1996

An adaptive algorithm for the design of distributed detection systems

Matthew C. Deans
Lehigh University

Deans, Matthew C.

An Adaptive Algorithm for the Design of Distributed Detection Systems

June 2, 1996

AN ADAPTIVE ALGORITHM FOR THE DESIGN OF DISTRIBUTED DETECTION SYSTEMS

by
Matthew C. Deans

A Thesis
Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Master of Science
in
Electrical Engineering

Lehigh University
May 1996

Approved and recommended for acceptance as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

5/6/96
(Date)

Dr. Rick S. Blum
(Thesis Advisor)

Dr. Alastair McAulay
(Chairman of Department)

Acknowledgments

I would like to thank my advisor Dr. Rick S. Blum for his patience, help and support and for the opportunities he has given me during the past year.

I would also like to thank my family for their confidence in me and their financial, emotional, and moral support which I have needed in varying degrees over the past 6 years.

Finally I would like to thank both the National Science Foundation and the Department of Defense MURI program for providing the funding for this work, and I hope that they will continue to give students opportunities like the one they have given me to pursue their research and educational goals.

Contents

List of Tables

List of Figures

Abstract

Combining signal detection decisions from multiple sensors is useful in some communications, radar, and sonar applications. There has been extensive investigation of optimum schemes for generating and combining the detector decisions for cases with independent observations from sensor to sensor. However, cases with dependent observations from sensor to sensor had received much less attention. Here a simple design approach is outlined. The focus is on the detection of a weak random signal in additive, possibly non-Gaussian noise. The design approach is based on an adaptive algorithm which attempts to learn the distributed detection scheme which provides a minimum mean-square error match to the best centralized detection scheme.

Chapter 1

Introduction

Distributing sensors over a large area is necessary in some signal detection, tracking, and surveillance applications. Such arrangements may provide advantages over single sensor systems in terms of reliability, survivability, and improved signal detection performance. These performance improvements are the result of the inherent spatial diversity combining that occurs in such cases, provided the sensors are separated by sufficient distances. In the interest of reduction of communication costs, simplification of processing, and preventing interception of one's communications, it is often advisable to use distributed detection schemes, which locate quantizers directly at each sensor. As illustrated in Figure 1.1, these quantizers reduce each sensor's observations to a multi-bit decision, and attempt to retain the essential information in these individual decisions needed to make a final signal detection decision.

Specifying the form of the quantizers at the sensors (the sensor decision rules) and specifying how the quantized observations will be used in the final signal detection decision (the fusion rule) are of fundamental importance for obtaining optimum performance. The need to specify each sensor's decision rule and the fusion rule makes distributed detection schemes inherently more complicated to design than the more common centralized detection schemes, where all observations are available in their original form at a central location. In this thesis, the focus is on the case of weak signals, so we consider locally optimum tests [1]. Further, we consider the

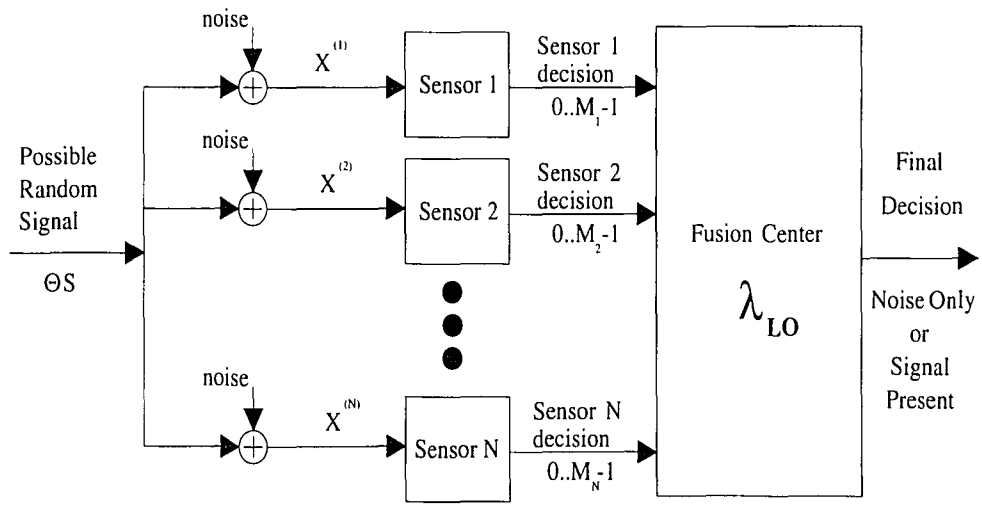


Figure 1.1: Block diagram of a distributed signal detection system.

detection of a common random signal in additive noise.

The results in [2, 3] show that for large observation sample sizes and weak signals [1] the optimum sensor test statistics are composed of some known functions with unknown parameters. The results [2, 3] also show that for large observation sample sizes and weak signals the optimum sensor rules minimize the mean-square error between the test statistic used to make a final decision in the distributed detection system and the test statistic used in the optimum centralized detection system for the same problem. Even for small sample sizes, this mean-square error is a meaningful performance measure. This suggests using an adaptive algorithm to find the parameters that minimize this mean-square error as in Figure 1.2.

Our training scheme implements the stochastic LMS gradient algorithm, a gradient descent algorithm which has been used successfully in adaptive filter and neural network applications [4]. It is important to note that a gradient descent algorithm does not always find the global minimum. However, here we are interested in looking for any local minimum which gives better performance than the best test under the assumption of independent observations, which others have suggested using.

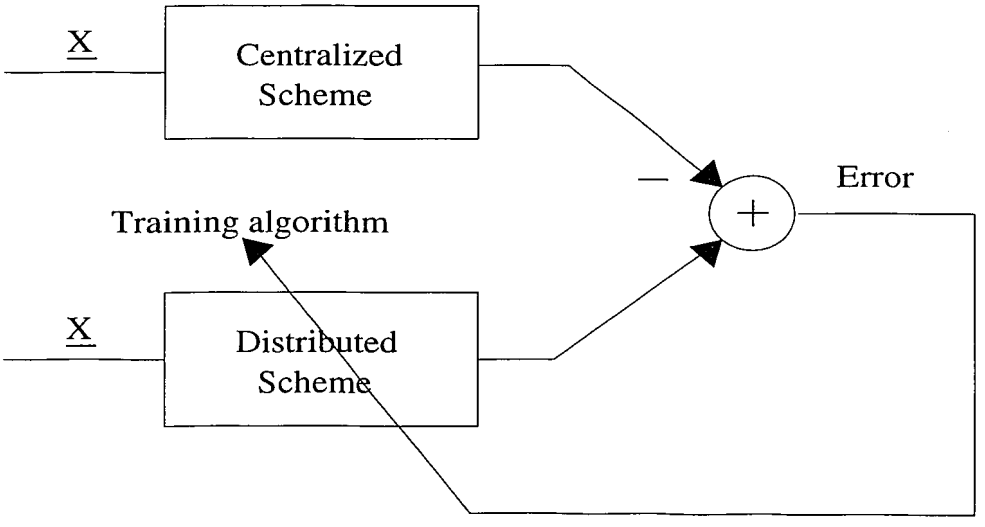


Figure 1.2: Block diagram of the training scheme.

