
- [14] W. A. Gardner, " A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM-28, pp. 807-816, June 1980.
- [15] R. S. Ellis, "Large deviations for a general class of random vectors," *Ann. Probab.*, vol. 12, pp. 1-12, 1976.

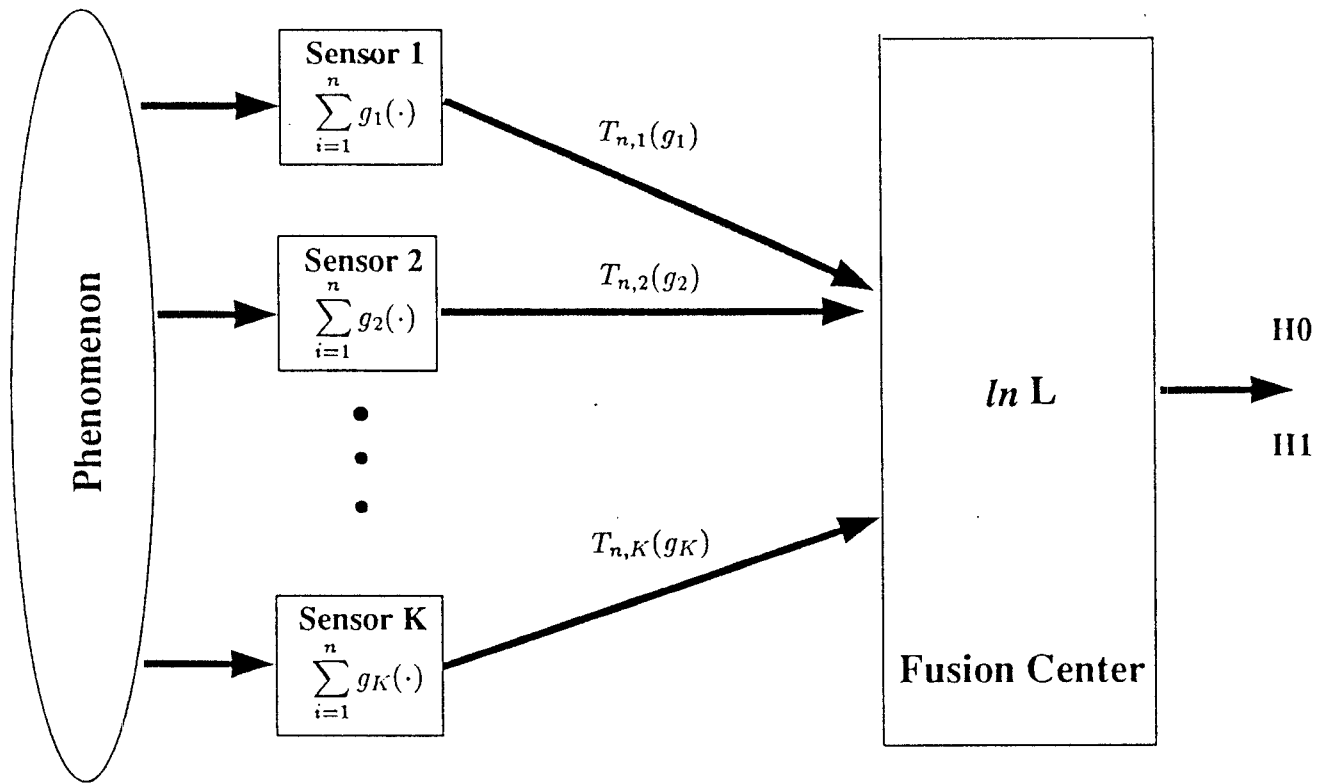


Fig. C.1 Fusion of Unquantized Sensor Test Statistics ( Scheme 1 )

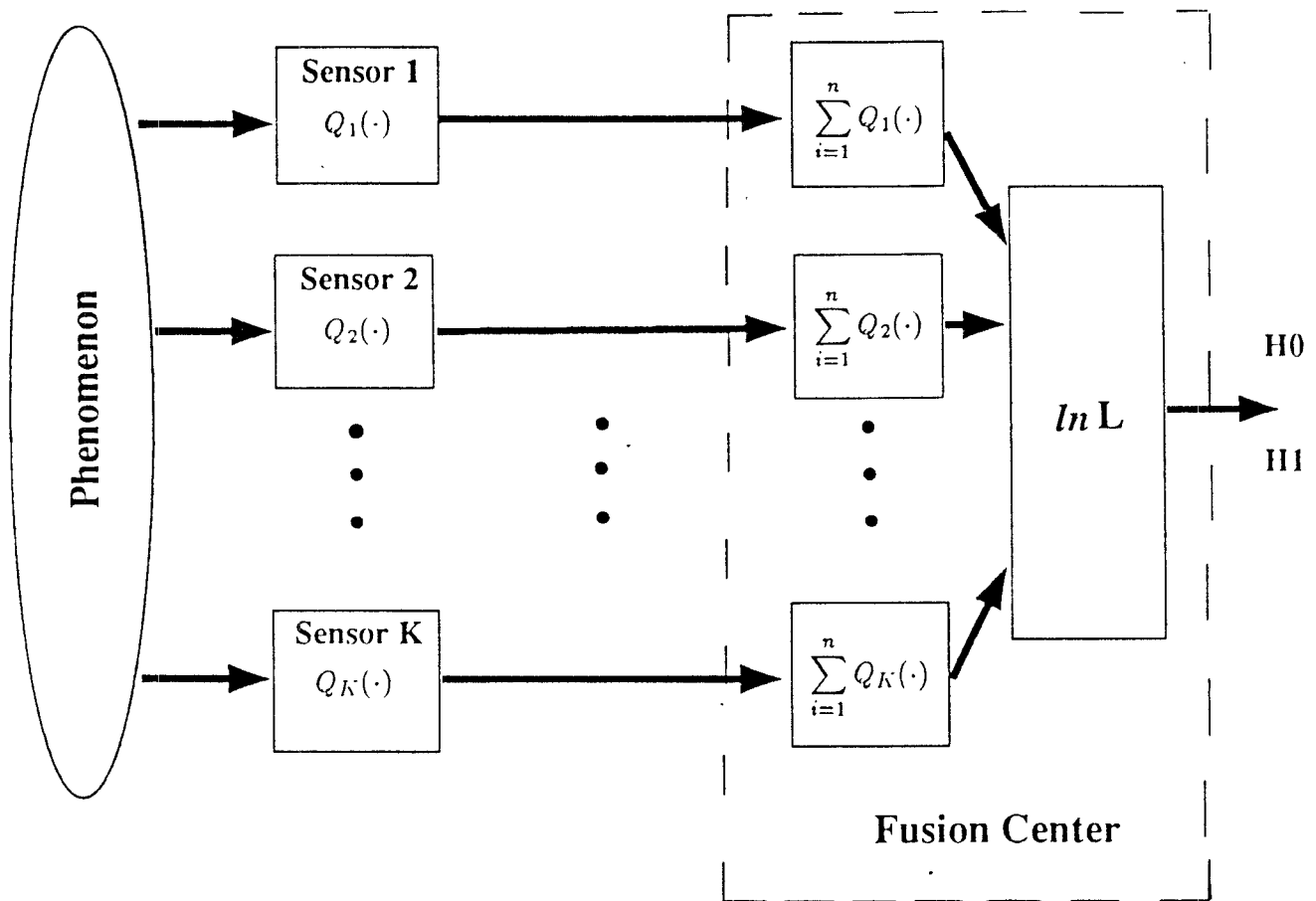


Fig. C.2 Fusion of Quantized Sensor Observations ( Scheme 2 )

