AFCRL-67-0071

41

Estimation And Detection Of Optical Signals Distorted By Diffraction, Background Noise, And Detection Noise

Mark C. Austin
Craig K. Rushforth

Electro-Dynamics Laboratories
Departments of Electrical Engineering, Physics, and Chemistry
UTAH STATE UNIVERSITY
Logan, Utah

Contract No. AF19(628)-3825
Project No. 8663

SCIENTIFIC REPORT NO. 9

31 December 1966

Dr. G. A. Vanasse

This research was sponsored by the Advanced Research Projects Agency under ARPA Order No. 450

Distribution of this document is unlimited.

Prepared for

Air Force Cambridge Research Laboratories
Office of Aerospace Research
UNITED STATES AIR FORCE
Bedford, Massachusetts

AFCRL-67-0071

ESTIMATION AND DETECTION OF OPTICAL SIGNALS DISTORTED BY DIFFRACTION, BACKGROUND NOISE, AND DETECTION NOISE

Mark C. Austin
Craig K. Rushforth

Electro-Dynamics Laboratories
Departments of Electrical Engineering, Physics, and Chemistry
UTAH STATE UNIVERSITY
Logan, Utah

Contract No. AF19(628)-3825

Project No. 8663

SCIENTIFIC REPORT NO. 9

31 December 1966

Dr. G.A. Vanasse

This research was sponsored by the Advanced Research Projects Agency under ARPA Order No. 450

Distribution of this document is unlimited.

Prepared for

Air Force Cambridge Research Laboratories
Office of Aerospace Research
UNITED STATES AIR FORCE
Bedford, Massachusetts

ABSTRACT

Estimation and detection of optical signals distorted by diffraction, additive background noise, and multiplicative (detection) noise are studied. Assuming that the output of the detector is a Poisson process, that the signal and noise are additive, and that they have prescribed moans and covariance matrices, the optimum linear estimate of the optical signal or object is obtained. In the physical detection process, the interaction between the incident radiation and the detector produces an effect called multiplicative noise which must be taken into account in obtaining the optimum linear estimate. The performance of the estimation procedure is evaluated for several special cases. Both white and colored noise are considered in the estimation problem. The problem of discriminating between optical signals is considered. Optimum procedures are derived for detecting known and unknown optical signals using fixed-sample detectors. The properties of sequential detectors which are optimum for the detection of random or unknown optical signals are investigated. A comparison is made of the average test lengths of these optimum random signal detectors with those of a detector designed for particular optical signals. The test lengths of the fixed-sample detector and sequential detector are compare for a particular example.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

In an optical system the final image is not an exact representation of the original object. In general the image differs from the object due to diffraction and stray light or additive background noise. The problem is further complicated when the image is measured. When measurements are made, detection noise (multiplicative noise) is introduced.

In the absence of any noise, with distortion due only to diffraction, Harris (1964) showed that the object can in principle be reconstructed exactly if the object is known to be spatially bounded. In general, however, additive and multiplicative noise will be present and will give rise to error in any restoration procedure. In establishing such a procedure, we need to take into account any known statistics since the restoration procedure in the presence of noise may be different from the procedure used when noise is absent.

In this paper, methods of detecting and estimating optical signals which have been distorted by diffraction, additive noise, and multiplicative noise are investigated. The estimation procedures considered are the minimum mean-square-error estimate, the maximum a posteriori estimate, the maximum likelihood estimate, and the Bayes' estimate. The main emphasis will be on the minimum mean-square-error estimate. For the detection procedures, both fixed-sample detection and sequential detection are studied. Comparisons are made between detecting known signals and unknown signals to determine the deterioration in performance due to ignorance about the unknown signals.

STATEMENT OF THE PROBLEM

Throughout this paper the conditions necessary for Fraunhofer diffraction will be assumed to be satisfied (Born and Wolf, 1964; Stone, 1963). In essence, these conditions require that the effective distances from a point in the object plane (or observation point in the image plane) to any two points in the aperture plane differ by not more than a small fraction of a wavelength. Also, the radiation will be assumed to be spatially incoherent and quasi-monochromatic. By quasi-monochromatic we mean that the radiation has a frequency bandwidth which is much smaller than the frequency itself. Unless stated otherwise, the observed quantities will be number-of-photoelectrons and the estimated quantities will be average-number-of-photons per unit time (mean rate).

Under the conditions of Fraunhofer diffraction, the techniques of Fourier analysis can be used to investigate the characteristics of the optical system (O'Neill, 1963). The optical system can then be treated as a linear filter of spatial frequencies whose properties are described by a transfer function $A\ (f_\xi,\ f_\zeta)$ where ξ and ζ are image-plane coordinates. For incoherent illumination, the spatial frequency spectrum of the image $V\ (f_\xi,\ f_\zeta)$ is found by multiplying the spatial frequency spectrum of the object $W\ (f_\alpha,\ f_\beta)$, where α and β are the object-plane coordinates, by the system transfer function $A\ (f_\xi,\ f_\zeta)$ (see Appendix). Alternately, by the convolution theorem, the image intensity distribution $v(\xi,\ \zeta)$ is obtained by convolving the object intensity distribution $w(\alpha, \beta)$ with the point spread function $a(\xi, \zeta)$ of the optical system. Hence,

$$v(\xi, \zeta) = w(\xi, \zeta) * a(\xi, \zeta) \quad (1)$$

where $w(\xi, \zeta)$ is the object intensity distribution referred to the image plane and * denotes convolution.

The image $v(\xi, \zeta)$ is further distorted by additive background noise $q(\xi, \zeta)$ and the resulting image intensity distribution is $r(\xi, \zeta) = v(\xi, \zeta) + q(\xi, \zeta)$. During the detection of $r(\xi, \zeta)$ the interaction between the impinging radiation and the detector produces a multiplicative effect or detection noise resulting in an image $u(\xi, \zeta)$ or a stream of photoelectrons z. Our objective is to count the number of photoelectrons in the output and from this, estimate $w(\xi, \zeta)$ or discriminate between two alternative signals $w_1(\xi, \zeta)$ and $w_2(\xi, \zeta)$. The estimation and discrimination procedures we develop depend upon the statistics of the additive and multiplicative noise as well as any a priori information available concerning the optical signals to be estimated or detected.

STATISTICAL MODEL

Radiation can be observed only through its interaction with matter. The interaction process we will consider results from the photolectric effect. The receptor in the image plane will be assumed to be a photosensitive surface divided into a large number of very small regions or cells. It is assumed that the cells are small enough that the illuminance is approximately constant within a given cell. Consider the radiation incident upon the elementary regions or cells to be streams of photons each with energy $h\nu$ where h is Planck's constant and ν is the frequency of the incident radiation. The average number of photons $\bar{y}$ incident upon a cell in the time interval τ is equal to the incident energy of that cell τr divided by $h\nu$, where r is the received intensity due to the diffracted object and additive noise. The cells are labeled with the index i, and y_i represents the number of photons incident upon the i^{th} cell. The number of photoelectrons z_i emitted from the i^{th} cell depends upon the incident energy and also upon the multiplicative effect of the receptor. Because of the stochastic nature of the interaction between radiation and matter, for a given y_i the quantity z_i is a random variable rather than a determinintic quantity and must be described in probabilistic terms.

The number of photoelectrons z_i emitted from each cell consitiutes the observed data. It is assumed that the location where each photoelectric event takes place can be determined.

The photons that strike the light sensitive surface of the receptor will cause some type of reaction that can be measured. For example, in the photographic film case, the photons will cause many of the silver halide grains to become developable. The pattern that results on the developed photographic film will be a measure of the number of the photons reaching the image plane. In this case, film grandularity and saturation must be taken into account when determining the number of incoming photons. In the photomultiplier tube case, a single photon that strikes the light sensitive plate gives rise to many electrons in the output. By scanning the image plane with a photomultiplier tube it is possible to obtain an estimate of the number of photons that are incident upon each of the incremental cells. For a simple photon-electron converter, a photon gives rise to a single photoelectron with probability η. The quantity η is called the quantum efficiency. The photon-electron converter is a degenerate case of the photomultiplier tube case in which we consider only the first stage of the photomultiplier tube.

Throughout this paper we will assume that if the incident energy per unit time v_i (or mean rate of signal photons $\bar{s}_i$) from an optical signal is known, the signal photons statistics are Poisson with the probability that exactly s_i signal photons will impinge upon the i^{th} cell in time τ given by

$$p(\bar{s}_i) = \frac{(\tau\bar{s}_i)^{s_i} e^{-\tau\bar{s}_i}}{s_i!} \qquad (3)$$

where the mean rate $\bar{s}_i = v_i/h\nu$. Likewise, if the incident energy per unit time q_i (or mean rate of noise photons $\bar{n}_i$) from the additive noise is known, the noise photon statistics are assumed Poisson with the probability that exactly n_i noise photons will impinge upon the i^{th} cell in time τ given by

$$p(n_i) = \frac{(\tau \bar{n}_i)^{n_i} e^{-\tau \bar{n}_i}}{n_i!} \tag{3}$$

where the mean rate $\bar{n}_i = q_i/h\nu$.

For the unknown signal case, the mean rate of signal photons $\bar{s}_i$ incident upon the i^{th} cell in the image plane is a random variable. The statistical properties of $\bar{s}_i$ need to be considered since they influence the statistical properties of the signal photoelectrons emitted during the reception of a signal. For this case the prior distribution of $\bar{s}_i$ will be assumed to be the following gamma distribution (Goodman, 1965; Farrell, 1966):

$$f(\bar{s}_i) = \frac{\beta_i (\beta_i \bar{s}_i)^{u_i - 1} e^{-\beta_i \bar{s}_i}}{\Gamma(u_i)}, \quad s_i \geq 0 \tag{4}$$

$$= 0, \qquad \text{otherwise}$$

where $E(\bar{s}_i) = \bar{\bar{s}}_i = u_i/\beta_i$ and var $(\bar{s}_i) = \sigma_{\bar{s}_i}^2 = u_i/\beta_i^2$. For this unknown signal case the probability that s_i signal photons will impinge upon the i^{th} cell in time τ is given by

$$p(s_i) = \int_0^\infty p(s_i / \bar{s}_i) \, f(\bar{s}_i) \, d\bar{s}_i \tag{5}$$

where

$$p(s_i/\bar{s}_i) = \frac{(\tau \bar{s}_i)^{s_i} e^{-\tau \bar{s}_i}}{s_i !} \tag{6}$$

and $f(\bar{s}_i)$ is the gamma distribution in (4).

For the case of known signal and noise both the signal and noise photons obey Poisson statistics. Because of the additive nature of the Poisson distribution the total photon stream has a Poisson distribution. The probability that y_i signal-plus-noise photons will impinge upon the i^{th} cell in time τ is given by

$$p(y_i/\bar{s}_i, \bar{n}_i) = \frac{[\tau(\bar{s}_i + \bar{n}_i)]^{y_i} e^{-[\tau(\bar{s}_i + \bar{n}_i)]}}{y_i !} \tag{7}$$

$$= \frac{(\tau \bar{y}_i)^{y_i} e^{-\tau \bar{y}_i}}{y_i !} .$$

For the case where the mean rate of signal photons $\bar{s}_i$ is unknown and the mean rate of noise photons is known the probability of exactly y_i photons impinging upon the i^{th} cell in time τ is given by

$$p(y_i/\bar{n}_i) = \int_0^\infty p(y_i/\bar{s}_i, \bar{n}_i) f(\bar{s}_i) d\bar{s}_i .$$

$$= \left(\frac{\beta_i}{\tau + \beta_i}\right)^{u_i} \frac{e^{-\tau \bar{n}_i} (\bar{n}_i \tau)}{\Gamma(u_i)} \sum_{j=o}^{y_i} \frac{\Gamma(u_i + j)}{(y_i - j)! \, j! \, [\bar{n}_i (\tau + \beta_i)]^j} . \tag{8}$$

When the mean rate of noise photons $\bar{n}_i$ is unknown we will assume its prior distribution to be the following gamma distribution:

$$f(\bar{n}_i) = \frac{\lambda_i(\lambda_i \bar{n}_i)^{u_i - 1} e^{-\lambda_i \bar{n}_i}}{\Gamma(u_i)}, \quad \bar{n}_i \geq 0 \tag{9}$$

$$= 0, \quad \text{otherwise.}$$

When the mean rates of noise photons $\bar{n}_i$ and signal photons $\bar{s}_i$ are unknown (i. e., $\bar{y}_i = \bar{s}_i + \bar{n}_i$ is unknown) we will assume that the prior distribution of the sum $\bar{n}_i + \bar{s}_i$ is given by the following gamma distribution:

$$f(\bar{y}_i) = f(\bar{s}_i + \bar{n}_i) = \frac{\alpha_i(\alpha_i \bar{y}_i)^{u_i - 1} e^{-\alpha_i \bar{y}_i}}{\Gamma(u_i)}, \quad \bar{y}_i \geq 0 \tag{10}$$

$$= 0, \quad \text{otherwise.}$$

MATRIX REPRESENTATION

In this paper the spatially bounded objects are divided into small cells over which the intensity is approximately constant. If these cells are made small enough, they may represent point sources. By knowing the point spread function of the system, the image can be approximated for incoherent light by superposition of the point spread functions resulting from all the point sources of the disected object. Here we are assuming spatial invariance. By this we mean that the object is small enough that points of a given intensity located anywhere on the object gives rise to the same point spread function in the image plane. The location of the point spread function is determined by the position of the point source (O'Neill, 1963).