Chapter 2

Distributed Detection Scheme

Consider the detection of a common weak random signal where the observation at the i^{th} discrete-time instant and the p^{th} sensor is given by

$$X_i^{(p)} = \theta S_i + V_i^{(p)} \tag{2.1}$$

where $\{V_i^{(p)}; i = 1, \cdots, n, \quad p = 1, \cdots, N\}$ is a set of independent and identically distributed (iid) zero-mean noise samples with a known common univariate probability density function (pdf) f, θ is a scalar value which represents the signal strength, which is assumed to be small, and $\{S_i; i = 1, \cdots, n\}$ is a set of zero-mean unit-variance random signal samples. For simplicity, we assume the signal samples at different time instants are independent.

In the locally optimum centralized detection scheme, sensor p makes a set of observations, $(X_1^{(p)}, X_2^{(p)}, \cdots, X_n^{(p)}) = (x_1^{(p)}, x_2^{(p)}, \cdots, x_n^{(p)})$. The final test statistic is given by

$$\lambda_{UQ} = \sum_{i=1}^{n}\sum_{p=1}^{N} \frac{f''(x_i^{(p)})}{f(x_i^{(p)})} + \sum_{i=1}^{n}\sum_{p=1}^{N}\sum_{q=1,q\neq p}^{N} \frac{f'(x_i^{(p)})}{f(x_i^{(p)})}\frac{f'(x_i^{(q)})}{f(x_i^{(q)})}. \tag{2.2}$$

where $f'(x)$ and $f''(x)$ are the first and second derivative, respectively, of the noise pdf $f(x)$. The final decision is made by comparing λ_{UQ} to a threshold. The value of the threshold is determined by the required false alarm probability.

It is important to note that if the noise pdf $f(x)$ is an even symmetric function then the second derivative $f''(x)$ is even symmetric and the first derivative $f'(x)$ is

odd symmetric

The first term of (2.2) can be interpreted as a generalized measure of the energy of the observations. The second term can be interpreted as a generalized measure of the correlation between observations made at different sensors. In a situation where the observations from sensor to sensor are uncorrelated, the correlation measure becomes unimportant and the second term in (2.2) can be ignored. However, ignoring the correlation term in situations where observations are dependent from sensor to sensor may degrade performance.

In the distributed detection scheme, sensor p makes a set of observations, $(X_1^{(p)}, X_2^{(p)}, \cdots, X_n^{(p)}) = (x_1^{(p)}, x_2^{(p)}, \cdots, x_n^{(p)})$. If the observation vector falls within a quantization region $A_j^{(p)}$, the quantizer at sensor p produces the multi-bit decision symbol $j \in \{0, \cdots, M_p - 1\}$. The fusion center makes a final decision based on the complete set of sensor decisions by calculating the locally optimum test statistic [3]

$$\lambda_{LO} = \sum_{p=1}^{N} \sum_{j=0}^{M_p-1} l_j^{(p)} z_j^{(p)} + \sum_{p=1}^{N} \sum_{j=0}^{M_p-1} \sum_{q=1, q\neq p}^{N} \sum_{k=0}^{M_p-1} \tilde{l}_j^{(p)} z_j^{(p)} \tilde{l}_k^{(q)} z_k^{(q)} \tag{2.3}$$

where

$$l_j^{(p)} = \frac{\int_{x \in A_j^p} \frac{f''(x)}{f(x)} f(x)\, dx}{\int_{x \in A_j^p} f(x) dx} \,. \tag{2.4}$$

and

$$\tilde{l}_j^{(p)} = \frac{\int_{x \in A_j^p} \frac{f'(x)}{f(x)} f(x)\, dx}{\int_{x \in A_j^p} f(x)\, dx} \tag{2.5}$$

are conditional expectation values and $z_j^{(p)}$ denotes an indicator function such that

$$z_j^{(p)} = \begin{cases} 1, & \text{if } x_1^{(p)} \in A_j^p \\ 0, & \text{otherwise} \end{cases} \,. \tag{2.6}$$

The final detection decision is made by comparing λ_{LO} to a threshold. The value of the threshold is determined by the required false alarm probability.

The first summation in (2.3) is analogous to the first summation in (2.2) and represents the energy of the observations. Likewise, the second summation in (2.3) is analogous to the second summation in (2.2) and represents the correlation between observations at different sensors.

Consider the mean-square error between the locally optimum quantized test statistic and the locally optimum centralized test statistic as given by

$$J = E\left[(\lambda_{UQ} - \lambda_{LO})^2 \,|\theta = 0\right] \tag{2.7}$$

Under the assumptions outlined, the sensor processing schemes which minimize J are found in [3]. The regions $A_j^{(p)}$ are given by the intersection of $M_p - 1$ regions, each of which comes from comparing a test statistic

$$\hat{\lambda}_{j,k}^{(p)}(\underline{x}^{(p)}) = \sum_{i=1}^{n} a_{i,j,k}^{(p)} \frac{f''(x_i^{(p)})}{f(x_i^{(p)})} + \sum_{i=1}^{n} b_{i,j,k}^{(p)} \frac{f'(x_i^{(p)})}{f(x_i^{(p)})} \tag{2.8}$$

to a threshold $\hat{t}_{j,k}^{(p)}$. The scalar parameters $a_{i,j,k}^{(p)}$, $b_{i,j,k}^{(p)}$, and thresholds $\hat{t}_{j,k}^{(p)}$ are antisymmetric with respect to j and k. Because of this antisymmetric property, there are a total of $(2n+1)M_p(M_p-1)/2$ unknowns. Region $A_j^{(p)}$ is given by

$$A_j^{(p)} = \bigcap_{k=0,k\neq j}^{M_p-1} \left\{\underline{x} : \hat{\lambda}_{j,k}^{(p)}(\underline{x}) \geq \hat{t}_{j,k}^{(p)}\right\} \tag{2.9}$$

To simplify matters we consider the case where N individual sensor quantizers each make a single observation. This is the case where $n = 1$ in (2.1) and the complete set of sensor observations is $\{x_1^{(p)}; p = 1, \ldots, N\}$. The test statistics in (2.8) then take the form

$$\hat{\lambda}_{j,k}^{(p)}(x) = a_{1,j,k}^{(p)} \frac{f''(x)}{f(x)} + b_{1,j,k}^{(p)} \frac{f'(x)}{f(x)} \tag{2.10}$$

Normalizing by $||(a_{1,j,k}^{(p)}, b_{1,j,k}^{(p)})||$ and transforming to polar coordinates yields

$$\lambda_{j,k}^{(p)}(x) = \cos(\phi_{j,k}^{(p)}) \frac{f''(x)}{f(x)} + \sin(\phi_{j,k}^{(p)}) \frac{f'(x)}{f(x)} \tag{2.11}$$

where

$$\phi_{j,k}^{(p)} = \arctan\left(\frac{b_{1,j,k}^{(p)}}{a_{1,j,k}^{(p)}}\right). \tag{2.12}$$

The threshold is normalized as

$$t_{j,k}^{(p)} = \frac{\hat{t}_{j,k}^{(p)}}{||(a_{1,j,k}^{(p)}, b_{1,j,k}^{(p)})||}. \tag{2.13}$$

This transformation results in fewer unknown parameters without a loss of generality, since any region definition which can be obtained by comparing (2.10) to a threshold can be obtained by comparing (2.11) to a threshold.