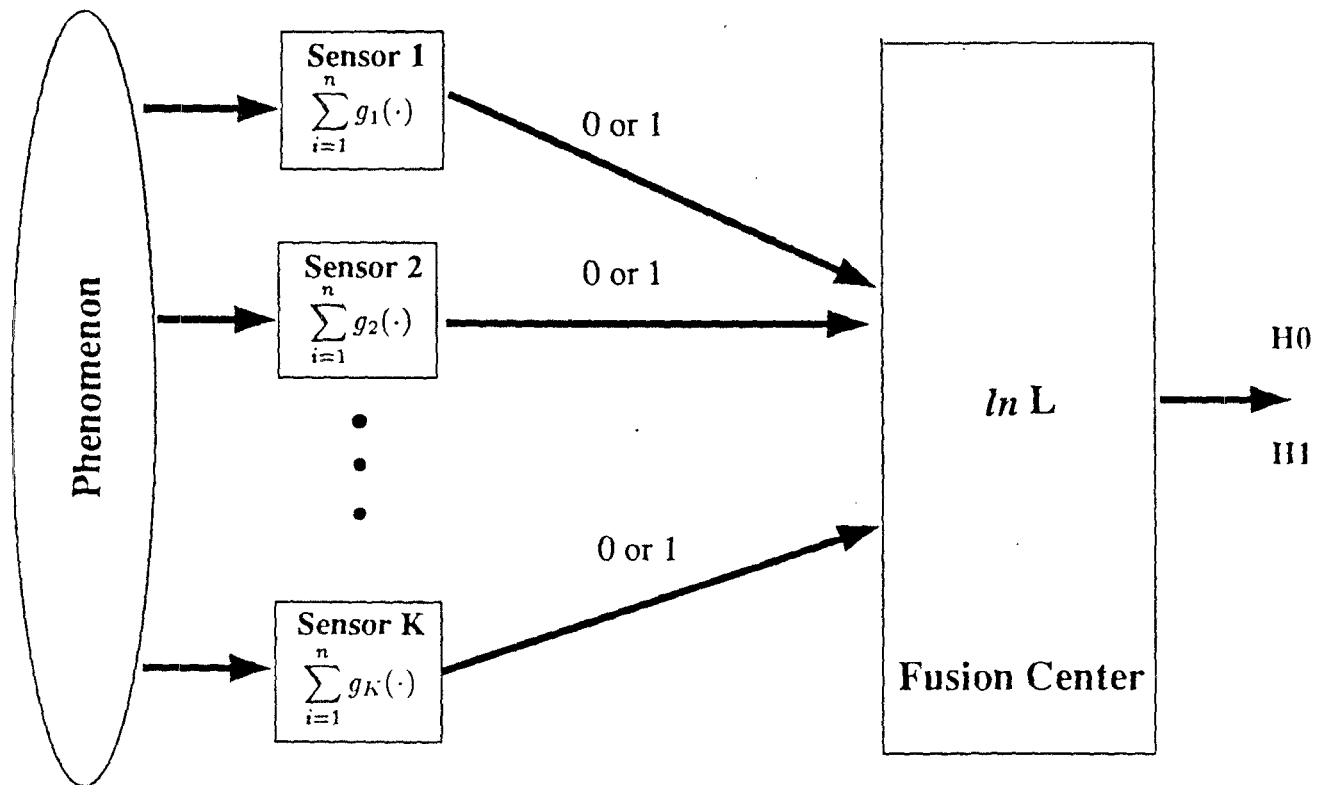


Fig. C.3 Fusion of Binary Sensor Decisions ( Scheme 3 )



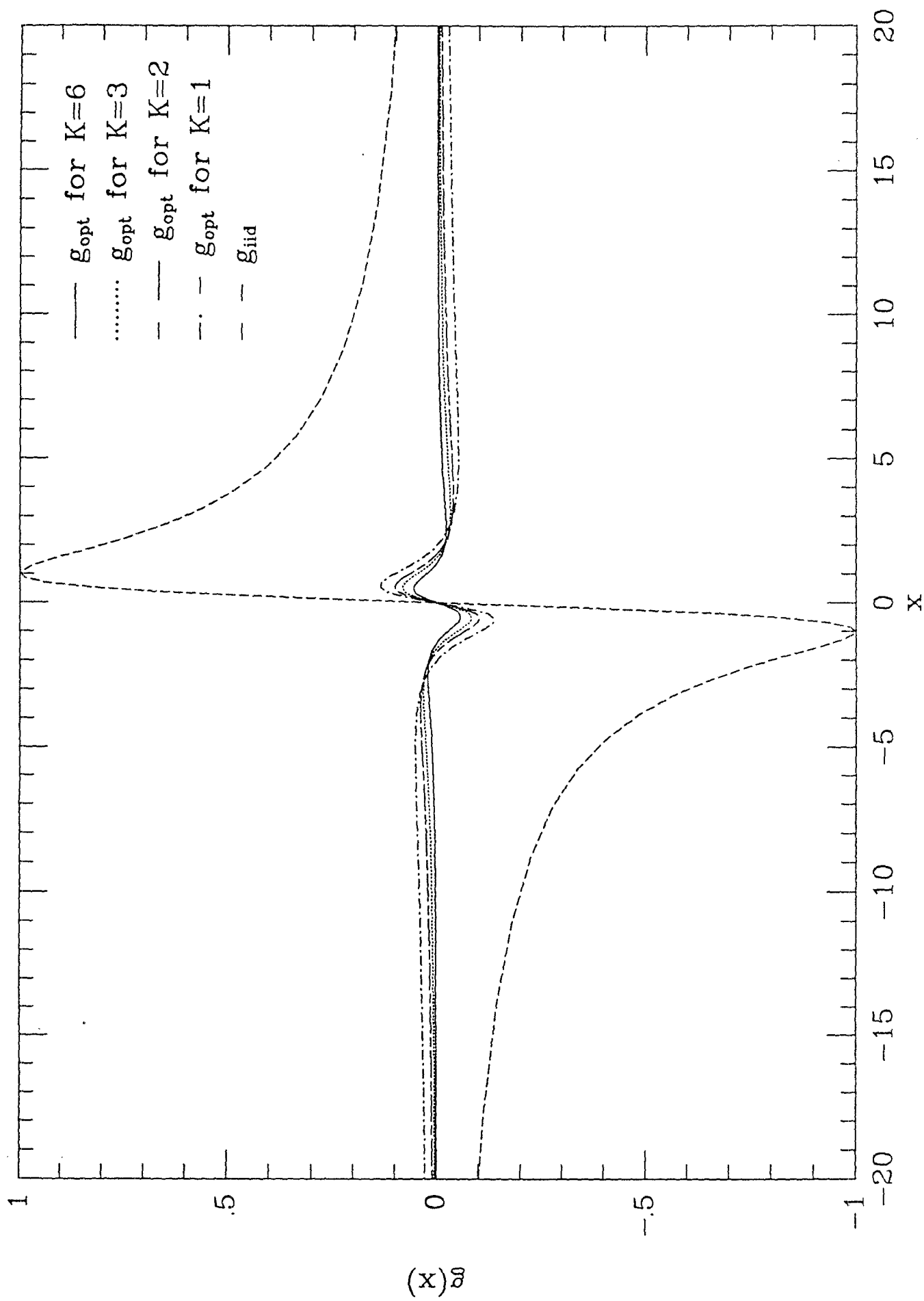


Fig. 1.1 Optimal Nonlinearities  $g_{\text{opt}}$  vs.  $g_{\text{iid}}$  in Scheme 1