The number of photons emitted from the j^{th} region (point source) of the object plane $x_j(\alpha_j, \beta_j)$ can be represented by a delta function $x_j(\alpha_j, \beta_j) = x_j\,\delta(\alpha-\alpha_j, \beta-\beta_j)$ where α and β are the coordinate representation in the object plane. Consider the system impulse response or point spread function $a(\xi, \zeta)$ where ξ and ζ are the coordinate representation in the image plane. The optical image or point spread function in the image plane due to the single point source $x_j(\alpha_j, \beta_j)$ is, using image plane coordinates, $x_j(\xi_j, \zeta_j) * a(\xi, \zeta) =$ $x_j\,\delta(\xi-\xi_j, \zeta-\zeta_j) * a(\xi, \zeta) = x_j a(\xi-\xi_j, \zeta-\zeta_j) = s_j(\xi-\xi_j, \zeta-\zeta_j)$. The total image is the superposition of the images of all m point sources.

That is,

$$s(\xi, \zeta) = \sum_{j=1}^{m} x_j a(\xi-\xi_j,\ \zeta-\zeta_j) = \sum_{j=1}^{m} s_j(\xi-\xi_j,\ \zeta-\zeta_j). \tag{11}$$

The image at the point (ξ^i, ζ^i) will be due to the images of the m point sources, and hence we have

$$s(\xi^i, \zeta^i) = \sum_{j=1}^{m} s_j(\xi^i-\xi_j,\ \zeta^i-\zeta_j) = \sum_{j=1}^{m} a(\xi^i-\xi_j,\ \zeta^i-\zeta_j)x_j. \tag{12}$$

Let $s_i = s(\xi^i, \zeta^i)$ and $a_{ij} = a(\xi^i - \xi_j, \zeta^i - \zeta_j)$. The quantity s_i is the number of photons due to the optical signal incident upon the i^{th} cell of the image plane. We can then write

$$s_i = \sum_{j=1}^{m} a_{ij}x_j. \tag{13}$$

In matrix notation this can be written as

$$s = Ax \tag{14}$$

where

$$s = \begin{pmatrix} s_1 \\ \vdots \\ s_\ell \end{pmatrix},\quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_\ell \end{pmatrix},\ \text{and } A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{\ell 1} & \cdots & a_{\ell m} \end{pmatrix}.$$

If additive noise is present at the image, each measurement of s_i will be corrupted by an additive noise element n_i; hence, we have

$$y_i = s_i + n_i = \sum_{j=1}^{m} a_{ij} x_j + n_i. \qquad (15)$$

The quantity n_i is the number of photons due to additive noise, incident upon the i^{th} cell of the image plane. In matrix notation the observed vector is

$$y = s + n = Ax + n \qquad (16)$$

where

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_\ell \end{pmatrix} \text{ and } n = \begin{pmatrix} n_1 \\ \vdots \\ n_\ell \end{pmatrix}.$$

The vector y will, during the detection process, be contaminated by multiplicative noise, the form of which will be discussed later.

We will assume throughout this paper that the vector representation of the object is sufficiently accurate that any error associated with it is negligible.

MINIMUM MEAN-SQUARE-ERROR ESTIMATE

Introduction

In this section we investiage the problem of obtaining optimum estimates of an optical signal or object distorted by diffraction, additive background noise, and multiplicative noise using the criterion of minimum mean-square error.

Consider the photon-stream vector y impinging upon the light sensitive surface in the image plane and giving rise to the output vector z. The quantity y_i is the number of photons due to the optical signal and additive noise incident upon the i^{th} cell of the image plane. The quantity z_i is the number of photoelectrons due to the optical signal, additive noise, and multiplicative noise, being emitted from the i^{th} cell of the image plane. The quantity z_i can be thought of as the number of counts (i.e., photoelectrons for the photomultiplier tube case and developable grains in the photographic film case) in the output of the detector. The system being considered is illustrated in Figure 1.

For a given mean rate vector $\bar{n}$ we will assume that the noise n has a conditional covariance matrix K_n. We will assume that for a given mean rate vector $\bar{x}$ the object x has a conditional covariance matrix K_x and that x and n are conditionally independent (i.e., conditioned on knowing $\bar{n}$ and $\bar{x}$). Also, we will assume that the mean rate vector

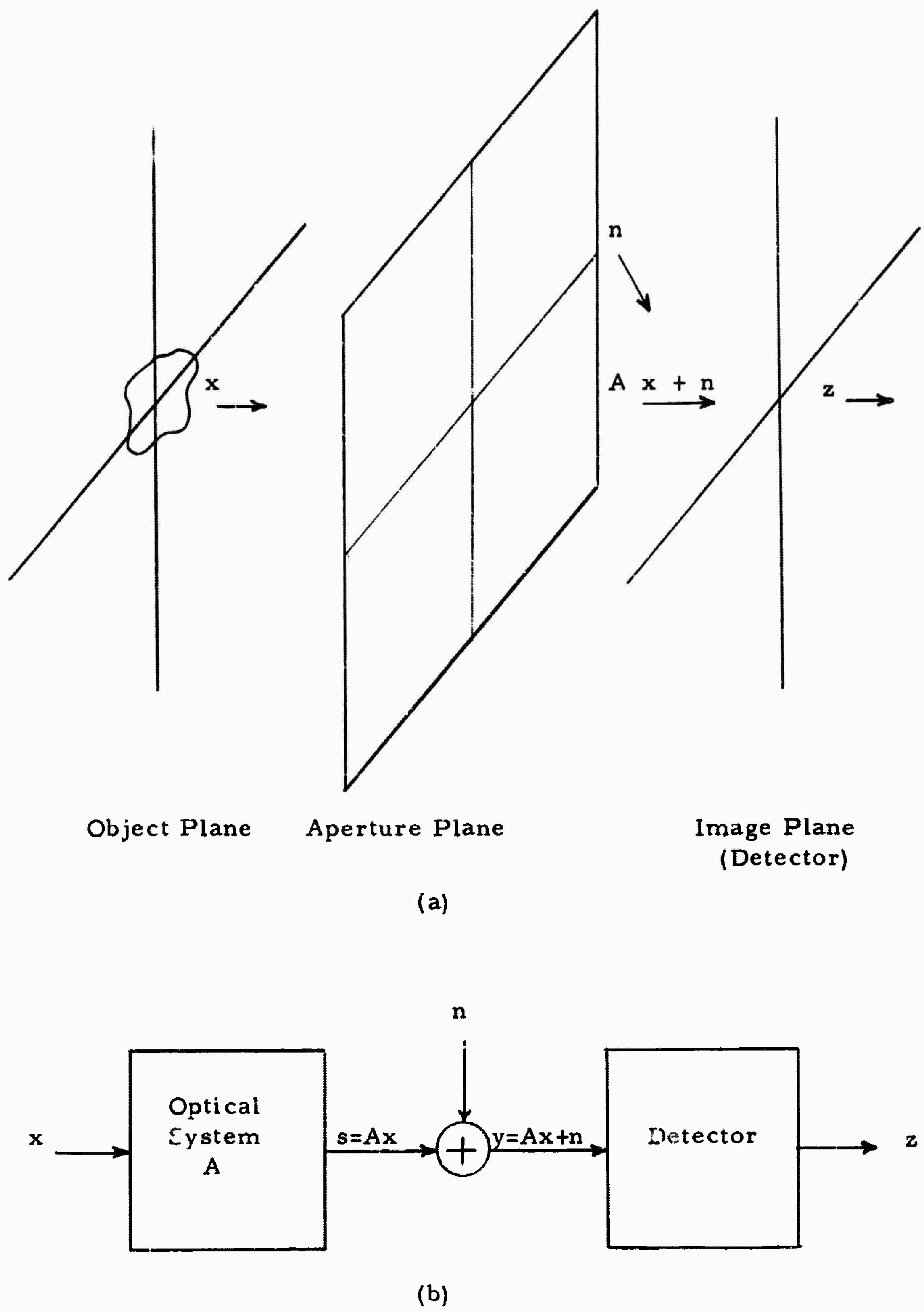

Figure 1. Optical system configuration.

$\bar{n}$ has mean $\bar{\bar{n}}$ and a covariance matrix $K_{\bar{n}}$ and that the mean rate vector $\bar{x}$ has a mean $\bar{\bar{x}}$ and a covariance matrix $K_{\bar{x}}$ and that $\bar{x}$ and $\bar{n}$ are independent. $K_{\bar{x}}$ and $\bar{\bar{x}}$ as used here represent our prior information about the mean rate of the object rather than any actual statistical fluctuation of the object. Large terms in $K_{\bar{x}}$ indicate small prior information about the mean rate of the object and small terms imply large prior information.

Photomultiplier Tube Detector

For the case where the photomultiplier tube acts as the detector of the optical image, a photon k gives rise to B_k electrons in the output of the detector. The output due to each photon is assumed to be independent of the outputs due to other photons but identically distributed with mean b and variance σ_b^2. From the photon stream incident upon the i^{th} cell of the image plane we have

$$z_i = \sum_{k=1}^{y_i} B_k. \tag{17}$$

The element z_i is a random number of independent random variables (Parzen, 1962). Because of the Poisson nature of the photon stream the conditional mean and conditional variance of y_i are equal (i.e., $\mathrm{var}\ (y_i / \bar{x}, \bar{n}) = E(y_i/\bar{x}, \bar{n})$). The mean and variance of z_i are respectively

$$E(z_i) = E(y_i)E(B) = (a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\,\tau b, \tag{18}$$

$$\begin{aligned} Var\,(z_i) &= E(y_i)Var(B) + Var(y_i)E^2(B) \\ &= \tau_b^2\,\tau(a_i\bar{\bar{x}} + \bar{\bar{n}}_i) + b^2(a_i\bar{K}_x a_i' + \overline{K_{n_i n_i}}) + \tau^2 b^2(a_i K_{\bar{x}} a_i' + K_{\bar{n}_i \bar{n}_i}) \\ &= (\sigma_b^2 + b^2)\,\tau(a_i\bar{\bar{x}} + \bar{\bar{n}}_i) + \tau^2 b^2(a_i K_{\bar{x}} a_i' + K_{\bar{n}_i \bar{n}_i}) \end{aligned} \tag{19}$$

where a_i is defined as the i^{th} row of the system matrix A and the prime on a_i denotes transpose. Thus,

$$a_i\bar{\bar{x}} = \left(a_{i1}, \ldots, a_{im}\right) \begin{pmatrix} \bar{\bar{x}}_1 \\ \vdots \\ \bar{\bar{x}}_m \end{pmatrix} = a_{i1}\bar{\bar{x}}_1 + \ldots\ a_{im}\bar{\bar{x}}_m. \tag{20}$$

The quantities $\bar{\bar{x}}$ and $\bar{\bar{n}}_i$ are defined as $E(\bar{x})$ and $E(\bar{n}_i)$ respectively. Point spread functions a_{ij} for various optical system apertures are derived in the Appendix.