As an example, consider a binary quantizer ($M_p = 2$) at sensor p. There is one comparison which defines each region $A_j^{(p)}$. The regions are defined as

$$A_1^{(p)} = \left\{x : \lambda_{1,0}^{(p)}(x) > t_{1,0}^{(p)}\right\}$$

$$A_0^{(p)} = \overline{A_1^{(p)}}. \tag{2.14}$$

For a ternary quantizer ($M_p = 3$) at sensor p, there are two comparisons which define each region $A_j^{(p)}$. The regions are defined as

$$A_2^{(p)} = \left\{x : \lambda_{2,1}^{(p)}(x) \geq t_{2,1}^{(p)}\right\} \cap \left\{x : \lambda_{2,0}^{(p)}(x) \geq t_{2,0}^{(p)}\right\}$$
$$A_1^{(p)} = \left\{x : \lambda_{2,1}^{(p)}(x) < t_{2,1}^{(p)}\right\} \cap \left\{x : \lambda_{1,0}^{(p)}(x) \geq t_{1,0}^{(p)}\right\}$$

$$A_0^{(p)} = \overline{A_2^{(p)} \cup A_1^{(p)}}. \tag{2.15}$$

Chapter 3

Training Algorithm

The mean-square error J in (2.7) depends on the parameters of the sensor quantizers which are collectively denoted by the system parameter vector $\vec{W}$. For a set of binary quantizers, the system parameter vector would be

$$\vec{W} = \left[t_{1,0}^{(1)}, \phi_{1,0}^{(1)}, t_{1,0}^{(2)}, \phi_{1,0}^{(2)}, \cdots, t_{1,0}^{(N)}, \phi_{1,0}^{(N)}, \right]^T . \tag{3.1}$$

In the gradient descent method [4], the system parameter vector is updated by adjusting it in the direction of steepest descent of the cost function, which is opposite the gradient $\vec{\nabla}_W J$ as in

$$\vec{W}(m+1) = \vec{W}(m) - \mu \vec{\nabla}_W J, \tag{3.2}$$

where μ is a small positive constant called the *learning rate* and m is an index which denotes the iteration number in the training. The updated vector $\vec{W}(m+1)$ will be closer than $\vec{W}(m)$ to the minimum of J provided [4] J is well behaved and has a minimum, μ is sufficiently small, and the initial condition $W(0)$ is in the domain of attraction of the minimum of J. Since λ_{UQ} in (2.2) does not depend on the parameters of the distributed scheme, $\vec{\nabla}_W \lambda_{UQ}$ is zero so that $\vec{\nabla}_W J$ can be written as

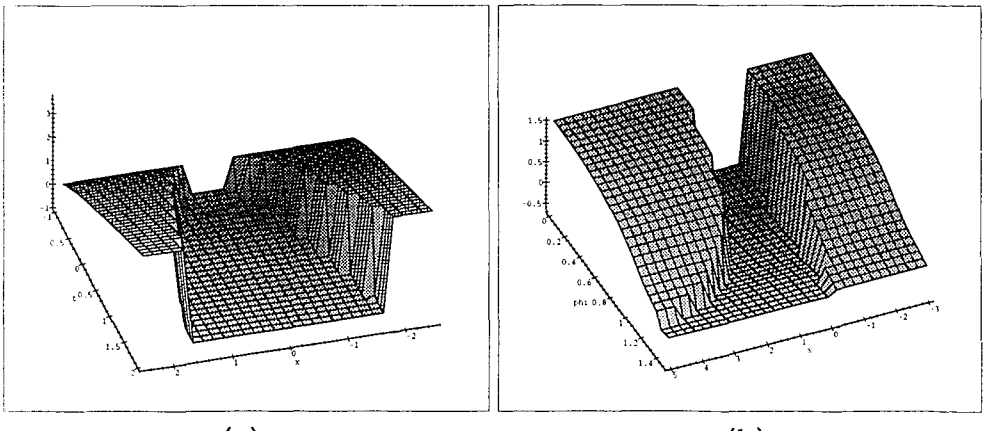
$$\vec{\nabla}_W J = -2E \left[(\lambda_{UQ} - \lambda_{LO}) \vec{\nabla}_W \lambda_{LO} | \theta = 0 \right] . \tag{3.3}$$

For simplicity, the expectation value in the gradient will be approximated during training (with $\theta = 0$), as in

$$\vec{\nabla}_W J \approx -2(\lambda_{UQ} - \lambda_{LO})\vec{\nabla}_W \lambda_{LO}, \tag{3.4}$$

which is simply the instantaneous value. This approximation reduces the training to the stochastic LMS gradient algorithm which has been previously studied [4]. At each iteration, $\vec{W}$ is updated as

$$\vec{W}(m+1) = \vec{W}(m) + 2\mu(\lambda_{UQ} - \lambda_{LO})\vec{\nabla}_W \lambda_{LO}. \tag{3.5}$$



(a) (b)

Figure 3.1: λ_{LO} for a one sensor system.

Consider a single sensor with a binary quantizer. The observation $x_1^{(1)}$ is quantized, causing the test statistic λ_{LO} to be discontinuous in $x_1^{(1)}$. However, the quantization regions are determined by $\vec{W} = \left[t_{1,0}^{(1)}, \phi_{1,0}^{(1)}\right]^T$. This means that for some $x_1^{(1)}$ and $t_{1,0}^{(1)}$, the test statistic is discontinuous in $\phi_{1,0}^{(1)}$, and for some $x_1^{(1)}$ and $\phi_{1,0}^{(1)}$, the test statistic is discontinuous in $t_{1,0}^{(1)}$. The discontinuity present in λ_{LO} causes the gradient to be undefined for some parameter settings and observation values. In Figure 3.1(a) the test statistic λ_{LO} is shown as a function of $x_1^{(1)}$, labeled as 'x', and $t_{1,0}^{(1)}$, labeled as 't', with $\phi_{1,0}^{(1)} = 0$. In Figure 3.1(b) the test statistic λ_{LO} is shown as a function of $x_1^{(1)}$, labeled as 'x', and $\phi_{1,0}^{(1)}$, labeled as 'phi', with $t_{1,0}^{(1)} = 0$. We have introduced a sigmoid function which smoothes out the discontinuities in λ_{LO}

by providing a continuous and differentiable approximation to $\sum_{j=0}^{M_p-1} l_j^{(p)} z_j^{(p)}$ in (2.3) as

$$L^{(p)}(x) = l_0^{(p)} + \sum_{j=1}^{M_p-1} \frac{l_j^{(p)} - l_0^{(p)}}{1 + \sum_{k=0,k\neq j}^{M_p-1} exp\left(-\gamma(\lambda_{j,k}(x) - t_{j,k})\right)} \tag{3.6}$$