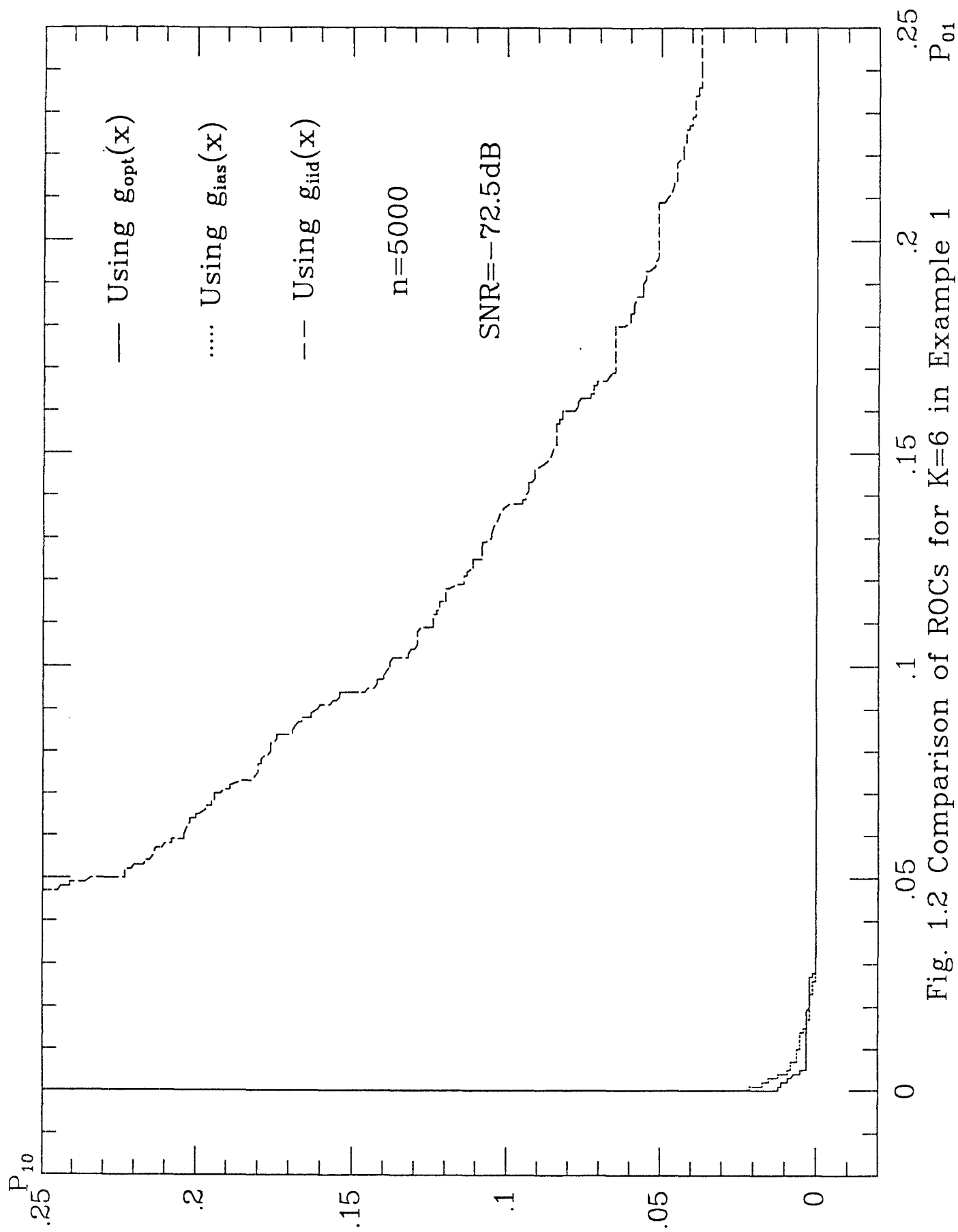


Fig. 1.2 Comparison of ROCs for  $K=6$  in Example 1

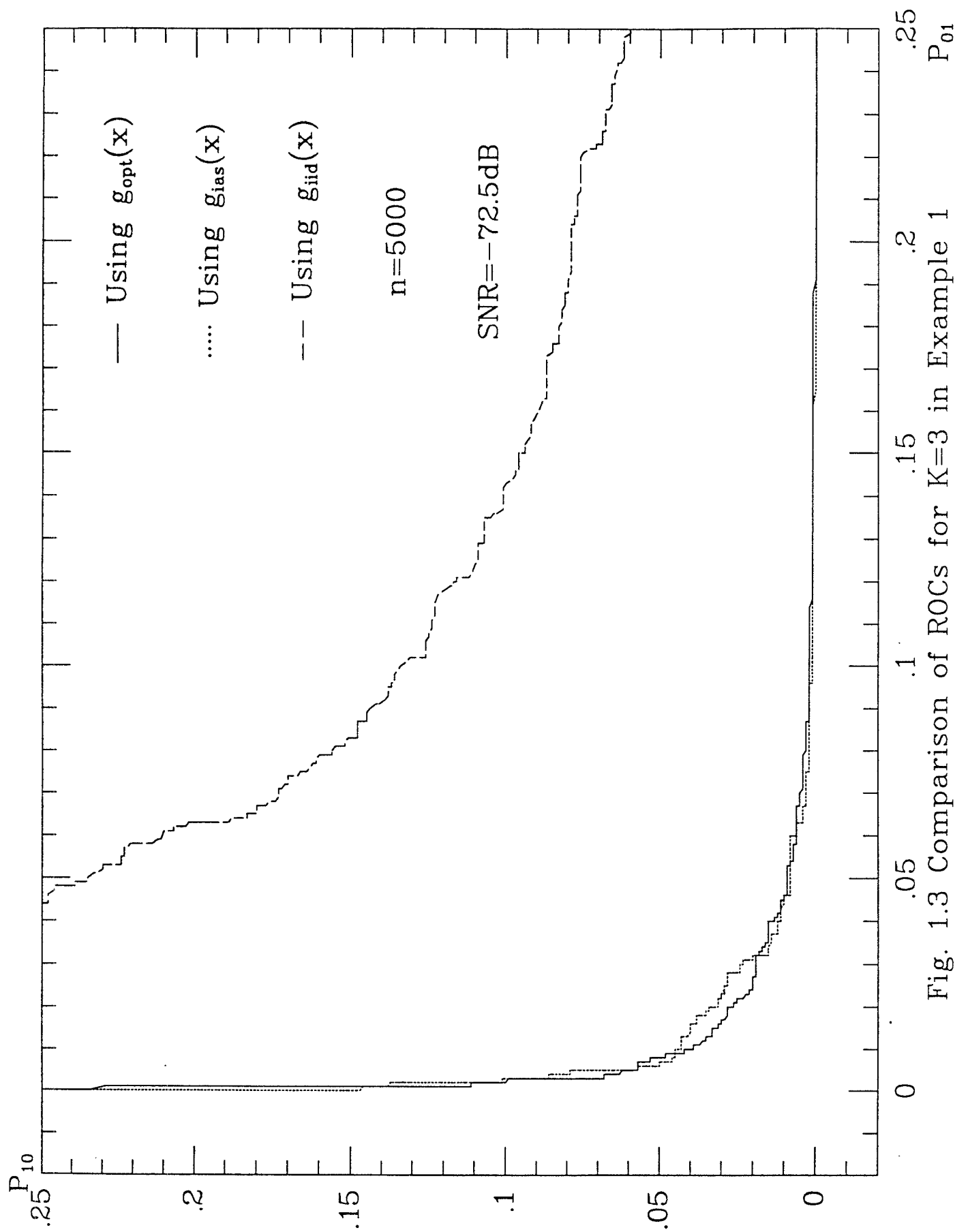


Fig. 1.3 Comparison of ROCs for  $K=3$  in Example 1

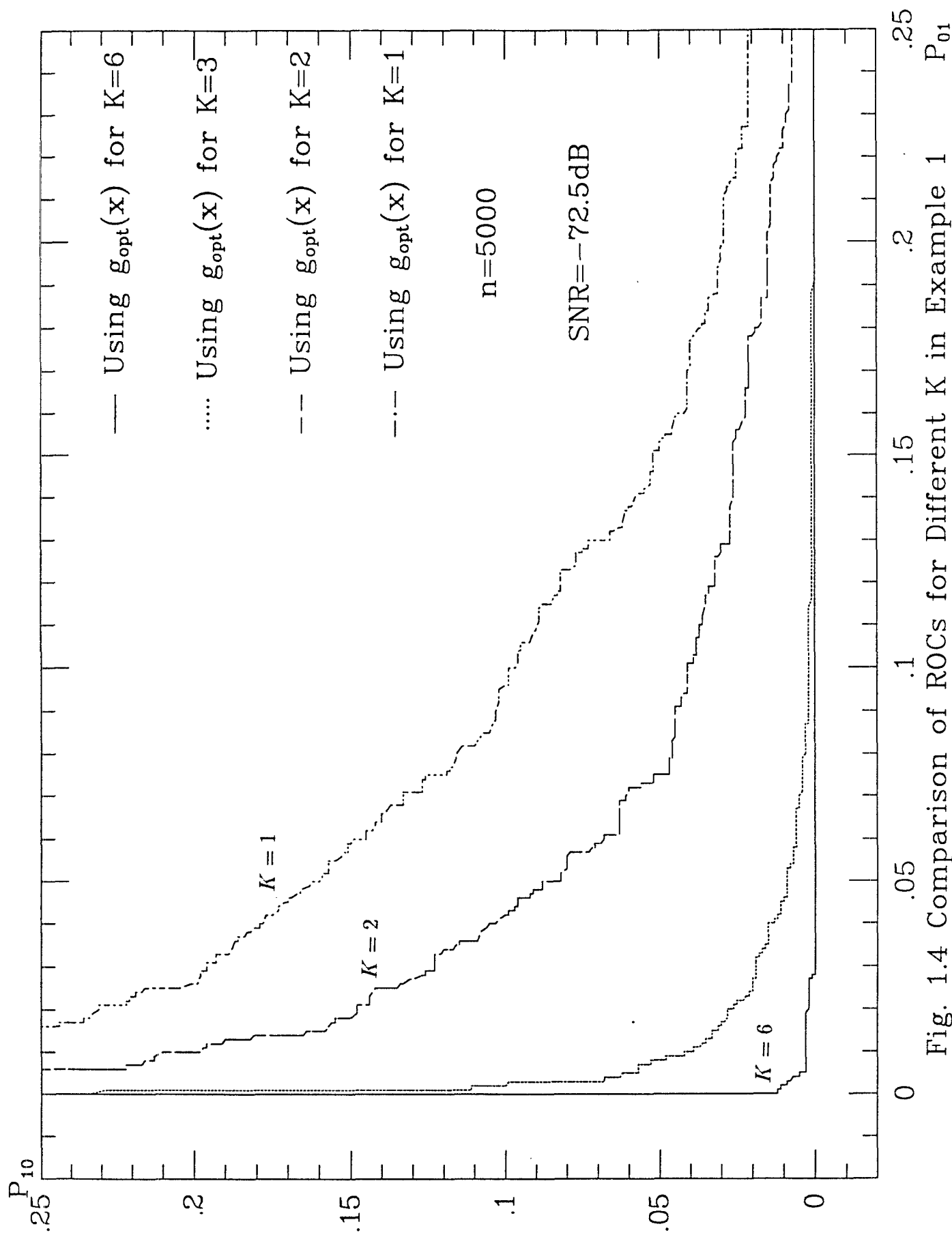


Fig. 1.4 Comparison of ROCs for Different  $K$  in Example 1

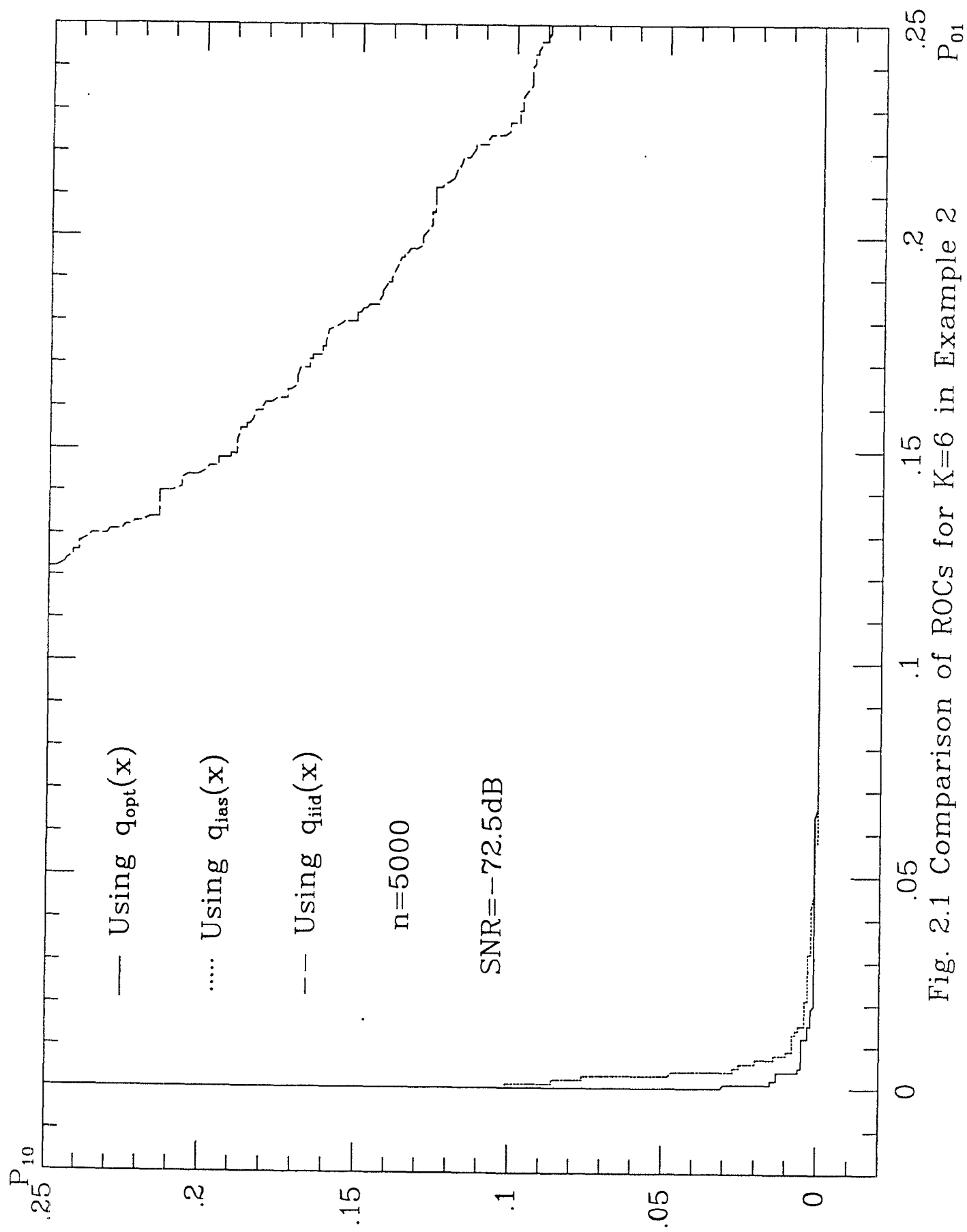


Fig. 2.1 Comparison of ROCs for  $K=6$  in Example 2

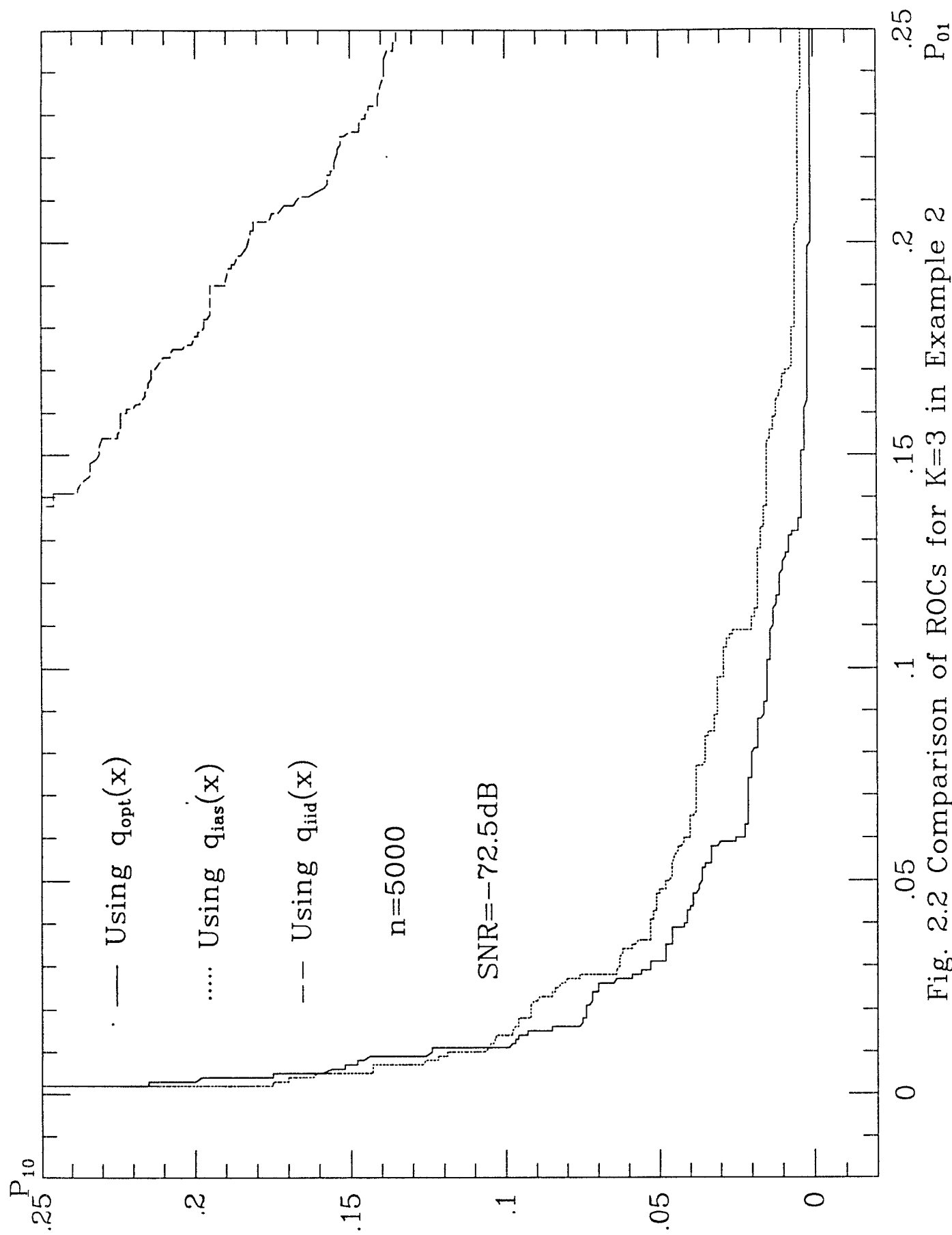


Fig. 2.2 Comparison of ROCs for  $K=3$  in Example 2

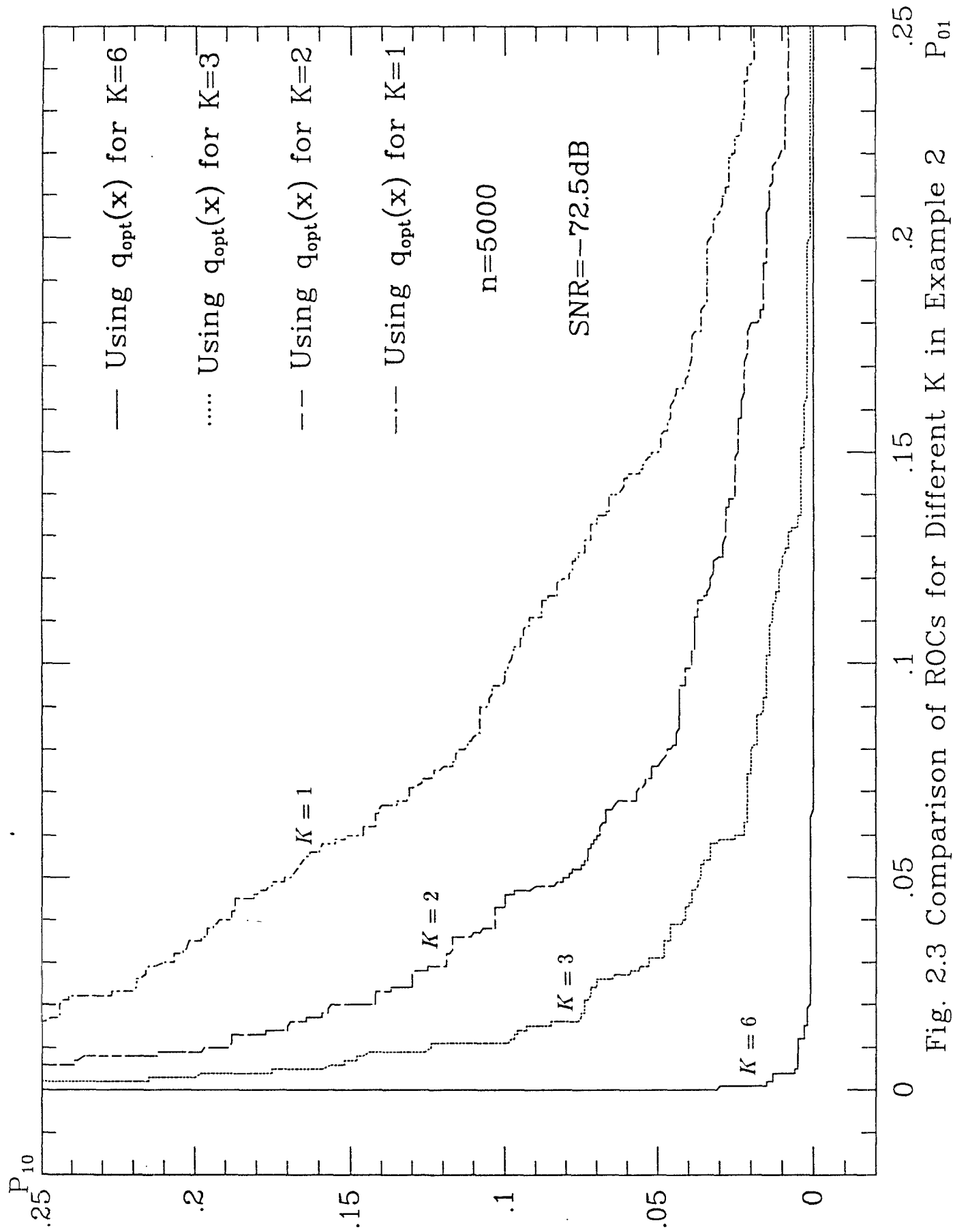


Fig. 2.3 Comparison of ROCs for Different  $K$  in Example 2

$q_{opt}$ for $K = 2$		$q_{opt}$ for $K = 1$	