For the vector case, the expected mean of z is

$$E(z) = \tau\, b(A\bar{\bar{x}} + \bar{\bar{n}}) \tag{21}$$

and the covariance matrix of z is

$$Cov\,(z_i, z_j) = E(zz') - E(z)\;E(z') \tag{22}$$

where

$$E(z)\,E(z') = \tau^2 b^2\,(A\bar{\bar{x}}\bar{\bar{x}}'A' + A\bar{\bar{x}}\bar{\bar{n}}' + \bar{\bar{n}}\bar{\bar{x}}'A' + \bar{\bar{n}}\bar{\bar{n}}'). \tag{23}$$

We need to calculate $E(zz')$ which can be written as

$$E(zz') = \begin{bmatrix} E(z_1 z_1) & \cdots & E(z_1 z_m) \\ \vdots & & \vdots \\ E(z_m z_1) & \cdots & E(z_m z_m) \end{bmatrix}. \tag{24}$$

For the diagonal terms of (24) we have

$$\begin{aligned} E(z_i z_j) &= E[E(z_i z_i / x, n)] \\ &= \tau^2 b^2 [a_i (K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}') a_i' + \bar{\bar{n}}_i \bar{\bar{x}}' a_i' + a_i \bar{\bar{x}}\bar{\bar{n}}_i + K_{\bar{n}_i \bar{n}_i} + \bar{\bar{n}}_i \bar{\bar{n}}_i] \\ &\quad + b^2 [a_i \bar{K}_x a_i' + \overline{K_{n_i n_i}}] + \tau \sigma_b^2 (a_i \bar{\bar{x}} + \bar{\bar{n}}_i) \\ &= (\sigma_b^2 + b^2) \tau (a_i \bar{\bar{x}} + \bar{\bar{n}}_i) + \tau^2 b^2 [a_i (K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}') a_i' + \bar{\bar{n}}_i \bar{\bar{x}}' a_i' + a_i \bar{\bar{x}}\bar{\bar{n}}_i \\ &\quad + K_{\bar{n}_i \bar{n}_i} + \bar{\bar{n}}_i \bar{\bar{n}}_i]. \end{aligned} \tag{25}$$

Now consider the off-diagonal elements of (24) which are

$$\begin{aligned} E(z_i z_j) &= E[E(z_i z_j / x, n)] \\ &= \tau^2 b^2 [a_i (K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}') a_j' + a_i \bar{\bar{x}}\bar{\bar{n}}_j + \bar{\bar{n}}_i \bar{\bar{x}}' a_j' + K_{\bar{n}_i \bar{n}_j} + \bar{\bar{n}}_i \bar{\bar{n}}_j] \\ &\quad + b^2 [a_i \bar{K}_x a_j' + \overline{K_{n_i n_j}}] + \rho_{ij} \sigma_b^2 \, E\sqrt{(a_i x + n_i)(a_j x + n_j)} \end{aligned} \tag{26}$$

where ρ_{ij} is the correlation coefficient of z_i and z_j. We will make the following definition:

$$C_z = [C_{z_{ij}}] \triangleq \left[\left(\rho_{ij} E\sqrt{(a_i x + n_i)(a_j x + n_j)}\right) \right]. \tag{27}$$

Thus we can write

$$E(zz') = b^2\tau^2[A(K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}')A' + \bar{\bar{n}}\bar{\bar{x}}'A' + A\bar{\bar{x}}\bar{\bar{n}}' + K_{\bar{n}} + \bar{\bar{n}}\bar{\bar{n}}'] \tag{28}$$
$$+ b^2[A\bar{K}_x A' + \bar{K}_n] + \sigma_b^2 C_z.$$

In this paper we will assume that the number of photoelectrons emitted from disjoint regions or cells of the detector are statistically independent (Helstrom, 1964; Farrell, 1966). Hence, the multiplicative noise will be assumed to be uncorrelated. For uncorrelated multiplicative noise (i.e., $\rho_{ij} = 0$ $(i \neq j)$ and $\rho_{ii} = 1$) we have

$$C_z = \tau(a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\delta_{ij} = \begin{bmatrix} \tau(a_1\bar{\bar{x}} + \bar{\bar{n}}_1) & & 0 \\ & \ddots & \\ 0 & & \tau(a_m\bar{\bar{x}} + \bar{\bar{n}}_m) \end{bmatrix}. \tag{29}$$

The covariance matrix of z can now be written as

$$\mathrm{Cov}\,(z_i, z_j) = \sigma_b^2 C_z + b^2(A\bar{K}_x A' + \bar{K}_n) \tag{30}$$
$$= \tau\sigma_b^2\,(a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\,\delta_{ij} + b^2(A\bar{K}_x A' + \bar{K}_n).$$

We will assume for convenience that the number of photons impinging upon different regions or cells of the detector are statistically independent. This assumption is by no means essential. This assumption makes the matrix sum of $A\bar{K}_x A' + \bar{K}_n$ diagonal, and due to the Poisson nature of the photon stream this sum becomes $\tau\,(a_i\bar{x} + \bar{n}_i)\,\delta_{ij}$. The covariance matrix of z is then

$$\mathrm{Cov}\,(z_i, z_j) = (\sigma_b^2 + b^2)\tau(a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\,\delta_{ij}\,. \tag{31}$$

Also,

$$E(zz') = (\sigma_b^2 + b^2)C_z + \tau^2 b^2[A(K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}')A' + \bar{\bar{n}}\bar{\bar{x}}'A' + A\bar{\bar{x}}\bar{\bar{n}}' + K_{\bar{n}} + \bar{\bar{n}}\bar{\bar{n}}']. \tag{32}$$

We now want to find the linear estimate $\hat{\bar{x}}$ of $\bar{x}$ which will minimize the mean-square error (MSE). That is, find $\hat{\bar{x}}$ to minimize

$$e = E[(\hat{\bar{x}}-\bar{x})'\,(\hat{\bar{x}}-\bar{x})] = \mathrm{tr}\,E[(\hat{\bar{x}}-\bar{x})(\hat{\bar{x}}-\bar{x})'] \tag{33}$$

where tr denotes the trace of a matrix.

For the linear estimate of $\bar{x}$ we write

$$\hat{\bar{x}} = Hz/\tau + v. \tag{34}$$

To simplify the mathematics later on let

$$v = -Hb(A\bar{\bar{x}} + \bar{\bar{n}}) + \omega. \tag{35}$$

The linear estimate of $\bar{x}$ is then

$$\hat{\bar{x}} = H[z/\tau - b(A\bar{\bar{x}} + \bar{\bar{n}})] + \omega. \tag{36}$$

We need to find a matrix H (discrete linear filter) and a vector ω such that the mean-square error is minimized. Substituting (36) into (33) and expanding yields

$$\epsilon = \left(\frac{1}{\tau^2}\right) \mathrm{tr}\, E\Big[H\big(zz' - b\tau z\bar{\bar{x}}'A' - b\tau z\bar{\bar{n}}' + b^2\tau^2 A\bar{\bar{x}}\bar{\bar{x}}'A'$$

$$+ b^2\tau^2 A\bar{\bar{x}}\bar{\bar{n}}' + b^2\tau^2 \bar{\bar{n}}\bar{\bar{n}}'A' + b^2\tau^2 \bar{\bar{n}}\bar{\bar{n}}' - b\tau A\bar{\bar{x}}z' - b\tau \bar{\bar{n}}z'\big)H'$$

$$+ H\big(\tau z\omega' - b\tau^2 A\bar{\bar{x}}\omega' - b\tau^2 \bar{\bar{n}}\omega' - \tau z\bar{x}' + b\tau^2 A\bar{\bar{x}}\bar{x}' + b\tau^2 \bar{\bar{n}}\bar{x}'\big) \tag{37}$$

$$+ \big(\tau\omega z' - b\tau^2 \omega\bar{\bar{x}}'A' - b\tau^2 \omega\bar{\bar{n}}' - \tau\bar{x}z' + b\tau^2 \bar{x}\bar{\bar{x}}'A' + b\tau^2 \bar{x}\bar{\bar{n}}'\big)H'$$

$$+ \big(\tau^2\omega\omega' - \tau^2\omega\bar{x}' - \tau^2\bar{x}\omega' + \tau^2\bar{x}\bar{x}'\big)\Big].$$

Carrying out the expectation operation we obtain

$$E(zz') = (\sigma_b^2 + b^2)C_z + b^2\tau^2\big[A(K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}')A' + \bar{\bar{n}}\bar{\bar{x}}'A' + A\bar{\bar{x}}\bar{\bar{n}}' + K_{\bar{n}} + \bar{\bar{n}}\bar{\bar{n}}'\big], \tag{38}$$

$$E(z\bar{x}') = b\tau(A\bar{\bar{x}}\bar{\bar{x}}' + \bar{\bar{n}}\bar{\bar{x}}'), \tag{39}$$

$$E(\bar{x}\bar{x}') = K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}', \tag{40}$$

$$E(\omega\bar{x}') = \omega\bar{\bar{x}}', \tag{41}$$

$$E(\omega z') = b\tau(\omega\bar{\bar{x}}'A' + \omega\bar{\bar{n}}'), \tag{42}$$

$$E(\bar{x}z') = b\tau^2(\bar{\bar{x}}\bar{\bar{x}}' + K_{\bar{x}})A' + b\tau^2\bar{\bar{x}}\bar{\bar{n}}', \tag{43}$$

etc.

Substituting the expected values of (38)-(43) into (37) yields

$$\begin{aligned} e = \frac{1}{\tau^2} tr\Big[& H[b^2\tau^2(AK_{\bar{x}}A' + K_{\bar{n}}) + (\sigma_b^2 + b^2)C_z]H' \\ & -b\tau^2 HAK_{\bar{x}} - b\tau^2 K_{\bar{x}}A'H' + \tau^2\omega\omega' - \tau^2\omega\bar{\bar{x}}' - \tau^2\bar{\bar{x}}\omega' \\ & +\tau^2(K_{\bar{x}} + \bar{\bar{x}}\bar{\bar{x}}')\Big]. \end{aligned} \tag{44}$$

Minimizing (44) with respect to ω requires that $\omega = \bar{\bar{x}}$ and minimizing with respect to H requires that

$$\begin{aligned} H &= \frac{K_{\bar{x}}}{b}A'(gC_z + K_{\bar{n}} + AK_{\bar{x}}A')^{-1} \\ &= \frac{1}{b}[A'(gC_z + K_{\bar{n}})^{-1}A + K_{\bar{x}}^{-1}]^{-1}A'(gC_z + K_{\bar{n}}^{-1}) \end{aligned} \tag{45}$$

where $g = (\sigma_b^2 + b^2)/\tau^2 b^2$.

This is the optimum discrete linear filter in the MSE sense. The optimum linear estimate of $\bar{x}$ is now

$$\begin{aligned} \hat{\bar{x}} &= H[z/\tau - b(A\bar{\bar{x}} + \bar{\bar{n}})] + \bar{\bar{x}} = H(z/\tau - b\bar{\bar{n}}) + (I-bHA)\bar{\bar{x}} \\ &= \frac{1}{b}[A'(gC_z + K_{\bar{n}})^{-1}A + K_{\bar{x}}^{-1}]^{-1}A'(gC_z + K_{\bar{n}})^{-1}(z/\tau - b\bar{\bar{n}}) \\ &\quad + [A'(gC_z + K_{\bar{n}})^{-1}A + K_{\bar{x}}^{-1}]^{-1} K_{\bar{x}}^{-1}\bar{\bar{x}}. \end{aligned} \tag{46}$$

If the mean $E(\hat{\bar{x}})$ of an estimate $\hat{\bar{x}}$ equals $\bar{x}$, the estimate $\hat{\bar{x}}$ of $\bar{x}$ is said to be unbiased; if not, the difference $E(\hat{\bar{x}})-\bar{x}$ is defined as the

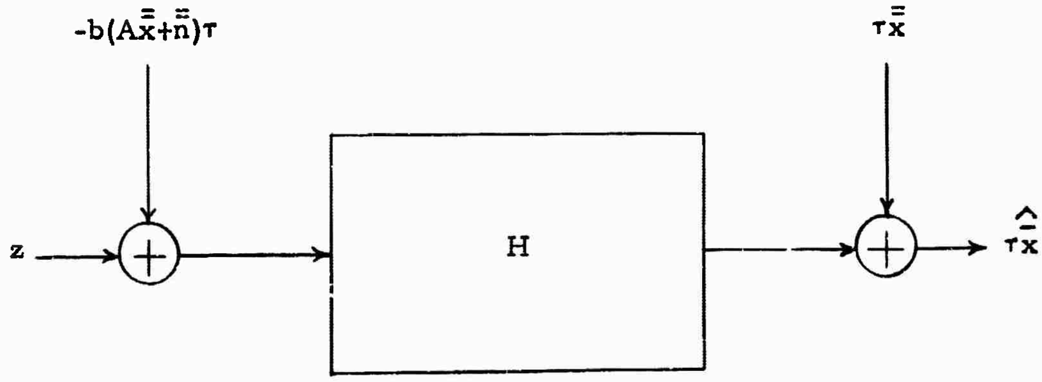

Figure 2. Minimum MSE estimation system.

bias of the estimate. In the absence of a priori information about the estimate, it is desirable that the bias of the estimate be small or zero. We will now check to see if the estimate $\hat{\bar{x}}_o$ is biased where $\bar{x}_o$ is some actual but unknown mean rate object vector. If $\hat{\bar{x}}_o$ is unbiased then $E(\hat{\bar{x}}_o) = E(\bar{x}_o) = \bar{x}_o$. For the case being considered

$$
\begin{aligned}
\text{bias} &= E[\,H(z - b\tau(A\bar{\bar{x}} + \bar{\bar{n}})) + \tau\bar{\bar{x}}\,] - \bar{x}_o\tau \\
&= H(bA\bar{x}_o\tau - b\tau A\bar{\bar{x}}) + \bar{\bar{x}}\tau - \bar{x}_o\tau \qquad (47) \\
&= (bHA - I)\,(\bar{x}_o - \bar{\bar{x}})\tau
\end{aligned}
$$

where I is the identity matrix. The estimate $\hat{\bar{x}}_o$ is unbiased if either

1) $\bar{x}_o = \bar{\bar{x}}$ or

2) $bHA = I$.

The second condition implies that $K_{\bar{x}}^{-1} = \underline{0}$ where $\underline{0}$ is a matrix of zeros.

The covariance matrix $K_{\bar{x}}$ is related to the a priori information about the object. If the elements of $K_{\bar{x}}$ are large (particularly the diagonal elements) the prior information about the signal is small. Hence, for large a priori uncertainty about the object, $K_{\bar{x}}^{-1} \approx \underline{0}$. By $K_{\bar{x}}^{-1} \approx \underline{0}$ we mean that the elements of $K_{\bar{x}}^{-1}$ are small in comparison with $A'(gC_z + K_{\bar{n}})^{-1}A$.