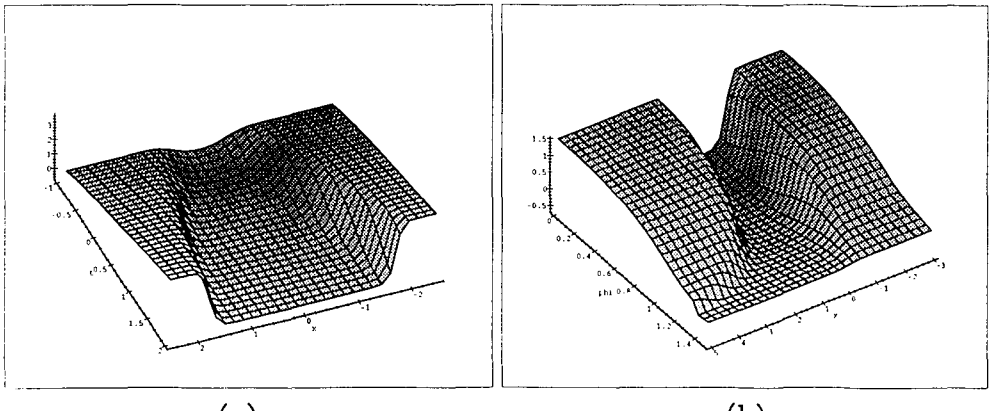
In a similar way we have defined a continuous, differentiable approximation to $\sum_{j=0}^{M_p-1} \tilde{l}_j^{(p)} z_j^{(p)}$ in (2.3) as

$$\tilde{L}^{(p)}(x) = \tilde{l}_0^{(p)} + \sum_{j=0}^{M_p-1} \frac{\tilde{l}_j^{(p)} - \tilde{l}_0^{(p)}}{1 + \sum_{k=0,k\neq j}^{M_p-1} exp\left(-\gamma(\lambda_{j,k}(x) - t_{j,k})\right)} \tag{3.7}$$

The scaling parameter γ controls the size of the region over which the discontinuity is smoothed. The test statistic is then formed at the fusion center using the approximate values $L^{(p)}(x_1^{(p)})$ and $\tilde{L}^{(p)}(x_1^{(p)})$ as

$$\lambda_{LO} \approx \sum_{p=1}^{N} L^{(p)}(x_1^{(p)}) + \sum_{p=1}^{N} \sum_{q=1,q\neq p}^{N} \tilde{L}^{(p)}(x_1^{(p)}) \tilde{L}^{(q)}(x_1^{(q)}). \tag{3.8}$$

The resulting test statistic is a smooth function of $x_1^{(1)}$ and $\vec{W}$ as shown in Figure 3.2. In Figure 3.2(a) the test statistic λ_{LO} is shown as a function of $x_1^{(1)}$ and $t_{1,0}^{(1)}$, labeled as 'x' and 't', with $\phi_{1,0}^{(1)} = 0$. In Figure 3.2(b) the test statistic λ_{LO} is shown as a function of $x_1^{(1)}$ and $\phi_{1,0}^{(1)}$, labeled as 'x' and 'phi', with the parameter $t_{1,0}^{(1)} = 0$.



(a) (b)

Figure 3.2: λ_{LO} for a one sensor system after smoothing with sigmoids.

Chapter 4

Implementation

The training algorithm of Chapter 3 has been implemented in IBM XL Pascal on an IBM RS/6000 workstation. Two-sensor cases with binary and ternary quantizers were studied. Binary quantization was studied for observations with noise pdfs of the form

$$f(x) = \frac{k}{2\mathcal{A}(k)\Gamma(1/k)} \exp -[|x|/\mathcal{A}(k)]^k \tag{4.1}$$

where

$$\mathcal{A}(k) = \left(\frac{\Gamma(1/k)}{\Gamma(3/k)}\right)^{\frac{1}{2}} \tag{4.2}$$

and where $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$ is the gamma function. f is known as a generalized Gaussian pdf, which has zero mean and unit variance. If $k = 2.0$ then f is the Gaussian pdf and if $k < 2.0$ then f is a pdf with heavy tails. The heavy-tailed pdfs appear to be reasonable models for a number of practical cases [1] including impulsive noise. In this thesis, parameter values of $k = 2.0, 1.8$, and 1.6 are presented. Binary quantization is also studied for observations with Cauchy noise, with a pdf

$$f(x) = \frac{1}{\pi(x^2+1)}. \tag{4.3}$$

The Cauchy pdf is also considered a heavy-tailed pdf. Cases with two ternary quantizers and generalized Gaussian noise with parameter $k = 2.0$ were also studied.

The observation samples are generated using subroutines available in the IMSL Fortran Numerical Libraries. For generalized Gaussian noise, the IMSL routine

DRNGCT is used, which generates random samples given a table of values for a monotone increasing continuous cumulative distribution function (cdf). The routine implements the inverse cdf method using piecewise cubic interpolation to generate the samples. For generation of Cauchy noise, the routine DRNCHY is used, which uses a faster and more simple method than inverse cdf to generate Cauchy noise samples.

To calculate the conditional expectation values $l_j^{(p)}$ and $\tilde{l}_j^{(p)}$, the integrals in equations (2.4) and (2.5) are to be calculated. This requires that the boundaries of the quantization regions $A_j^{(p)}$ be determined. The points at which the test statistics $\lambda_{j,k}^{(p)}(x)$ are equal to the thresholds $t_{j,k}^{(p)}$ are the boundaries of the quantization regions. The functions $\lambda_{j,k}^{(p)}(x) - t_{j,k}^{(p)}$ are scanned from left to right at non-uniformly spaced intervals starting at a large negative value and ending at a large positive value such that the starting and ending points have absolute values large enough that the value of the pdf at the endpoints $f(\text{boundary})$ is negligible so that if a zero crossing exists outside of the range scanned, its existence is not important numerically. The intervals are cubicly spaced so that the intervals are large for large values of x where the function $f(x)$ changes slowly, and small for x near 0 where $f'(x)$ and $f''(x)$ can take large values. Once a zero crossing is found in an interval, the exact location of the zero crossing is found to arbitrary accuracy using the bisection method. After all zero crossings are found, every interval between two zero crossings is tested using (2.9) to determine which region $A_j^{(p)}$ the interval belongs in. It should be pointed out that this method does not necessarily perform well for finding all roots of an arbitrary function. With the class of functions considered here, it is easy to bound the number of possible zero crossings and make some general statements concerning their locations.

Once the boundaries of the quantization regions are found, the integrals in equations (2.4) and (2.5) can be calculated. Since an analytic expression for $f(x)$ and $f'(x)$ can be written, the integrals in the numerators of equations (2.4) and (2.5) can be evaluated analytically. For the denominators, a general purpose integration routine available in the IMSL Fortran Numerical Library is used.

The calculation of λ_{LO} follows from the calculation of the conditional expectation

values $l_j^{(p)}$ and $\tilde{l}_j^{(p)}$ and equations (3.6) and (3.7).

The calculation of $\vec{\nabla}_W \lambda_{LO}$ is accomplished by a finite difference calculation of the derivative of λ_{LO} with respect to each of the system parameters in $\vec{W}$.

The training algorithm is summarized in the steps given below.

1. Generate $x_1^{(p)}$ for each sensor.
2. Calculate λ_{UQ} from equation (2.2) and λ_{LO} from equation (3.8).
3. Find the error $\lambda_{UQ} - \lambda_{LO}$.
4. Calculate the gradient $\vec{\nabla}_W \lambda_{LO}$ from (3.8)
5. Update system parameters according to (3.5).
6. Repeat.