<i>breakpoints</i>	<i>levels</i>	<i>breckpoints</i>	<i>levels</i>
0.	$0.1955 \times 10^{-1}$	0.	$0.1867 \times 10^{-1}$
0.1261	$0.5956 \times 10^{-1}$	$0.9373 \times 10^{-1}$	$0.5676 \times 10^{-1}$
0.2894	$0.9205 \times 10^{-1}$	0.2013	$0.9461 \times 10^{-1}$
0.4953	$0.9395 \times 10^{-1}$	0.3340	0.1298
0.9514	$0.5378 \times 10^{-1}$	0.6054	0.1196
1.3044	$0.1191 \times 10^{-1}$	1.1727	$0.4398 \times 10^{-1}$
1.7963	$-0.2918 \times 10^{-1}$	1.9575	$-0.4021 \times 10^{-1}$
22.0413	$0.2369 \times 10^{-2}$	34.0785	$-0.4920 \times 10^{-2}$

Table 1: 16-level Optimal Quantizers for  $K = 1, 2$  in Example 2



$q_{opt}$ for $K = 6$		$q_{opt}$ for $K = 3$	
<i>breakpoints</i>	<i>levels</i>	<i>breakpoints</i>	<i>levels</i>
0.	$0.7710 \times 10^{-2}$	0.	$0.1155 \times 10^{-1}$
$0.8027 \times 10^{-1}$	$0.2354 \times 10^{-1}$	$0.8679 \times 10^{-1}$	$0.3446 \times 10^{-1}$
0.1732	$0.3852 \times 10^{-1}$	0.1807	$0.5635 \times 10^{-1}$
0.2741	$0.5062 \times 10^{-1}$	0.2913	$0.7463 \times 10^{-1}$
0.9123	$0.1701 \times 10^{-1}$	1.0027	$0.2490 \times 10^{-1}$
1.3859	$-0.1740 \times 10^{-1}$	1.5836	$-0.2624 \times 10^{-1}$
5.8829	$-0.2725 \times 10^{-2}$	4.9459	$-0.2325 \times 10^{-1}$
30.0644	$-0.1529 \times 10^{-3}$	10.3475	$-0.2286 \times 10^{-2}$