For large a priori uncertainity we can write H as follows:

$$H = \frac{1}{b}\,[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1}\,A'(gC_z + K_{\bar{n}})^{-1}. \tag{48}$$

Since $K_{\bar{x}}^{-1} \approx \underline{0}$ for large a priori uncertainty, the estimate $\hat{\bar{x}}_o$ for this case is unbiased.

To evaluate the optimum estimation or restoration procedure, we must find the MSE for the actual but unknown object vector $\bar{x}_o$. Assuming large a priori uncertainty we have for our minimum MSE estimate of $\bar{x}_o$

$$\hat{\bar{x}}_o = \frac{1}{b}\,[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1}\,A'(gC_z + K_{\bar{n}})^{-1}\,(z/\tau - b\bar{\bar{n}}). \tag{49}$$

Given that the object vector is $\bar{x}_o$, the MSE is given by

$$e = \mathrm{tr}\,E\,[(\hat{\bar{x}}_o - \bar{x}_o)(\hat{\bar{x}} - \bar{x}_o)']. \tag{50}$$

Substituting (49) into (50) yields

$$e = \mathrm{tr}\Big[[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1}\,A'[gC_z + K_{\bar{n}}]^{-1}[gC_z(\bar{x}_o) + K_{\bar{n}}] \cdot [gC_z + K_{\bar{n}}]^{-1}\,A[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1}\Big] \tag{51}$$

where $C_z(\bar{x}_o) = \tau\,(a_i\bar{x}_o + \bar{\bar{n}}_i)\,\delta_{ij}$.

Let us consider a simple example to investigate the weighting due to $\bar{x}_o$ (the actual object vector). Assume that we have a slit aperture, white noise, and that the object is made up of two point sources which are separated by the Rayleigh criterion distance. We will make k measurements at the peak of each point spread function. Also, let $D/\nu R$ equal one. See the Appendix for clarification of these assumptions. The 2k x 2 matrix A becomes

$$A = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ -k \\ \\ \\ -2k \end{matrix} .$$

Substituting this matrix into (51) and carrying out the indicated operations yields

$$e = \frac{1}{k}\left[\frac{\sigma_b^2 + b^2}{\tau b^2} (\bar{x}_{01} + \bar{x}_{02} + \bar{\bar{n}}_1 + \bar{\bar{n}}_2) + \sigma_{\bar{n}}^2\right] \tag{52}$$

where $k = 0, 1, 2, \ldots$ and the actual object vector is $\bar{x}_o = \begin{pmatrix}\bar{x}_{01} \\ \bar{x}_{02}\end{pmatrix}$. The weighting due to $\bar{x}_o$ is not significant if $\bar{\bar{n}}_1 + \bar{\bar{n}}_2 >> \bar{x}_{01} + \bar{x}_{02}$ (small signal-to-noise ration (SNR)), if τ becomes large, or if $\sigma_{\bar{n}}^2$ is large. Figure 3 shows how the error varies with τ for various signal values, a single noise value, and k = 1.

In general as τ becomes very large, $K_{\bar{n}} >> gC_z = \frac{(\sigma_b^2 + b^2)}{\tau b^2}(a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\delta_{ij}$. Hence, as τ becomes very large (51) reduces to

$$e = tr\,(A' K_{\bar{n}}^{-1} A)^{-1}. \tag{53}$$

For the special case of white noise (i. e., $K_{\bar{n}} = \sigma_{\bar{n}}^2 I$ where I is the identity matrix) we have

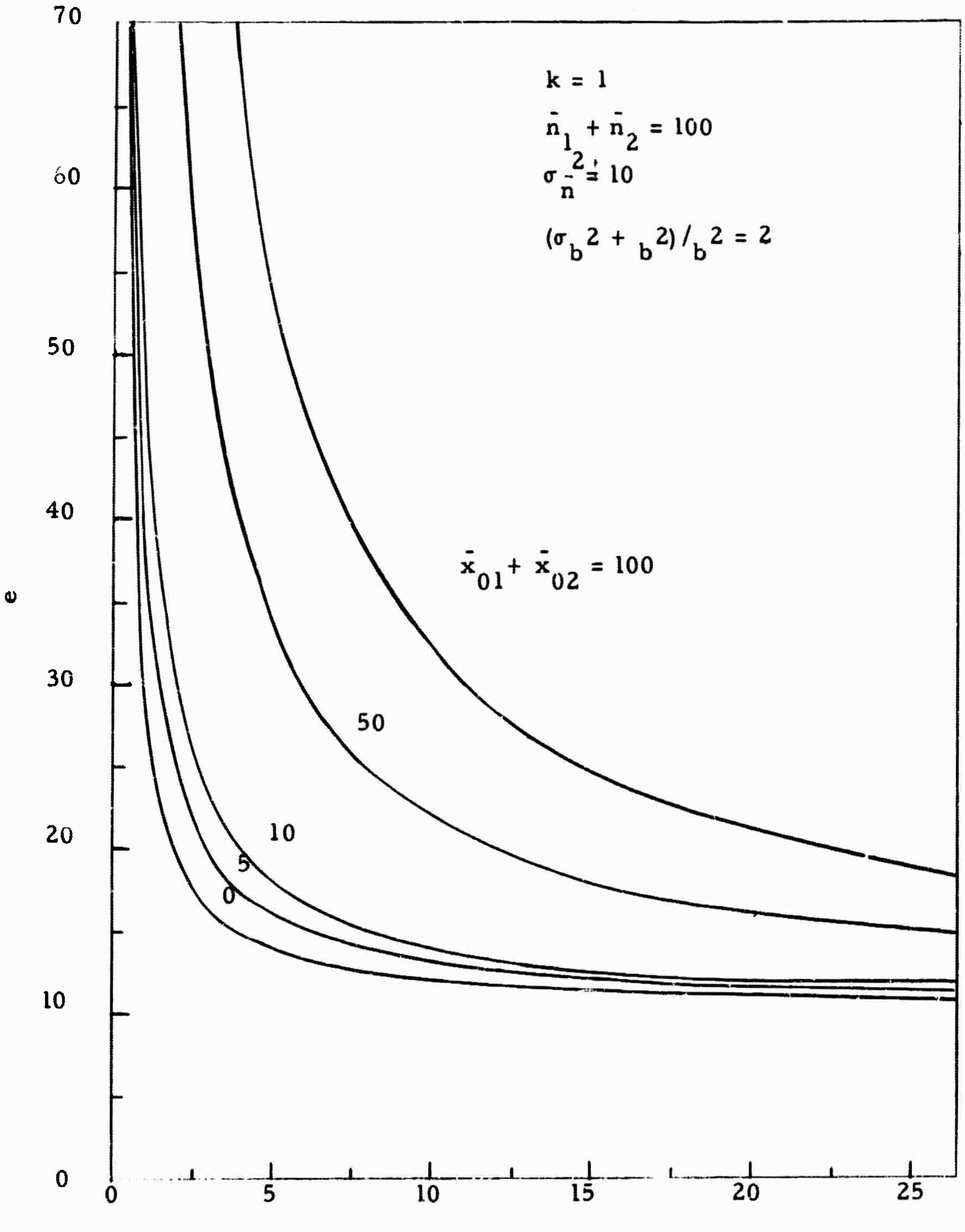

Figure 3. Error e versus observation time τ for the case of two measurements of the image of two point sources separated by the Rayleigh criterion distance.

$$e = \sigma_n^2 \operatorname{tr} (A'A)^{-1}. \tag{54}$$

The factor $(A'A)^{-1}$ in (54) can be thought of as an amplifier of the noise σ_n^2. The amplification increases with decreased aperture size and/or decreased spacing between the point sources. Let

$$G = \operatorname{tr} (A'A)^{-1}. \tag{55}$$

To investigate the nature of G let us consider the infinite slit aperture. The point spread function for this case is

$$\frac{D^2}{\nu^2 R^2} \; \frac{\sin^2 \pi [(D/\nu R)(\xi - h)]}{\pi^2 [(D/\nu R)(\xi - h)]^2} \tag{56}$$

where D is the width of the aperture, h is the distance of the point source from the origin, and R is the distance from the image plane and object plane to the aperture plane. For simplicity, consider an object consisting of two point sources (one at the origin and one at a distance h from the origin, see Figure 4). The A matrix becomes

$$A = \begin{bmatrix} (D^2/\nu^2 R^2) \operatorname{sinc}^2 [(D/\nu R)(\xi_1 - h)] & (D^2/\nu^2 R^2) \operatorname{sinc}^2 [D\xi_1/\nu R] \\ (D^2/\nu^2 R^2) \operatorname{sinc}^2 [(D/\nu R)(\xi_2 - h)] & (D^2/\nu^2 R^2) \operatorname{sinc}^2 [D\xi_2/\nu R] \end{bmatrix} \tag{57}$$

where ξ_1 and ξ_2 are the measurement positions in the image plane

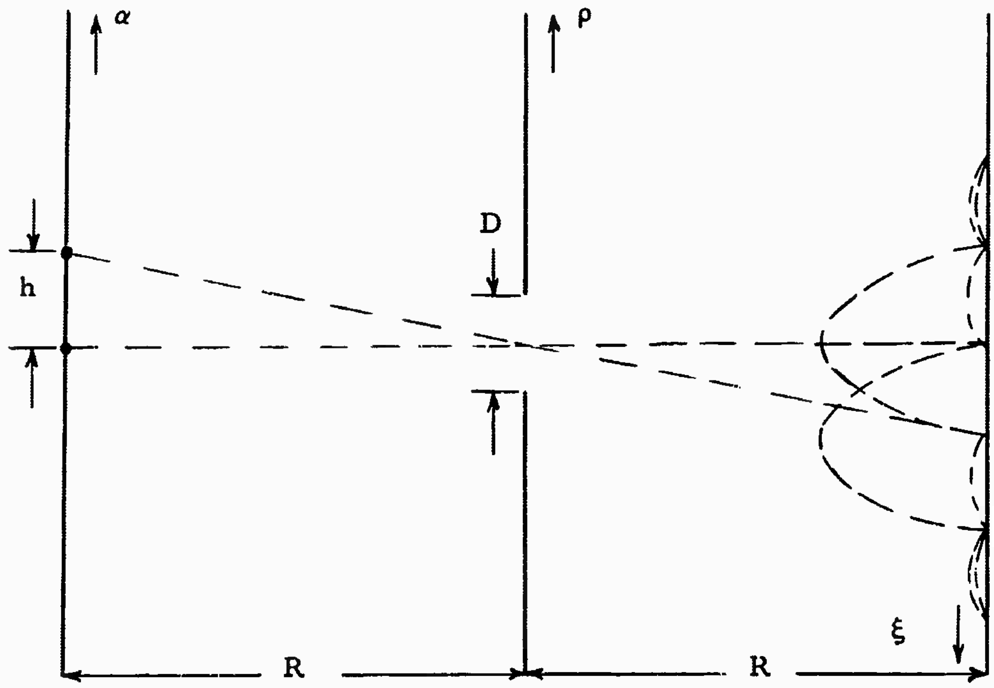

Figure 4. Optical system configuration for an infinite slit aperture and two point sources.

and sinc $x = \sin \pi x / \pi x$. If we make one measurement at the peak of each point spread function and let νR equal one, we have

$$A = \begin{bmatrix} D^2 & D^2 \operatorname{sinc}^2(Dh) \\ D^2 \operatorname{sinc}^2(Dh) & D^2 \end{bmatrix} . \tag{52}$$

The amplification factor G then becomes

$$G = \operatorname{tr} (A'A)^{-1} = \frac{2\,[1 + \operatorname{sinc}^4 (Dh)]}{D^4\,[1 - \operatorname{sinc}^4 (Dh)]^2} . \tag{53}$$

Figure 5 shows how the amplification factor G varies with aperture width D and separation of point sources h. The abrupt increase of $\log_{10} G$ occurs when the size of the object (separation of the two point sources) becomes approximately the size of the point spread function (see Harris and Rushforth (1966)).

Special Cases

Because of the complexity of the general estimate in (46) we will consider various special cases in order to gain a better understanding of the estimation procedure.