Chapter 5

Results

We consider a distributed signal detection scheme as shown in Figure 1.1 which is designed to detect a weak random signal in noise. For simplicity we consider the case of only two sensors and assume the observations come from the model given in (2.1) with noise samples with a generalized Gaussian or a Cauchy pdf.

The parameters in the sensor decision rules are learned using the algorithm outlined in Chapter 3 with $\mu = 0.00001$ in (3.5) and the scaling parameter $\gamma = 40$ in (3.6) and (3.7).

A system with two binary sensors was trained for the case of generalized Gaussian noise with $k = 2.0$. The initial conditions were set up so that the quantization rules were non-symmetric and different at each sensor. The initial and final values of the system parameters are shown in Table 5.1. Other initial conditions were tried and similar results were obtained. The time evolution of the parameters is shown in Figure 5.1. Notice that the angles $\phi_{1,0}^{(1)}$ and $\phi_{1,0}^{(2)}$ both converge to values near zero, indicating that each sensor uses an even symmetric quantization rule, i.e. if $x \in A_j^{(p)}$ then $-x \in A_j^{(p)}$. An even symmetric region $A_j^{(p)}$ also implies that $\tilde{l}_j^{(p)} = 0$ for that region, since $\tilde{l}_j^{(p)}$ is calculated using the integral of an even symmetric function. The result is the same quantization rule as would be obtained under the assumption of independent observations, where the correlation between observations at different sensors is unimportant. The final parameters in Table 5.1 describe a system similar to one which was found to be optimum in previous work [2].

Parameter	Initial	Final
$t_{1,0}^{(1)}$	0.0	1.19
$\phi_{1,0}^{(1)}$	$\pi/4$	0.0
$t_{1,0}^{(2)}$	0.0	1.18
$\phi_{1,0}^{(2)}$	$-\pi/4$	0.0

Table 5.1: Optimum system parameters for two binary sensors and k=2.0.

Parameter	Initial	Final
$t_{1,0}^{(1)}$	0.0	0.76
$\phi_{1,0}^{(1)}$	1.0	0.0
$t_{1,0}^{(2)}$	0.0	0.76
$\phi_{1,0}^{(2)}$	-1.0	0.0

Table 5.2: Optimum system parameters for two binary sensors and k=1.8.

A system of two binary sensors was trained for the case of generalized Gaussian noise with $k = 1.8$. The system parameters were intentionally set to non-symmetric initial conditions as in the previous example. The initial and final values are shown in Table 5.2. Other initial conditions were tried for this case and similar results were obtained. The time evolution of these parameters is shown in Figure 5.2. Notice again that the final results indicate that the quantization rules used at each sensor are even symmetric. The final parameters shown in Table 5.2 describe a system similar to one which was found to be optimum in previous work [2].

A system of two binary sensors was trained under generalized Gaussian noise with $k = 1.6$. The system parameters were intentionally set to initially represent an even symmetric quantization rule. The initial and final values are shown in Table 5.3. Again, other initial conditions were tried for this case and similar results were obtained. The time evolution of these parameters is shown in Figure 5.3. The initial angles $\phi_{1,0}^{(1)}$ and $\phi_{1,0}^{(2)}$ were set to 0 but converged to nonzero final values, indicating that even symmetric quantization rules are not best for this particular noise pdf. The final parameters shown in Table 5.3 describe a system similar to one which was found to be optimum in previous work [2].

Parameter	Initial	Final
$t_{1,0}^{(1)}$	-1.0	0.03
$\phi_{1,0}^{(1)}$	0.0	-1.3
$t_{1,0}^{(2)}$	-1.0	0.03
$\phi_{1,0}^{(2)}$	0.0	1.3

Table 5.3: Optimum system parameters for two binary sensors and k=1.6.

Parameter	Initial	Final
$t_{1,0}^{(1)}$	0.0	0.2
$\phi_{1,0}^{(1)}$	0.0	-1.57
$t_{1,0}^{(2)}$	0.0	0.2
$\phi_{1,0}^{(2)}$	0.0	1.57

Table 5.4: Optimum system parameters for two binary sensors and Cauchy noise

For Cauchy noise, the initial and final parameter values are shown in Table 5.4. Again, other initial conditions were tried for this case and similar results were obtained. The time evolution of these parameters is shown in Figure 5.4. In the case of Cauchy noise, unlike any of the previous cases, the angles $\phi_{1,0}^{(1)}$ and $\phi_{1,0}^{(2)}$ converged to values of $\pi/2$ and $-\pi/2$, respectively. This implies that the energy term is unimportant and that the test statistic λ_{LO} depends heavily on correlation terms. Since this result has not been previously reported, it was compared to the optimum system under the assumption of independent observations. Under the assumption of independent observations, the optimum system has a mean square error of 6.56. The result presented here has a mean-square error of 2.31.

A system of two three-level sensors was also trained under generalized Gaussian noise with parameter k=2.0. The sensor decisions are made using (2.15) and the fusion center generates a final decision using the test statistic in (3.8). Given two different sets of initial conditions, very different results were obtained. One set of initial and final values of the system parameters are shown in Table 5.5. Another appears in Table 5.6. The values which appear in Table 5.5 describe a system with two identical quantizers, each of which has symmetric quantization regions

	Initial	Final		Initial	Final
$t_{2,1}^{(1)}$	6.0	3.1	$t_{2,1}^{(2)}$	6.0	3.1
$t_{2,0}^{(1)}$	3.0	2.5	$t_{2,0}^{(2)}$	3.0	2.5
$t_{1,0}^{(1)}$	2.0	0.3	$t_{1,0}^{(2)}$	2.0	0.3
$\phi_{2,1}^{(1)}$	$\pi/4$	0.0	$\phi_{2,1}^{(2)}$	$-\pi/4$	0.0
$\phi_{2,0}^{(1)}$	$\pi/4$	0.1	$\phi_{2,0}^{(2)}$	$-\pi/4$	-0.2
$\phi_{1,0}^{(1)}$	$\pi/4$	0.0	$\phi_{1,0}^{(2)}$	$-\pi/4$	0.0

Table 5.5: Case 1: System parameters for two three-level sensors and k=2.0.

	Initial	Final		Initial	Final
$t_{2,1}^{(1)}$	2.0	1.1	$t_{2,1}^{(2)}$	2.0	1.1
$t_{2,0}^{(1)}$	1.0	0.5	$t_{2,0}^{(2)}$	1.0	0.5
$t_{1,0}^{(1)}$	2.0	1.1	$t_{1,0}^{(2)}$	2.0	1.1
$\phi_{2,1}^{(1)}$	$\pi/4$	1.78	$\phi_{2,1}^{(2)}$	$\pi/4$	1.78
$\phi_{2,0}^{(1)}$	0.0	0.5	$\phi_{2,0}^{(2)}$	0.0	0.5
$\phi_{1,0}^{(1)}$	$-\pi/4$	-1.78	$\phi_{1,0}^{(2)}$	$-\pi/4$	-1.78

Table 5.6: Case 2: System parameters for two three-level sensors and k=2.0.

$A_j^{(p)}$. Due to the symmetry in each quantization region, all of the $\tilde{l}_j^{(p)}$ values in the system are zero. This implies that the correlation term in λ_{LO} is not important. The resulting system is identical to the optimum scheme under the assumption of independent observations. We will call this result Case 1. The time evolution of the system parameters in Case 1 appears in Figure 5.5. The values which appear in Table 5.6 describe a system with anti-symmetric quantization regions, where the correlation term is non-zero for two of the quantization regions. These values are different from those in Case 1. We will label this result Case 2. The time evolution of the system parameters for Case 2 appears in Figure 5.6.