Table 2: 16-level Optimal Quantizers for  $K = 3, 6$  in Example 2

<i>q<sub>iid</sub></i>	
<i>breakpoints</i>	<i>levels</i>
0.	0.1227
0.1240	0.3523
0.2421	0.5680
0.3862	0.7722
0.5703	0.9393
2.0513	0.6281
4.2297	0.2594
31.2372	$3.1992 \times 10^{-2}$

Table 3: 16-level Quantizer under the i.i.d. Assumption in Example 2

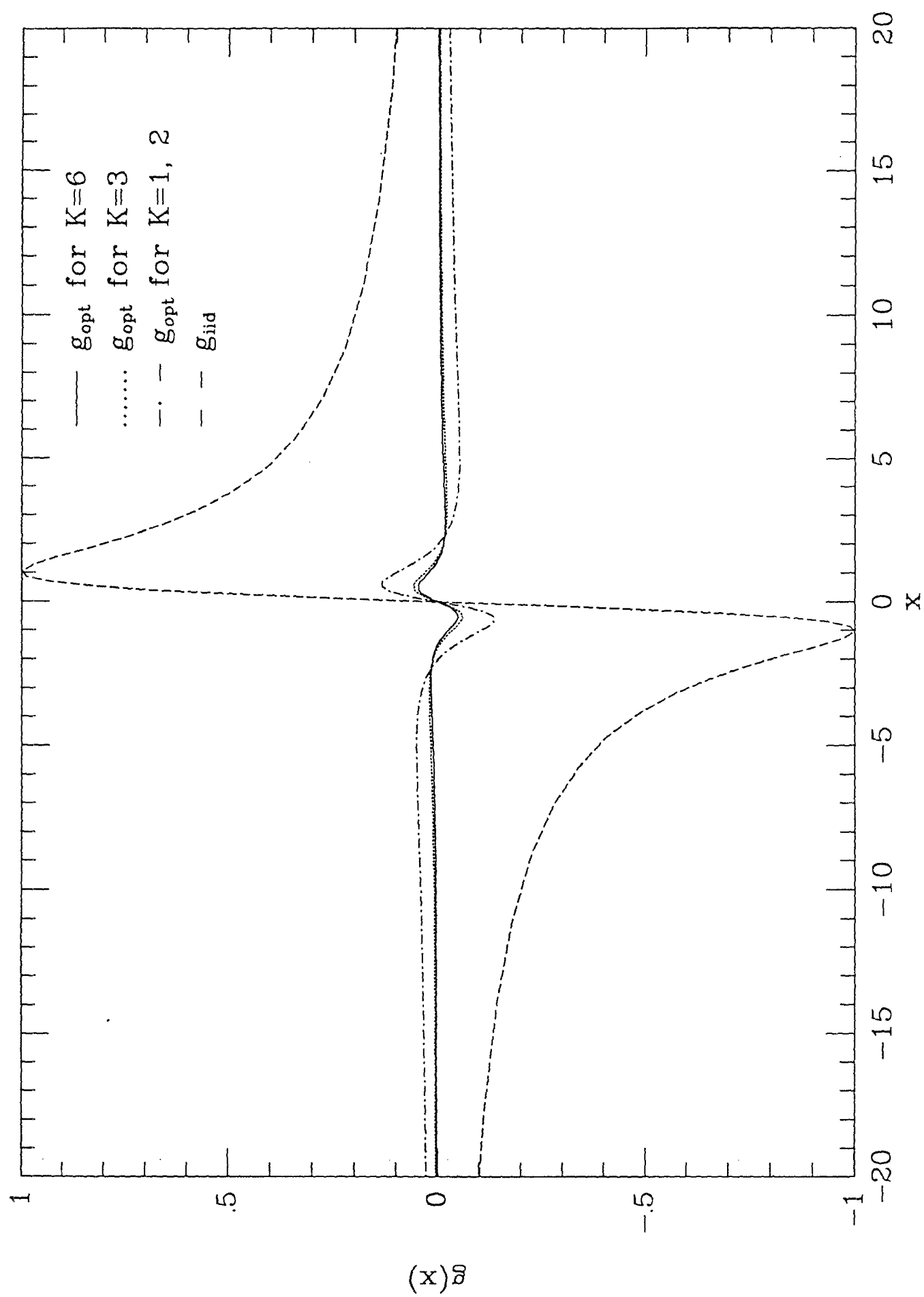


Fig. 3.1 Optimal Nonlinearities  $g_{\text{opt}}$  vs.  $g_{\text{iid}}$  in Scheme 3