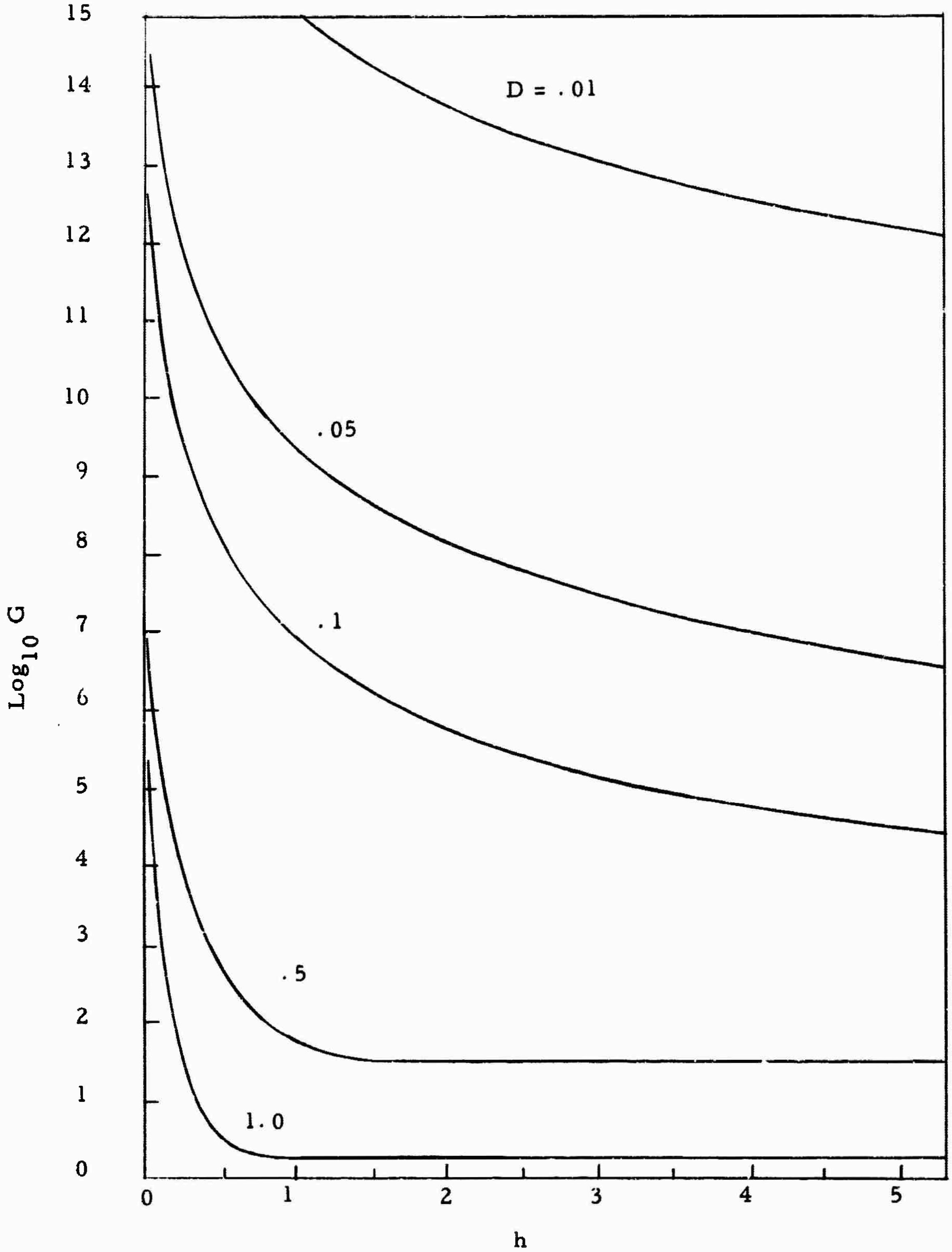

Figure 5. The logarithm of the amplification G versus point source separation h for several values of aperture width D.

<u>Prior information dependence</u>

Let $C_z = \sigma_z^2 I$, $K_{\bar{x}} = \sigma_{\bar{x}}^2$, and $K_{\bar{n}} = \sigma_{\bar{n}}^2 I$. Our estimate of $\bar{x}$ then becomes

$$\hat{\bar{x}} = \frac{1}{b\tau}\left[\frac{A'A}{g\sigma_z^2 + \sigma_{\bar{n}}^2} + \frac{I}{\sigma_{\bar{x}}^2}\right]^{-1} \frac{A'(z-b\tau\bar{\bar{n}})}{g\sigma_z^2 + \sigma_{\bar{n}}^2}$$

$$+\left[\frac{A'A}{g\sigma_z^2 + \sigma_{\bar{n}}^2} + \frac{I}{\sigma_{\bar{x}}^2}\right]^{-1} \frac{\bar{\bar{x}}}{\sigma_{\bar{x}}^2}. \tag{60}$$

For large a priori uncertainty, $\sigma_{\bar{x}}^2 >> g\,\sigma_z^2 + \sigma_{\bar{n}}^2$, we have

$$\hat{\bar{x}} \approx \frac{1}{\tau b}(A'A)^{-1}A'(z-b\tau\bar{\bar{n}}) \tag{61}$$

which indicates that we ignore the a priori mean $\bar{\bar{x}}$. For large prior information, $g\,\sigma_z^2 + \sigma_{\bar{n}}^2 >> \sigma_{\bar{x}}^2$, we have

$$\hat{\bar{x}} \approx \bar{\bar{x}} \tag{62}$$

which means that our prior mean is very reliable and that we learn very little from our experiment.

Large a priori uncertainty

Consider the case of large a priori uncertainty (i.e., $K_{\bar{x}}^{-1} \approx \underline{0}$). Our estimate of $\bar{x}$ becomes

$$\hat{\bar{x}} = \frac{1}{b\tau} [A'(gC_z + K_{\bar{n}})^{-1}A]^{-1} A'(gC_z + K_{\bar{n}})^{-1}(z - b\tau\bar{\bar{n}}) \tag{63}$$

and the MSE is

$$e = tr\Big[[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1} A'(gC_z + K_{\bar{n}})^{-1}(gC_z(\bar{x}_o) + K_{\bar{n}}) \tag{64}$$
$$\cdot (gC_z + K_{\bar{n}})^{-1} A[A'(gC_z + K_{\bar{n}})^{-1}A]^{-1} \Big].$$

Matrix A with inverse. If the system matrix A is square and invertible (63) reduces to

$$\hat{\bar{x}} = A^{-1}(z/b\tau - \bar{\bar{n}}) \tag{65}$$

and (64) reduces to

$$e = tr\,[A^{-1}(gC_z(\bar{x}_o) + K_{\bar{n}})A'^{-1}]. \tag{66}$$

The above estimate of $\bar{x}$ is an intuitive estimate since all we do is divide out the multiplicative effect, subtract the noise, and then pass this result through an inverse filter. This procedure is illustrated in Figure 6.

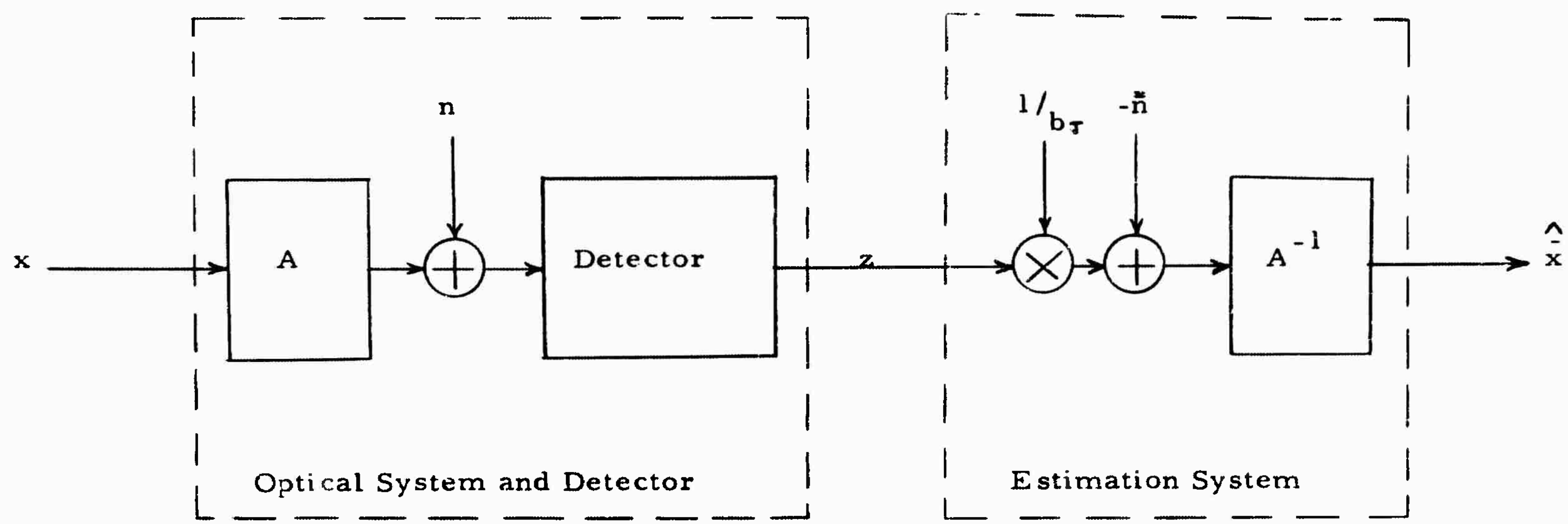

Figure 6. Diagram of the optical system, detector, and estimation system for cases where the optical system matrix A is square and invertible.

<u>Small signal-to-noise ratio.</u> For small signal-to-noise ratio (SNR) (i.e., $a_i \bar{\bar{x}} << \bar{\bar{n}}_i$) the linear filter H becomes

$$H = \frac{1}{b} [A'(g\tau \bar{\bar{n}}_i \delta_{ij} + K_{\bar{n}})^{-1} A]^{-1} A' (g\tau \bar{\bar{n}}_i \delta_{ij} + K_{\bar{n}})^{-1} \qquad (67)$$

and the MSE becomes

$$e = tr[A'(g\tau \bar{\bar{n}}_i \delta_{ij} + K_{\bar{n}})^{-1} A]^{-1}. \qquad (68)$$

The MSE error becomes independent of $\bar{x}_o$ for small SNR (see Figure 3).

If the noise is independent of i (uniform noise) then $\bar{\bar{n}}_i \delta_{ij} = NI$ where $\bar{\bar{n}}_i = N$ for all i. For uniform, white noise we have

$$H = \frac{1}{b} (A'A)^{-1} A' \qquad (69)$$

and

$$e = (g\tau N + \sigma_{\bar{n}}^2) \; tr \; (A'A)^{-1}. \qquad (70)$$

<u>Perfect detector.</u> For a perfect detector, B in (17) is a constant and hence each photon gives rise to exactly the same number, b, of photoelectrons (i.e., $\sigma_b^2 = 0$). For this case

$$H = \frac{1}{b} [A'(\tau(a_i \bar{\bar{x}} + \bar{\bar{n}}_i)\delta_{ij} + \tau^2 K_{\bar{n}})^{-1} A]^{-1} A'(\tau(a_i \bar{\bar{x}} + \bar{\bar{n}}_i)\delta_{ij} + \tau^2 K_{\bar{n}})^{-1} \qquad (71)$$

and

$$e = (1/\tau^2) \mathrm{tr}\Big[[A'(\tau(a_i\bar{\bar{x}} + \bar{\bar{n}}_i)\delta_{ij} + \tau^2 K_{\bar{n}})^{-1}A]^{-1}A'[\tau(a_i\bar{\bar{x}}+\bar{\bar{n}}_i)\delta_{ij}+\tau^2 K_{\bar{n}}]^{-1}$$

$$\cdot[\tau(a_i\bar{x}_o+\bar{\bar{n}}_i)\delta_{ij}+\tau^2 K_{\bar{n}}][\tau(a_i\bar{\bar{x}}+\bar{\bar{n}}_i)\delta_{ij}+\tau^2 K_{\bar{n}}]^{-1}A[A'(\tau(a_i\bar{\bar{x}}+\bar{\bar{n}}_i)\delta_{ij}+\tau^2 K_{\bar{n}})^{-1}A]^{-1}\Big], \quad (72)$$

Except for the constant b in the expression for H these are the same results that one would obtain if he were to count the photons incident upon the image plane and in turn find the minimum MSE estimate of the mean rate of signal photons emitted from the optical object (i. e., minimum MSE for no multiplicative or detector noise).

<u>Large additive noise covariance matrix.</u> For large $K_{\bar{n}}$ (i. e., $K_{\bar{n}} >> gC_z$) we can write

$$H = \frac{1}{b} (A'K_{\bar{n}}^{-1}A)^{-1}A'K_{\bar{n}}^{-1} \quad (73)$$

and

$$e = \mathrm{tr}\,(A'K_{\bar{n}}^{-1}A)^{-1}. \quad (74)$$

For this assumption the minimum MSE becomes independent of the multiplicative noise.