Because of the disparity in the two sets of results, the relative performance of the two resulting systems was compared by measuring the mean-square error J in each scheme. For Case 1, J was measured to be 4.67 while for Case 2, J was measured to be 3.56, much lower than the mean-square error for Case 1. This implies that the anti-symmetric quantization scheme in Case 2 performs better than the symmetric

quantization scheme in Case 1.

Consider the test statistic λ_{LO} in the two cases above and the test statistic λ_{UQ} in the centralized scheme. Figure 5.7 shows the test statistic values at the fusion center for each of these three schemes. Note that in Figure 5.7(a), λ_{LO} in Case 1 is symmetric about the $x^{(1)}$ axis and about the $x^{(2)}$ axis due to the symmetric quantization regions. In Case 2, the regions are antisymmetric, and if $x^{(1)}$ and $x^{(2)}$ have opposite signs, then the test statistic in Figure 5.7(b) has a lower value than if the observations are of the same sign. This appears to more closely approximate λ_{UQ} in Figure 5.7(c) where the test statistic is small for observations of opposite sign and large for observations of the same sign.

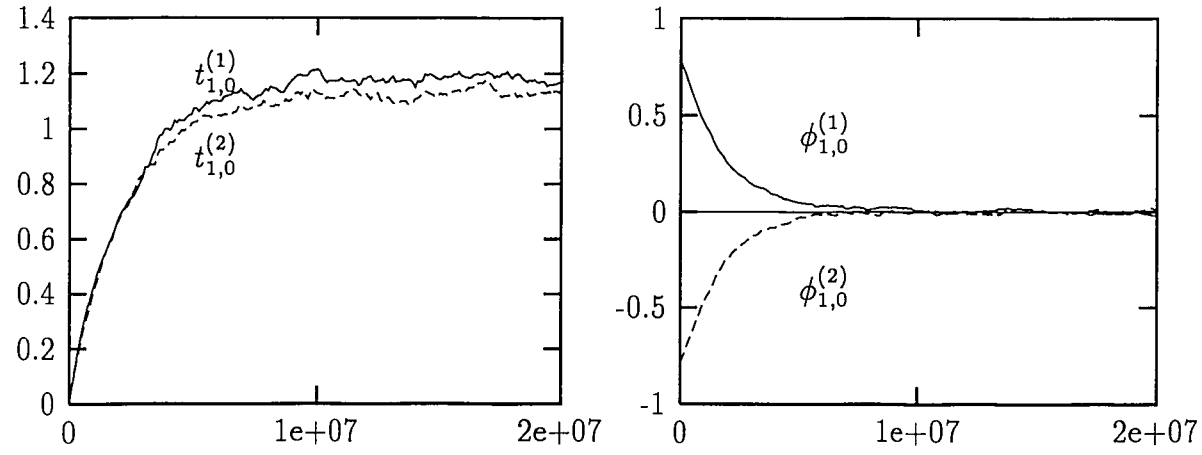


Figure 5.1: Time evolution of the parameters for a system with two binary sensors and generalized Gaussian noise with $k = 2.0$.

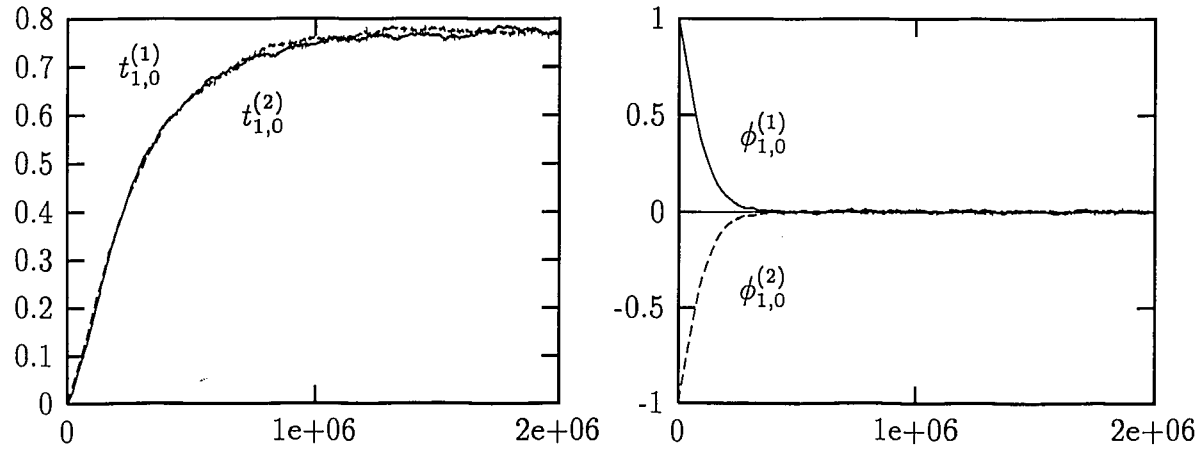


Figure 5.2: Time evolution of the parameters for a system with two binary sensors and generalized Gaussian noise with $k = 1.8$.

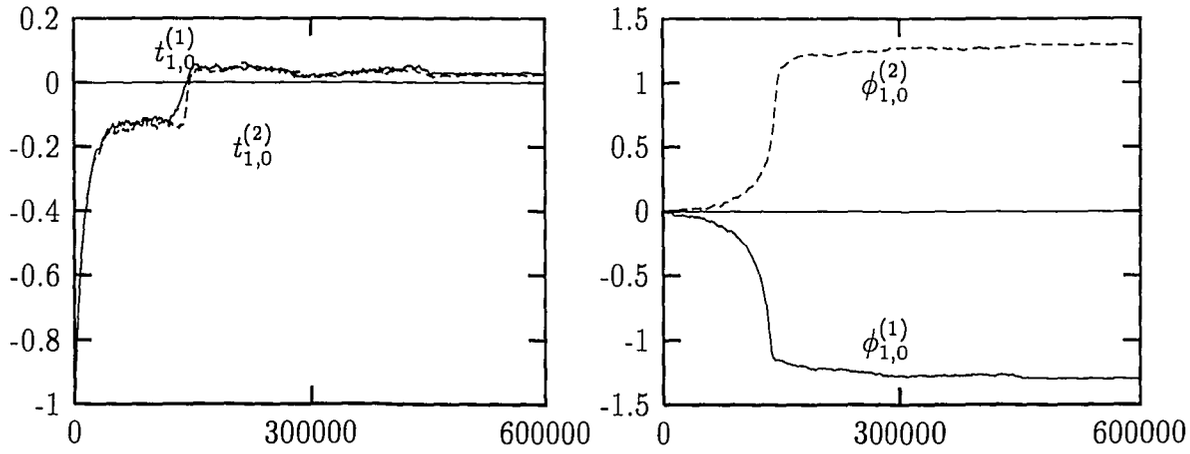


Figure 5.3: Time evolution of the parameters for a system with two binary sensors and generalized Gaussian noise with $k = 1.6$.