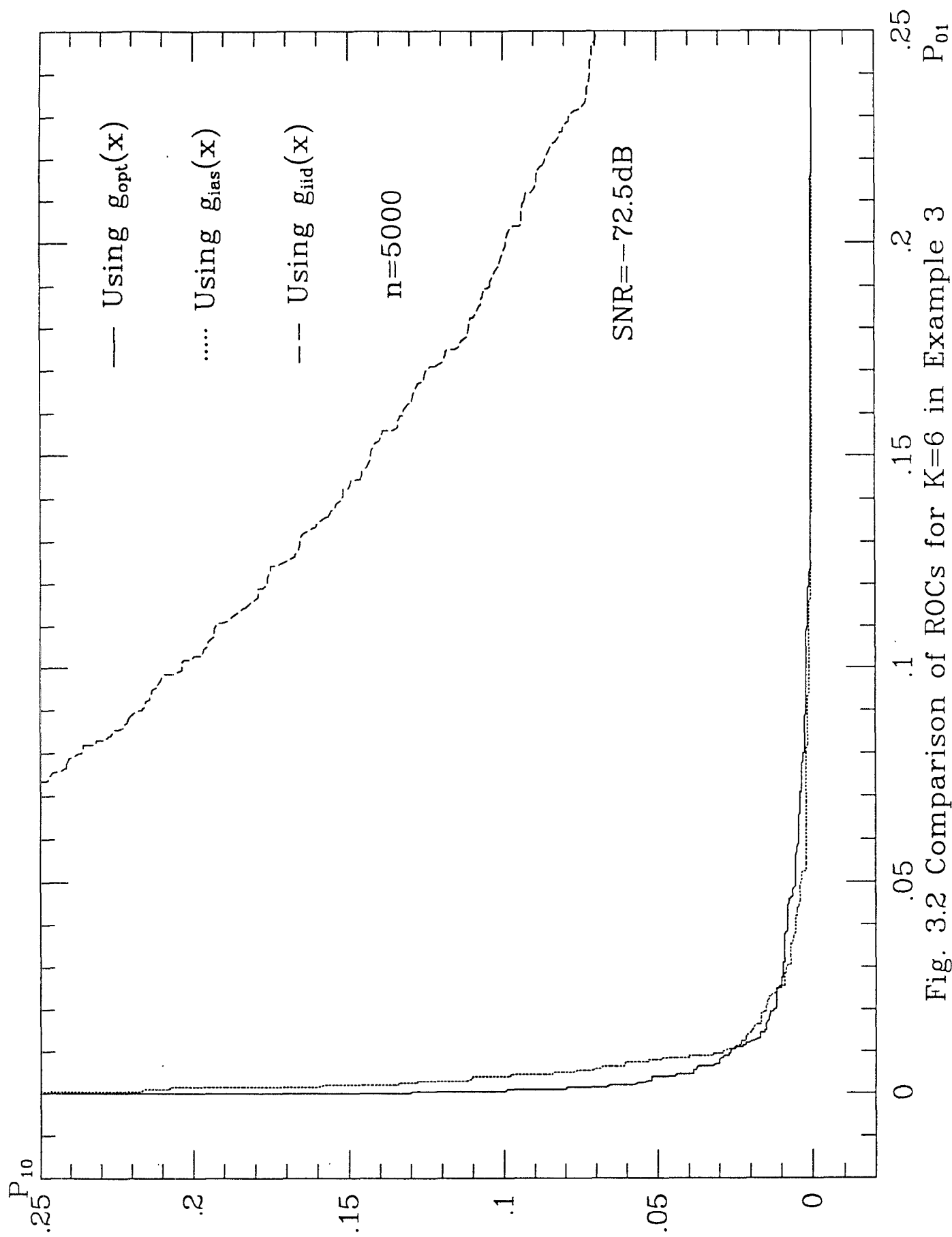


Fig. 3.2 Comparison of ROCs for  $K=6$  in Example 3

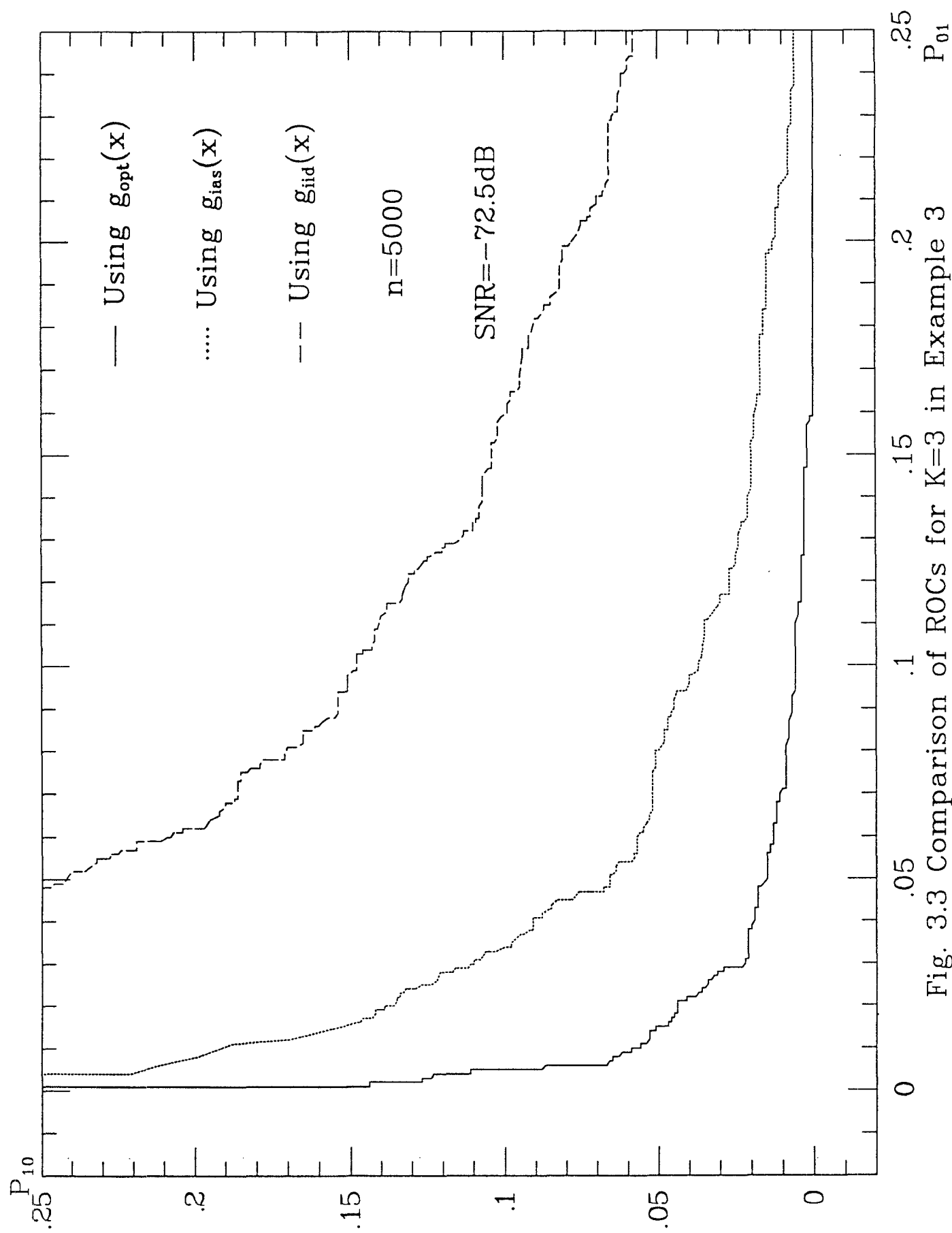


Fig. 3.3 Comparison of ROCs for  $K=3$  in Example 3

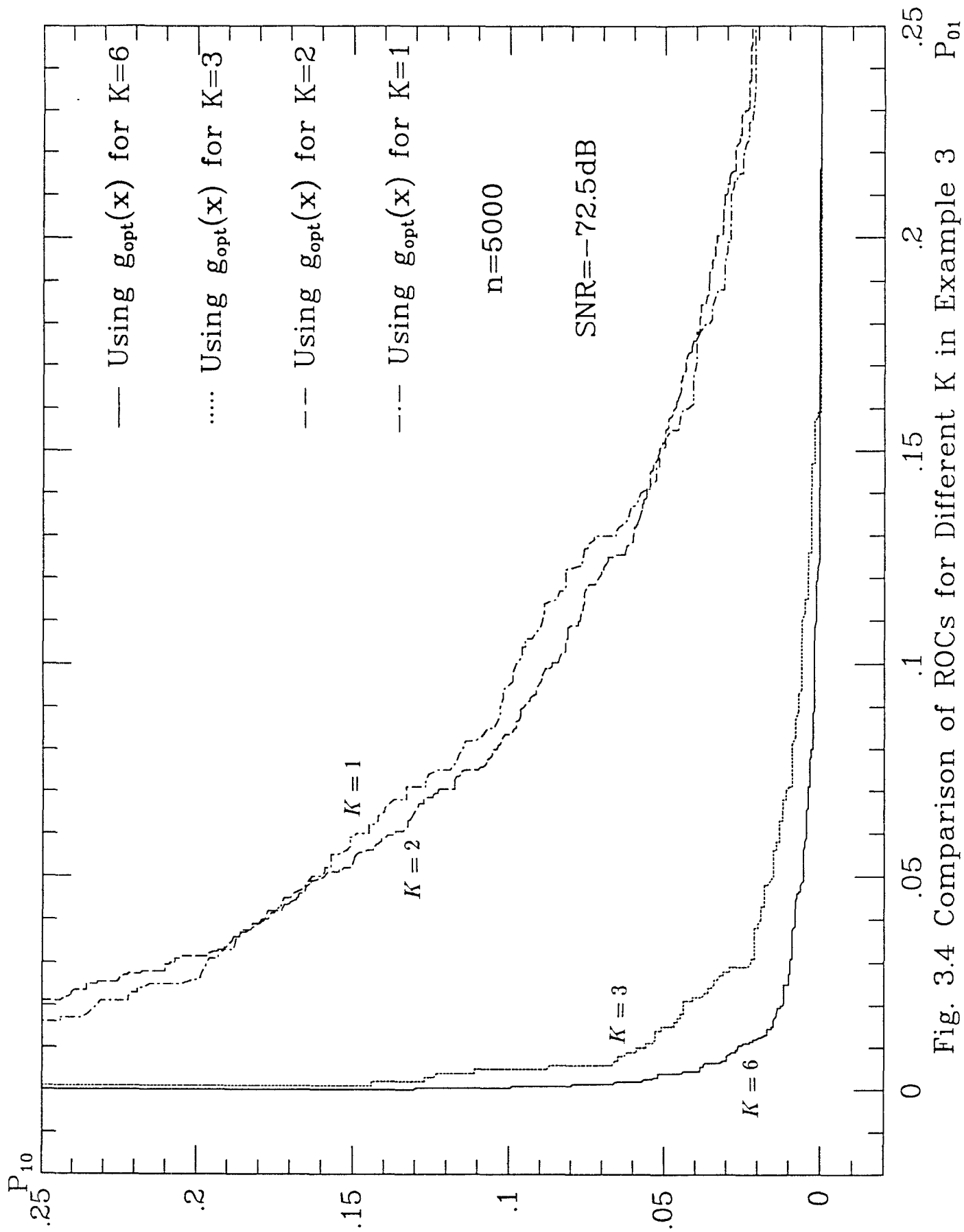