Photon-electron converter

The photon-electron converter is a degenerate case of the photomultiplier tube case in which we consider only its first stage. This is the detector model that Helstrom (1964) uses in some of his work on optical signal detection. We are assuming that the stream of photons y_i that impinge upon the i^{th} cell has a conditional Poisson distribution with a conditional mean rate of $\bar{y}_i = (a_i\bar{x} + \bar{n}_i)$; hence, $E(y_i/\bar{x},\bar{n}) = \text{var}(y_i/\bar{x},\bar{n}) = \bar{y}_i$. In the photon-electron converter each photon gives rise to an emitted electron with probability η. The stream of photoelectrons z_i being emitted from the i^{th} cell, therefore, has a conditional Poisson distribution with a conditional mean rate of $\eta(a_i\bar{x} + \bar{n}_i) = \eta\bar{y}_i$ since a Poisson process is preserved under random selection (Parzen, 1962). Hence, $\text{var}(z_i/\bar{x},\bar{n}) = E(z_i/\bar{x},\bar{n}) = \eta\bar{y}_i$. The incoming photons and emitted photoelectrons are related by their means: $\eta E(y_i) = \eta E(a_i x + n_i) = E(z_i)$. Since z is a Poisson process, $\text{var}(z_i/\bar{x},\bar{n}) = E(z_i/\bar{x},\bar{n}) = \eta E(y_i/\bar{x},\bar{n})$. From previous results and using the Poisson distribution properties, we have $E(z_i/\bar{x},\bar{n}) = bE(y_i/\bar{x},\bar{n})$ and $\text{var}(z_i/\bar{x},\bar{n}) = (\sigma_b^2 + b^2)E(y_i/\bar{x},\bar{n})$. Hence, we can use the previously obtained results and apply them to the photon-electron converter case. This is done by replacing b by η and replacing $\sigma_b^2 + b^2$ by η. Hence, for the photon-electron converter (assuming large a priori uncertainty) we have

$$H = \frac{1}{\eta}\,[A'(C_z/\eta + \tau^2 K_n \quad {}^1 A]^{-1} A'(C_z/\eta + \tau^2 K_{\bar{n}})^{-1} \tag{75}$$

and

$$e = (1/\tau^2)\,\mathrm{tr}\Big[[A'(C_z/\eta + \tau^2 K_{\bar{n}})^{-1} A]^{-1} A'(C_z/\eta + \tau^2 K_{\bar{n}})^{-1} (C_z(\bar{x}_o)/\eta + \tau^2 K_{\bar{n}})$$

$$\cdot (C_z/\eta + \tau^2 K_{\bar{n}})^{-1} A [A'(C_z/\eta + \tau^2 K_{\bar{n}})^{-1} A]^{-1}\Big]. \tag{76}$$

As η approaches unity these results approach those for the case of no multiplicative noise. The case of $\eta = 1$ (perfect detector) implies that every photon that impinges upon the detector causes an electron to be emitted with probability one. As η approaches zero, the error becomes extremely large.

<u>Estimation of mean rate $\bar{x}$ from observations of $\bar{z}$</u>

In many realistic cases where we have large number of counts, equipment capabilities allow us to measure only intensity or mean rate of photoelectrons z. (We are assuming that the sample mean equals $\bar{z}$ by the law of large numbers.) For this case we want to find the linear estimate $\hat{\bar{x}}$ of $\bar{x}$ from observations of $\bar{z}$ which will minimize the mean-square error; that is, find $\hat{\bar{x}}$ to minimize

$$e = \mathrm{tr}\, E[(\hat{\bar{x}} - \bar{x})(\hat{\bar{x}} - \bar{x})']. \tag{77}$$

From the estimate $\hat{\bar{x}}$ we can obtain an estimate of the intensity vector by multiplying $\hat{\bar{x}}$ by $h\nu$. From previous results $\bar{z} = b\bar{y}$. We also have

$$\bar{y} = A\bar{x} + \bar{n}, \tag{78}$$

$$E(\bar{y}) = \bar{\bar{y}} = A\bar{\bar{x}} + \bar{\bar{n}}, \tag{79}$$

$$E(\bar{z}) = \bar{\bar{z}} = b(A\bar{\bar{x}} + \bar{\bar{n}}), \tag{80}$$

$$\mathrm{Cov}(\bar{y}_i, \bar{y}_j) = (AK_{\bar{x}}A' + K_{\bar{n}}), \tag{81}$$

$$\mathrm{Cov}(\bar{z}_i, \bar{z}_j) = b^2(AK_{\bar{x}}A' + K_{\bar{n}}). \tag{82}$$

For the linear estimate of $\bar{x}$ we can write

$$\hat{\bar{x}} = H[\bar{z} - b(A\bar{\bar{x}} + \bar{\bar{n}})] + \omega. \tag{83}$$

Substituting this expression of $\hat{\bar{x}}$ into (77) and carrying out the expectation operation, we obtain

$$H = \frac{1}{b} K_{\bar{x}}A'(K_{\bar{n}} + AK_{\bar{x}}A')^{-1} = \frac{1}{b}(A'K_{\bar{n}}^{-1}A + K_{\bar{x}}^{-1})^{-1}A'K_{\bar{n}}^{-1}. \tag{84}$$

Hence, our estimate of $\bar{x}$ is

$$\hat{\bar{x}} = \frac{1}{b}(A'K_{\bar{n}}^{-1}A + K_{\bar{x}}^{-1})^{-1}A'K_{\bar{n}}^{-1}(\bar{z} - b\bar{\bar{n}}) + (A'K_{\bar{n}}^{-1}A + K_{\bar{x}}^{-1})^{-1}K_{\bar{x}}^{-1}\bar{\bar{x}}. \tag{85}$$

For $K_{\bar{x}}^{-1} \approx \underline{0}$ (large a priori uncertainty) the minimum MSE is

$$e = \mathrm{tr}(A'K_{\bar{n}}^{-1}A)^{-1}. \tag{86}$$

This is the same result that was obtained in (53) where we let τ become large. Except for a constant b, these results are the same as those for the case of measuring image intensity, and from this estimating the intensity of the object vector in the absence of any multiplicative noise (Rushforth, 1965).

Estimation of mean rate $\bar{y}$ from z

The purpose of considering the estimate of $\bar{y}$ from measurements of z is to use the results in a later section for the detection of unknown signals.

We want to find the linear estimate $\hat{\bar{y}}$ of $\bar{y}$ which will minimize the mean-square error

$$e = \mathrm{tr}E \left[(\hat{\bar{y}}-\bar{y})(\hat{\bar{y}}-\bar{y})'\right] \tag{87}$$

where $\hat{\bar{y}} = H(z/\tau - b\bar{\bar{y}}) + \bar{\bar{y}}$. When $\hat{\bar{y}}$ is substituted into (87) and the expectation operation carried out, we obtain

$$e = \mathrm{tr}\left[H(b^2 K_{\bar{y}} + (\sigma_b^2 + b^2) C_z/\tau^2) H' - bHK_{\bar{y}} - bK_{\bar{y}}H' + K_{\bar{y}}\right] . \tag{88}$$

The H that minimizes (88) is

$$H = \frac{K_{\bar{y}}}{b} (g C_z + K_{\bar{y}})^{-1} \tag{89}$$

where $g = (\sigma_b^2 + b^2)/\tau^2 b^2$. Hence, the minimum MSE estimate of

$\bar{y}$ is

$$\hat{\bar{\bar{y}}} = b\tau K_{\bar{y}}[(\sigma_b^2 + b^2) C_z + b^2\tau^2 K_{\bar{y}}]^{-1}(z - b\tau\bar{\bar{y}}) + \bar{\bar{y}}. \tag{90}$$

For the photon-electron converter $b = \eta$ and $(\sigma_b^2 + b^2) = \eta$; hence,

$$\hat{\bar{\bar{y}}} = \tau K_{\bar{y}} (C_z + \eta\tau^2 K_{\bar{y}})^{-1} (z - \eta\tau\bar{\bar{y}}) + \bar{\bar{y}}. \tag{91}$$

For $K_{\bar{y}} = \sigma_{\bar{y}_i}^2 \delta_{ij}$ we can write the estimate of $\bar{y}_i$ as

$$\hat{\bar{\bar{y}}}_i = \frac{\tau\sigma_{\bar{y}_i}^2 (z_i - \eta\tau\bar{\bar{y}}_i)}{\tau\bar{\bar{y}}_i + \tau^2\sigma_{\bar{y}_i}^2} + \bar{\bar{y}}_i. \tag{92}$$

For the statistical model we have been assuming $\bar{\bar{y}}_i = u_i/\alpha_i$ and $\sigma_{\bar{y}_i}^2 = u_i/\alpha_i^2$ (see (7) and (10)); hence, after substituting these values into (91) and rearranging we obtain

$$\hat{\bar{y}}_i = \frac{z_i + u_i}{\eta\tau + \alpha_i}. \tag{93}$$

Optimum Sampling Scheme

When the additive noise covariance matrix is large or for large τ the error expression reduces to that of the additive-noise-only case:

$$e = \mathrm{tr}\,(A'K_{\bar{n}}^{-1}A)^{-1}. \tag{94}$$

The problem we now face is to determine an optimum sampling procedure which will minimize the above error expression. We are able to vary the matrix A by varying our sampling positions. Thus, for some optimum sampling positions the MSE will be minimized. The optimum sampling procedure for this case also applies to the case of small SNR with uniform noise.

White noise

For white noise $K_{\bar{n}} = \sigma_{\bar{n}}^2 I$ and

$$e = \sigma_{\bar{n}}^2\,\mathrm{tr}\,(A'A)^{-1}. \tag{95}$$

Single point source. Consider an infinite slit aperture which is equivalent to reducing the optical problem to one dimension. The point spread function for this case is

$$\left(D^2/\nu^2R^2\right)\mathrm{sinc}^2\left[\frac{D}{\nu R}(\xi - h)\right] \tag{96}$$

(see Figure 4 and the Appendix). Consider the normalized case where the point spread function is

$$\text{sinc}^2(\xi-h) = \frac{\sin^2[\pi(\xi-h)]}{\pi^2 (\xi-h)^2}. \tag{97}$$

Our objective is to minimize $tr(A'A)^{-1}$ by properly locating our samples. For a single point source located at the origin in the object plane and making k samples in the image plane, the matrix A becomes a vector of the form

$$A = \begin{bmatrix} \text{sinc}^2 \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \text{sinc}^2 \xi_k \end{bmatrix} \tag{98}$$

where $\xi_1, \xi_2, \ldots, \xi_k$ represent the positions in the image plane where the samples are taken. $A'A$ is then $\sum_{i=1}^{k} \text{sinc}^4 \xi_i$ and $tr(A'A)^{-1} = (A'A)^{-1} = 1/\sum_{i=1}^{k} \text{sinc}^4 \xi_i$. Since we are free to make the measurements anywhere in the image plane, we want to make the measurements such that $\sum_{i=1}^{k} \text{sinc}^4 \xi_i$ is a maximum. It is obvious that we want to make all of the measurements at the origin (i.e., $\xi_i = 0$, $i = 1, 2, \ldots, k$). Hence, $tr(A'A)^{-1} = 1/k$. Then for the minimum error expression we have

$$e_{min} = \sigma_n^2 \, tr\,(A'A)^{-1} = \sigma_n^2/k. \tag{99}$$

The error is inversely proportional to the number of measurements.

As k approaches infinity, e_{min} approaches zero. The above sampling scheme applies to all well-behaved point spread functions for the case of white noise.

m point sources and ℓ measurements. Assume that our object consists of m point sources and that they are separated by the Rayleigh criterion distance. By the Rayleigh criterion distance, we mean that the maximum of the diffraction pattern of one point source overlaps the first minima of the diffraction patterns due to adjacent point sources. We want to minimize $e = \sigma_{\bar{n}}^2 \, tr \, (A'A)^{-1}$ when we have m point sources and make ℓ measurements.

Slit aperture. For a slit aperture the general expression for the point spread function is

$$\frac{D^2}{\nu^2 R^2} \, \text{sinc}^2 \left[\frac{D}{\nu R} (\xi - j) \right]. \tag{100}$$

To simplify the mathematics we assume that the point spread function is normalized with $D/\nu R$ set equal to unity. The point spread function is then $\text{sinc}^2(\xi - j)$ where $j = \ldots -3, -2, -1, 0, 1, 2, 3, \ldots$. The normalized system matrix for the slit aperture case becomes

$$A = \begin{bmatrix} \text{sinc}^2(\xi_1 - 1) & \ldots & \text{sinc}^2(\xi_1 - m) \\ \vdots & & \vdots \\ \text{sinc}^2(\xi_\ell - 1) & \ldots & \text{sinc}^2(\xi_\ell - m) \end{bmatrix}. \tag{101}$$

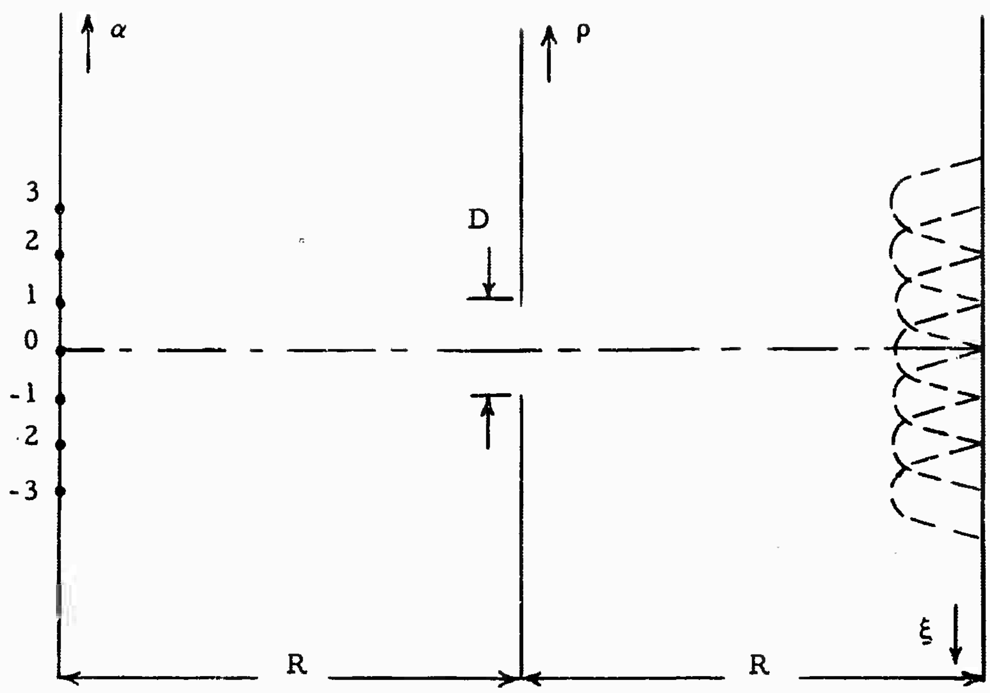

Figure 7. Optical system configuration for an infinite slit aperture and point sources separated by the Rayleigh criterion distance.