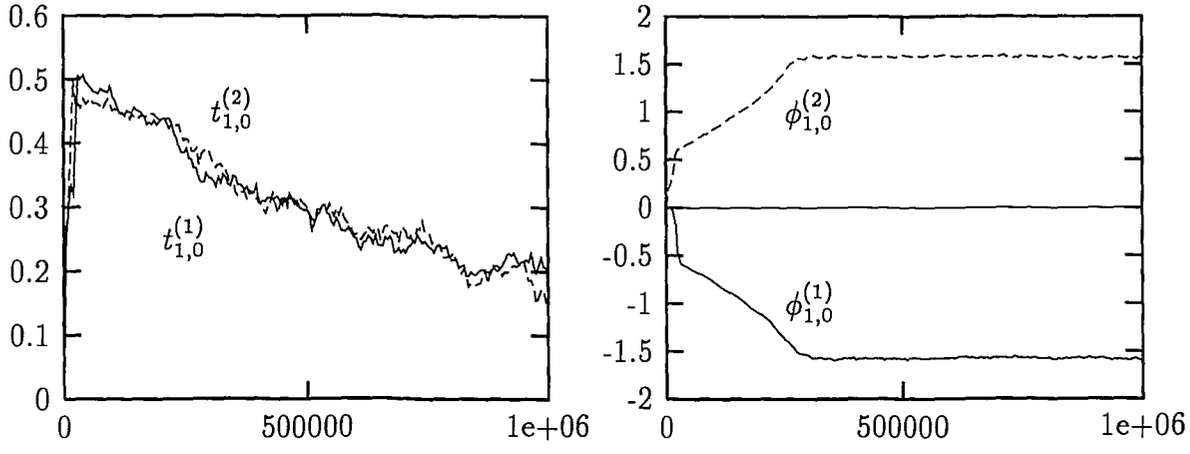


Figure 5.4: Time evolution of the parameters for a system with two binary sensors and Cauchy noise.

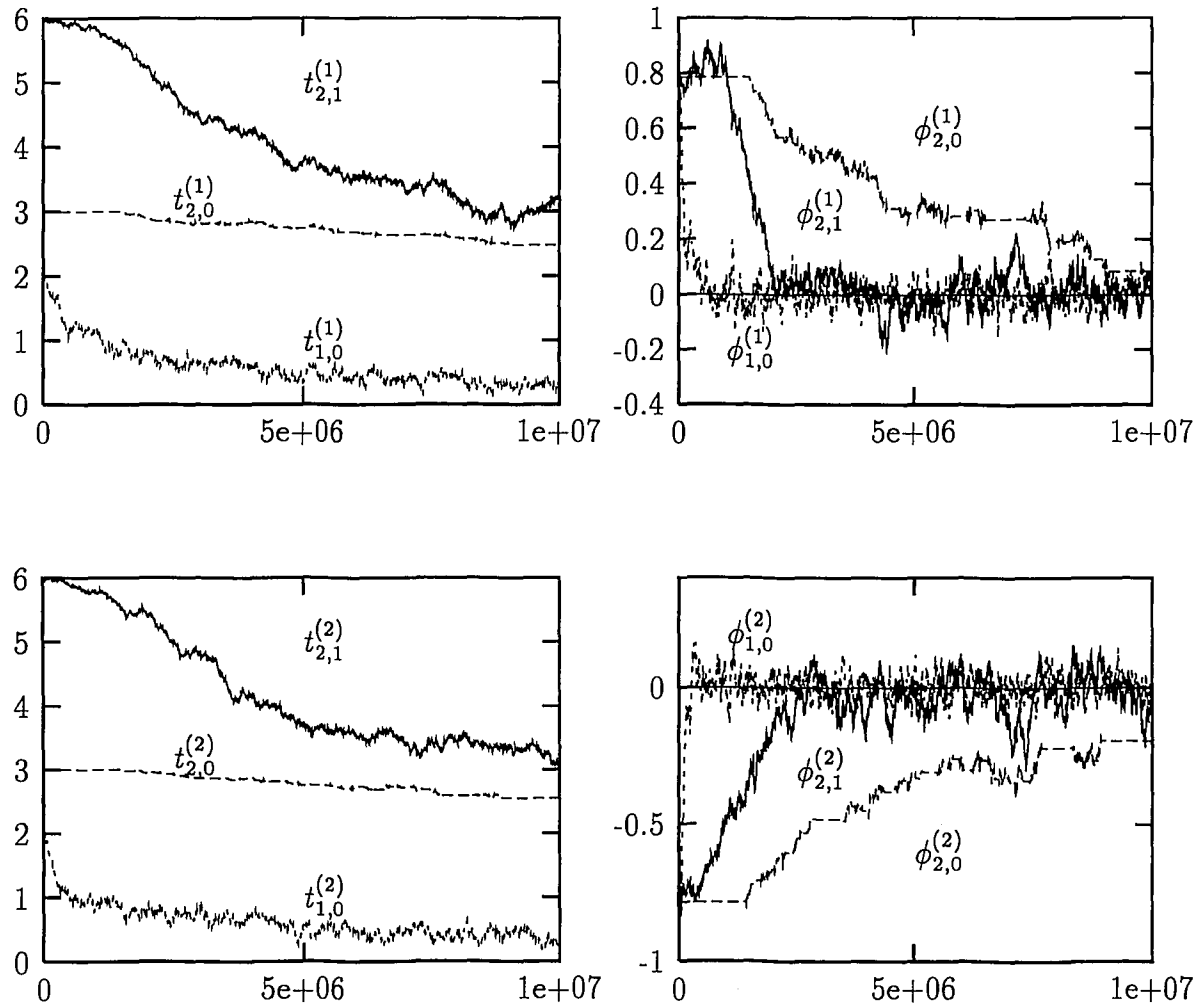


Figure 5.5: Case 1: Time evolution of the parameters for a system with two ternary sensors and generalized Gaussian noise with $k = 2.0$.