Fig. 3.4 Comparison of ROCs for Different  $K$  in Example 3

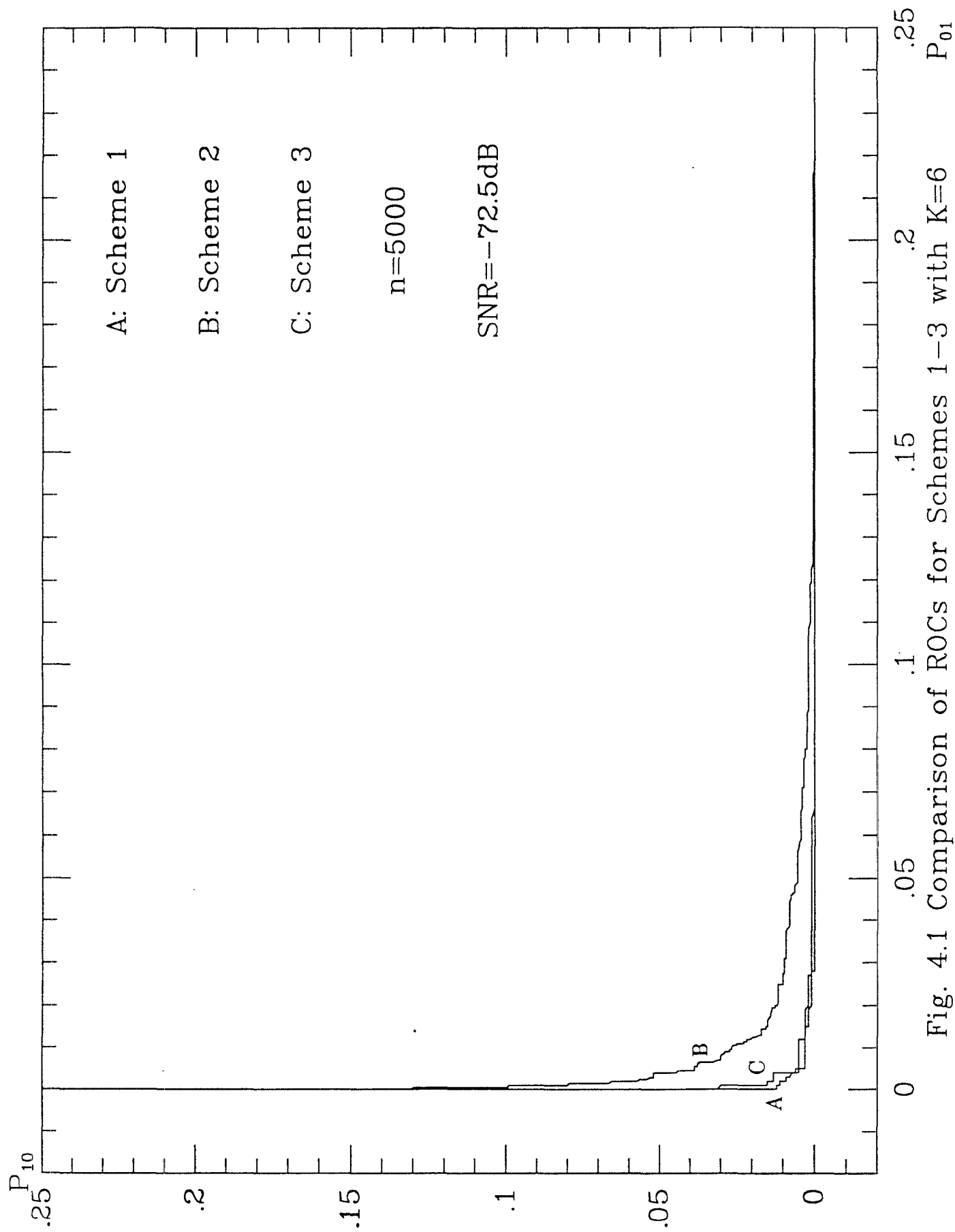


Fig. 4.1 Comparison of ROCs for Schemes 1-3 with K=6

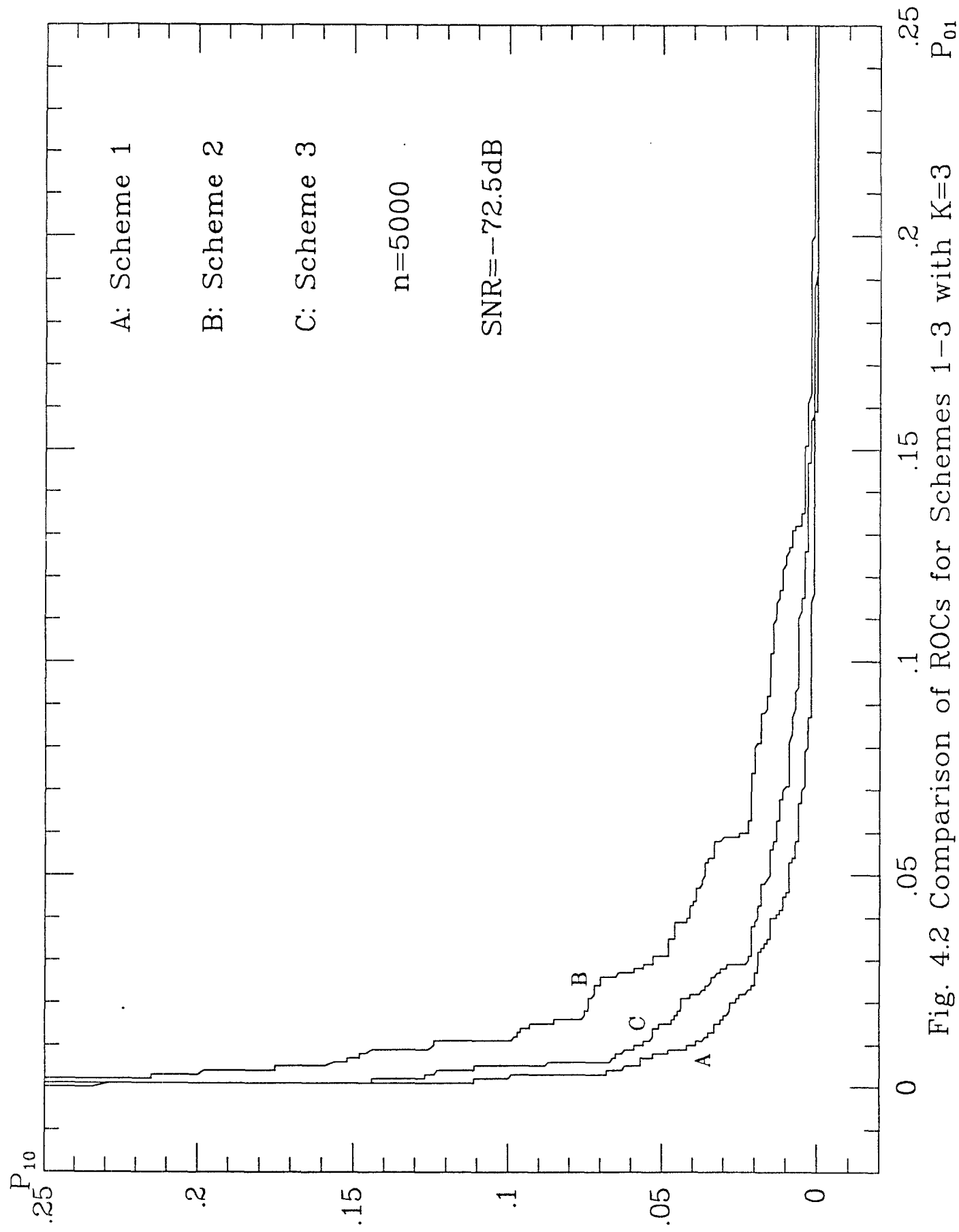


Fig. 4.2 Comparison of ROCs for Schemes 1-3 with K=3