The optimum sampling procedure for this case (see Appendix) is to sample ℓ/m times at the peak of each point spread function. If ℓ/m is not an integer (i.e., $\ell/m = k + n/m$) the optimum procedure is to make k measurements at the peak of each point spread function and one additional measurement at the peak of any n of the m point spread tions. If $\ell/m = k$ (integer)

$$e = \sigma_{\bar{n}}^2 \operatorname{tr}(A'A)^{-1} = \sigma_{\bar{n}}^2 m/k \tag{102}$$

where m is the number of point sources and k is the number of measurements per point source.

Rectangular aperture. For the retangular aperture the general expression for the point spread function is

$$\frac{a^2b^2}{\nu^2R^2} \operatorname{sinc}^2\left[\frac{a}{\nu R}(\xi - j)\right] \operatorname{sinc}^2\left[\frac{a}{\nu R}(\zeta - i)\right]. \tag{103}$$

Normalizing this point spread function by letting $a/\nu R = b/\nu R = 1$ yields $\operatorname{sinc}^2(\xi - j) \operatorname{sinc}^2(\zeta - i)$ where $i, j = \ldots -2, -1, 0, 1, 2, 3, \ldots$.

The optimum sampling procedure and MSE for this case are the same as for the slit aperture case.

For point source spacings less than the Raleigh criterion distance the optimum sampling procedure becomes very complicated and will not be considered here. Harris and Rushforth (1966) work out some specific examples of this case.

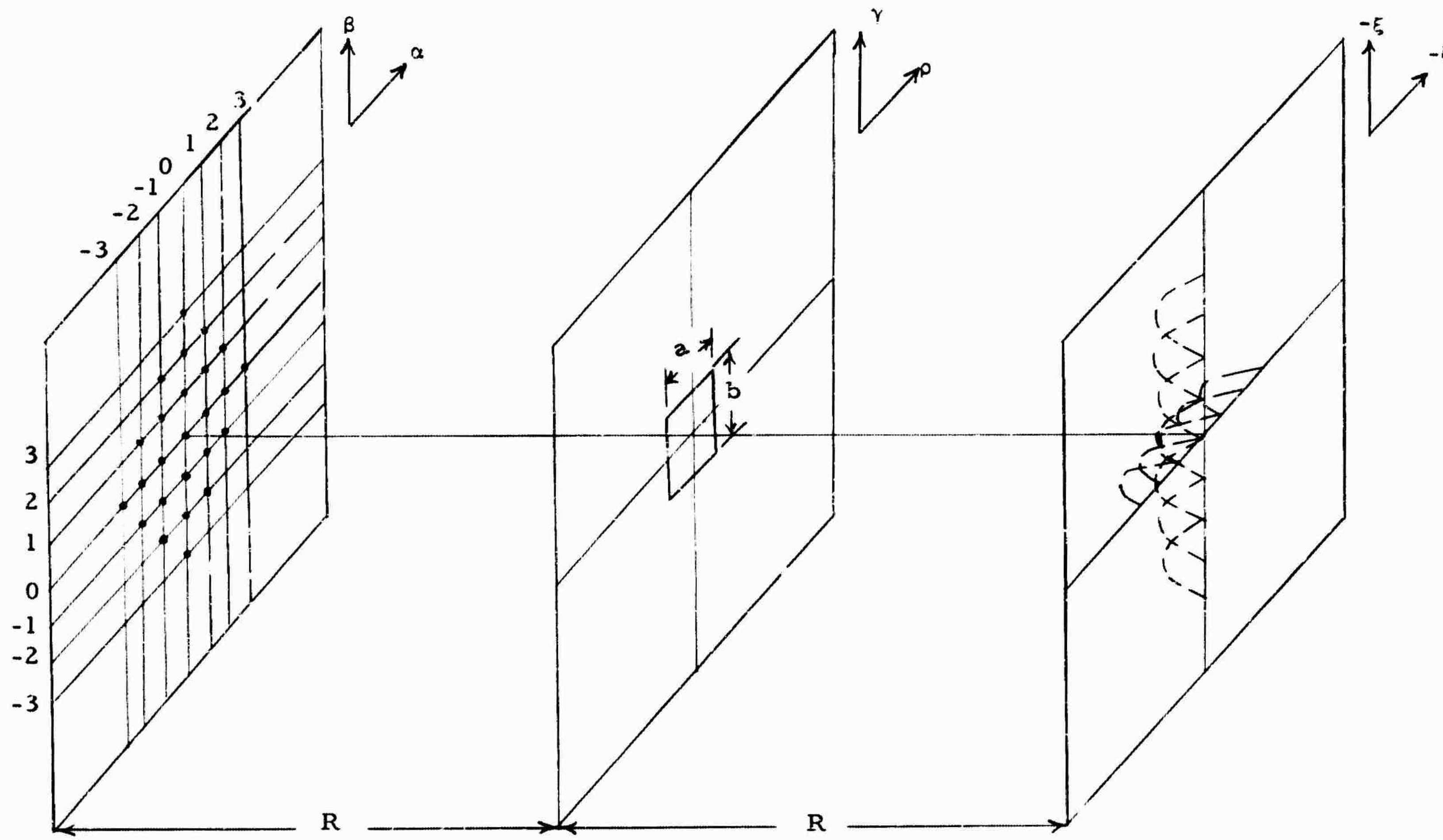

Figure 8. Optical system configuration for a rectangular aperture and point sources separated by the Rayleigh criterion distance.

Colored noise

Because of the complexity of the colored noise case only a single point source and two measurements in the image plane will be considered. We will assume a noise covariance matrix of the form

$$K_{\bar{n}} = \sigma_{\bar{n}}^2 \begin{pmatrix} 1 & e^{-c|d|} \\ e^{-c|d|} & 1 \end{pmatrix} \tag{104}$$

where $|d|$ is the distance between the measurement positions (i.e., $|d| = |\xi_1 - \xi_2|$) and c is a correlation constant. The inverse of the $K_{\bar{n}}$ matrix is

$$K_{\bar{n}}^{-1} = \frac{1}{\sigma_{\bar{n}}^2} \begin{pmatrix} \dfrac{\begin{matrix} 1 & -e^{-c|d|} \\ -e^{-c|d|} & 1 \end{matrix}}{1-e^{-2c|d|}} \end{pmatrix} . \tag{105}$$

For the present case of two measurements and one point source the system matrix A is a two-element vector. If we assume that the point source is located at the origin of the object plane, the normalized system matrix A becomes

$$A = \begin{pmatrix} \operatorname{sinc}^2 \xi_1 \\ \operatorname{sinc}^2 \xi_2 \end{pmatrix} \tag{106}$$

where ξ_1 and ξ_2 are the measurement positions in the image plane.

As c approaches infinity, $K_{\bar{n}}$ approaches $\sigma_{\bar{n}}^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \sigma_{\bar{n}}^2 I$ which implies that the additive noise is uncorrelated at any two positions or that we have white noise. As c approaches zero, $K_{\bar{n}}$ approaches

$\sigma_n^2 \binom{11}{11}$ which implies perfect correlation of the additive noise at any two positions.

We need to solve the error expression to find the measurement positions necessary to minimize it. The error expression (with $\sigma_n^2 = 1$) is

$$e = \mathrm{tr}\,(A'K_n^{-1}A)^{-1} = \frac{1 - e^{-2c|d|}}{\mathrm{sinc}^4\xi_1 - 2\,\mathrm{sinc}^2\xi_1\,\mathrm{sinc}^2\xi_2\,e^{-c|d|} + \mathrm{sinc}^4\xi_2} \tag{107}$$

There are four cases that can possibly minimize the error expressions of (107). Each of these cases need to be investigated in order to determine the minimum error conditions. These four cases are listed below.

Case 1. Let $\xi_1 = \xi_2 = 0$. This implies* that $e = 1$.

---

*Consider the general aperture case with one point source and two measurements and an arbitrary covariance matrix of the form

$$K_n = \sigma_n^2 \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$$

and let

$$A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Evaluate e for $a_1 = a_2$ and $a = 1$. In general

$$e = \mathrm{tr}\,(A'K_n^{-1}A)^{-1} = \frac{\sigma_n^2(1-a^2)}{a_1^2 - 2a_1a_2a + a_2^2}.$$

For $a_1 = a_2$,

$$e = \frac{\sigma_n^2(1+a)}{2a_1^2}.$$

Now for $a = 1$,

$$e = \sigma_n^2/a_1^2.$$

Case 2. Let $\xi_1 = -\xi_2$ which implies that

$$e = \frac{1 + e^{-2c|\xi_1|}}{2 \operatorname{sinc}^4 \xi_1} . \tag{108}$$

Case 3. Let $\xi_2 = 0$ and vary ξ_1 which implies that

$$e = \frac{1 - e^{-2c|\xi_1|}}{1 - 2 \operatorname{sinc}^2 \xi_1 \, e^{-c|\xi_1|}} . \tag{109}$$

Case 4. Let ξ_2 approach infinity and vary ξ_1. This implies that e has a minimum of unity for $\xi_1 = 0$.

Cases 2 and 3 need to be investigated and compared with cases 1 and 4. Cases 2 and 3 were programmed and the error determined for varying ξ_1 and also for different values of c. Cases 2 and 3 were found to be always as good or better than cases 1 and 4. The value of c determines which of cases 2 and 3 gives the minimum error. For c less than about 1.5, case 2 gives the minimum error; and for c larger than 1.5, case 3 gives the minimum error. Figure 9 shows the minimum error possible versus c. Figures 10 and 11 show the error obtained for varying ξ_1 for cases 2 and 3.

In Figure 10, for case 3, note that as the value of c becomes small the error becomes small and in fact approaches zero as c approaches zero. When c is zero the noise is constant from one position to another. Hence, if the noise plus signal is measured at the peak of