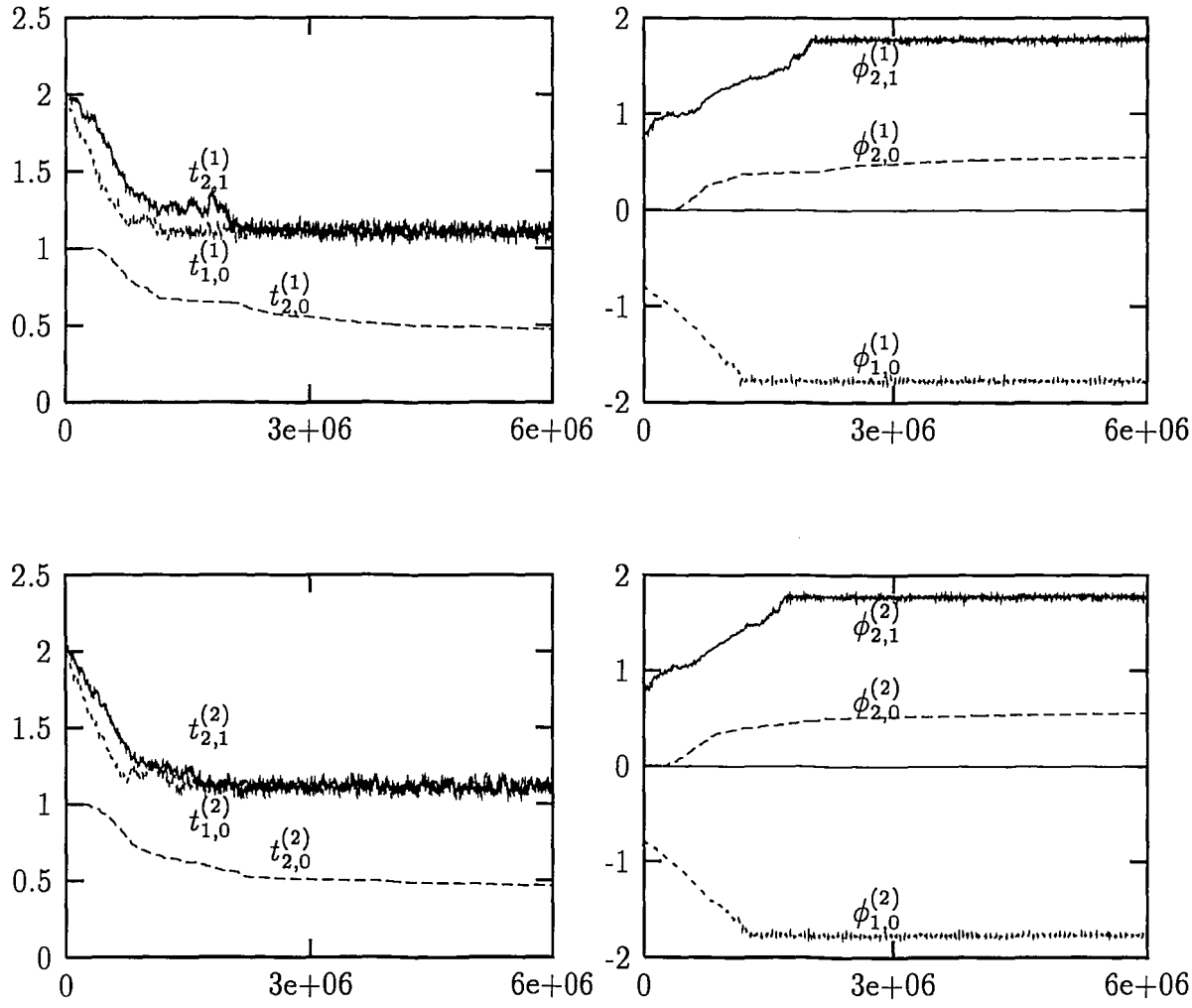
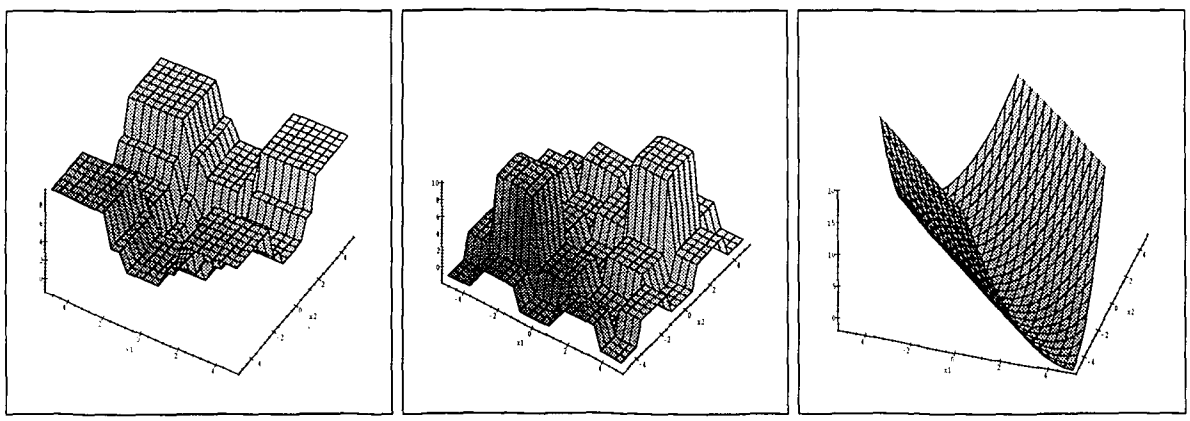


Figure 5.6: Case 2: Time evolution of the parameters for a system with two ternary sensors and generalized Gaussian noise with $k = 2.0$.



(a) (b) (c)

Figure 5.7: Comparison of fusion center test statistics as a function of two observations. (a) λ_{LO} for Case 1. (b) λ_{LO} for Case 2. (c) λ_{UQ}.

Chapter 6

Conclusion

We have presented an adaptive algorithm which can be used to train a distributed detection system to learn system parameters which provide good performance for the detection of weak random signals. The sensor test statistics and thresholds are adapted using a gradient descent method in an attempt to achieve a minimum mean-square error match to the best centralized test statistic. This method is easier to implement than techniques suggested in previous research [2].

In some of the cases we considered, the system parameters found using this method were close to those found to be optimum in previous research. In other cases, system parameters were found which offered lower mean-square error approximations to the corresponding centralized schemes than those found under the assumption of independent observations. Some results have been presented for cases which have not been studied elsewhere.

Bibliography

[1] S. A. Kassam, *Signal Detection in Non-Gaussian Noise,* Springer-Verlag: New York, NY, 1988.

[2] R. S. Blum, "Quantization in multisensor random signal detection," *IEEE Transactions on Information Theory*, Vol. 41, No. 1, pp. 204-215, Jan. 1995.

[3] X. Zhang and R. S. Blum, "Distributed quantization for signal detection in dependent sensors," *27th Annual Conference on Information Sciences and Systems*, Princeton University, Princeton, NJ, pp. 726-731, March 1994.

[4] S. Marcos, O. Macchi, and C. Vignat, "A unified framework for gradient algorithms used for filter adaptation and neural network training," *International Journal of Circuit Theory and Applications*, Vol. 20, No. 2, pp. 159-200, Mar 1992.

Biography

Matthew C. Deans was born to Mark and Cheryl Deans in May 1972 in Bristol, PA. He received the B.S degree in Electrical Engineering in 1994, the B.S. degree in Engineering Physics in 1995, and the M.S in Electrical Engineering in 1996 from Lehigh University, Bethlehem, PA. He has held positions with Johnson & Johnson, Inc. in New Brunswick, NJ and Electronic Technology, Inc. in Irvington, NJ. He has also held positions as a research assistant in the Physics Department in 1993 and the EECS Department from 1994 to the present at Lehigh University. His publications include:

K. R. Elder, Hao-Wen Xi, M. Deans, J. D. Gunton, "Spatiotemporal Chaos in the Damped Kuramoto-Sivashinsky Equation," p. 702 in *AIP Conf. Proc. 342,* CAM-94 Physics, Cancun, Mexico 1994, Arnulfo Zepeda (Ed.) AIP Press 1995.

M. C. Deans, R. S. Blum, "An Adaptive Algorithm for Distributed Signal Detection System Design," *Sensor Fusion and Networked Robotics VIII,* Proc. SPIE 2589, pp 172-179 (1995).

M. C. Deans, R. S. Blum, "Distributed Signal Detection System Design Using Adaptive Signal Processing Techniques," *Proc. CISS 96*, Princeton University, Princeton, NJ, March 1996.

END OF TITLE