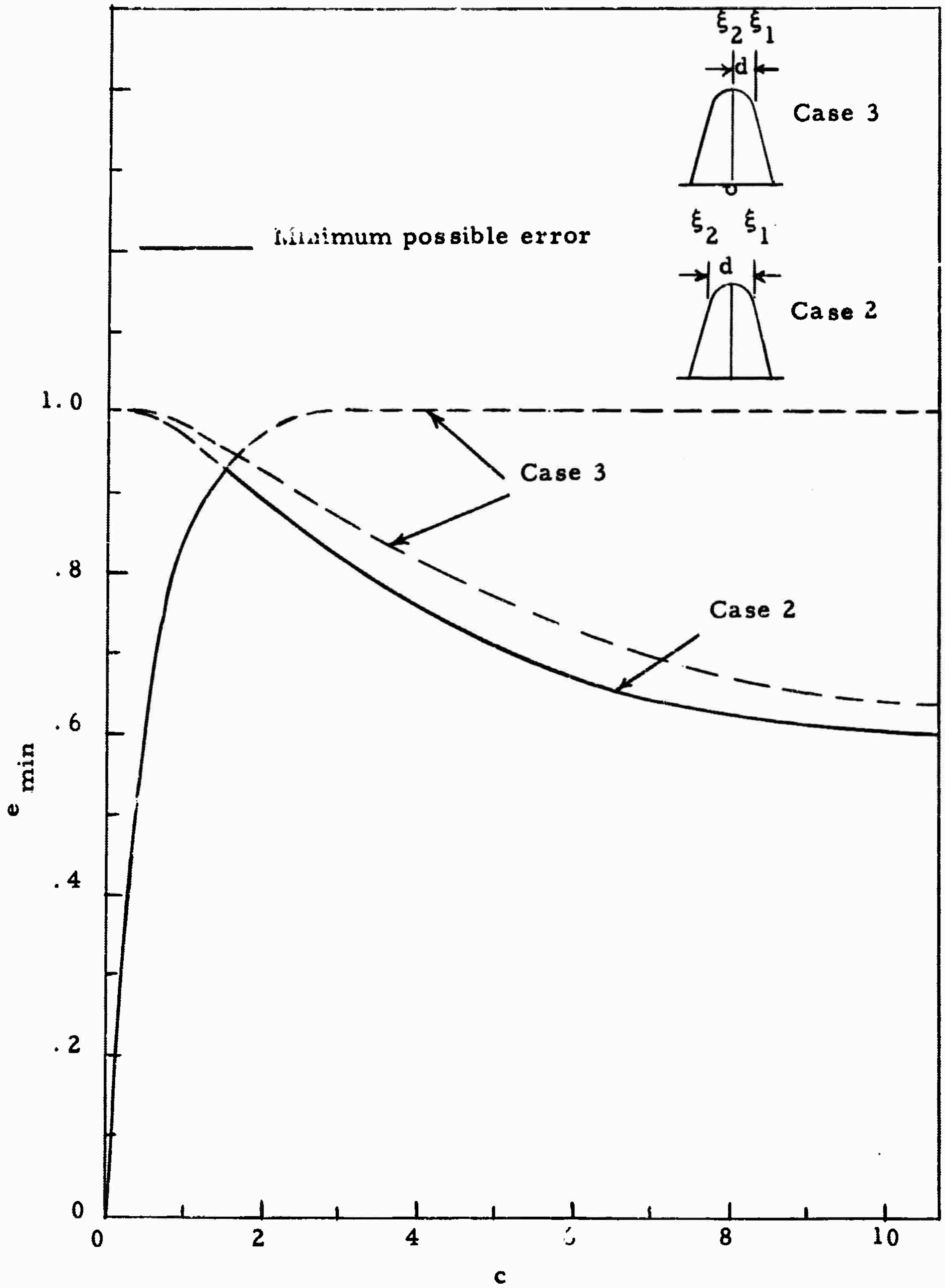

Figure 9. Minimum possible error versus correlation constant c for the case of colored noise, a single point source, and two image plane measurements.

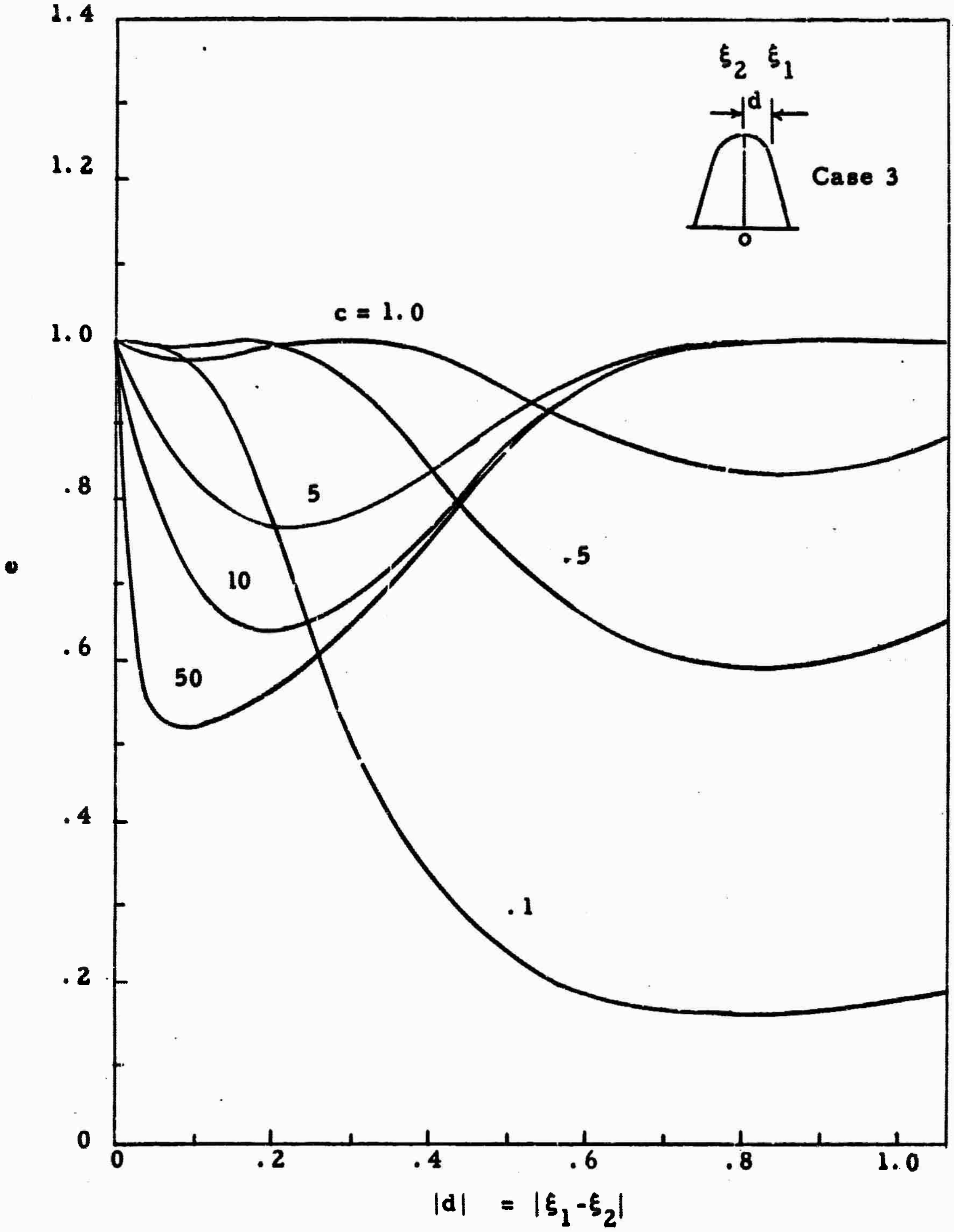

Figure 10. MSE e versus distance |d| between the two measurement positions of case 3.

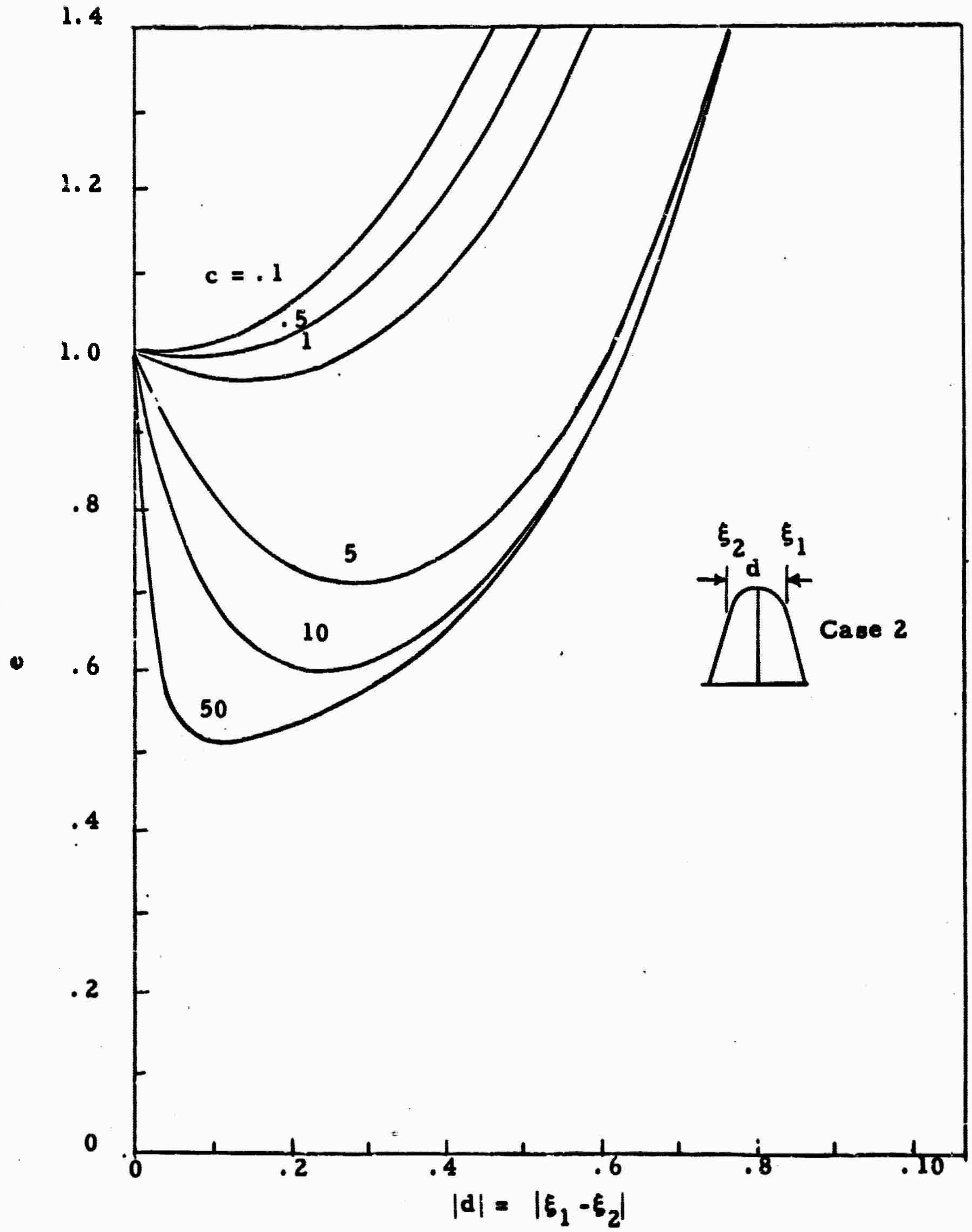

Figure 11. MSE e versus distance |d| between the two measurement positions of case 2.

the point spread function and the noise is measured at the null positions, the noise can be subtracted from the measurement at the peak and only the true signal will remain since the noise for both positions is the same. Note, however, that in this latter case the multiplicative noise must be considered.

OTHER ESTIMATES

Introduction

In this section we will investigate three other types of estimates: the Bayes' estimate, the maximum likehihood estimate, and the maximum a posteriori estimate. Throughout this section and the remaining sections we will consider only the photon-electron converter detector with a quantum efficiency of η. The incoming photons due to both known signals and known noise will be assumed to be Poisson distributed. Unless otherwise stated, whenever an a priori density function is needed for the mean rate of incident photons we will assume it to be a gamma distribution with known parameters (see Statistical Model section). Our motivation for using this distribution is due to its unique characteristic of generating another gamma distribution as an a posteriori density function when combined with a conditional Poisson distribution in the Bayes' formula. The gamma distribution is also physically reasonable (Goodman, 1965; Farrell, 1966).

We have defined $\bar{y}_i$ to be the mean number of photons that are incident upon the i^{th} cell of the image plane per unit time and $\eta\bar{y}_i$ as the mean number of photoelectrons that are emitted from the light sensitive i^{th} cell per unit time.

Since the stream of incoming photons are Poisson distributed for known signal and noise, the probability that y_i photons strike the i^{th} cell in time τ, given $\bar{y}_i$, is

$$p(y_i/\bar{y}) = \frac{(\bar{y}_i\tau)^{y_i} e^{-\bar{y}_i \tau}}{y_i!} \quad . \tag{110}$$

Likewise, the probability that z_i photoelectrons are emitted from the i^{th} cell in time τ, given $\bar{y}_i$, is

$$p(z_i/\bar{y}) = \frac{(\eta\tau\bar{y}_i)^{z_i} e^{-\bar{y}_i\tau\eta}}{z_i!} \quad . \tag{111}$$

If we assume that the photoelectrons or "counts" are independent for each region (i.e., the number of electrons emitted in each region is independent of those emitted from other regions or cells) we can write

$$\begin{aligned} p(z/\bar{y}) &= p(z_1, z_2, \ldots, z_m/\bar{y}) \\ &= \prod_{i=1}^{m} p(z_i/\bar{y}) = \prod_{i=1}^{m} \frac{(\eta\tau\bar{y}_i)^{z_i} e^{-\eta\tau\bar{y}_i}}{z_i!} \quad . \end{aligned} \tag{112}$$

This is the probability that z_i electrons are emitted from cell 1, z_2 electrons are emitted from cell 2, ..., and z_m electrons are emitted from cell m all in time τ. We will also assume that the mean rate of noise photons is fixed and known when the noise is considered independently of the signal.

The estimates of $\bar{y}_i$ obtained in this section will be used for the estimator correlator detector in the section on fixed-sample detection.

Bayes' Estimate

By definition, the Bayes' estimate is the $\hat{y}(z)$ which minimizes the average risk. When an incorrect decision or estimate is made a loss or a cost results. If $\bar{y}$ is the true state of nature and we say that $\hat{y}$ is the state of nature, we lose an amount $c(\bar{y}, \hat{y})$. The information about the experiment is contained in the conditional density function $p(z/\bar{y})$, which is assumed known for each $\bar{y}$. With the a priori density function $f(\bar{y})$, the loss function $c(\bar{y}, \hat{y})$, and the conditional density function $p(z/\bar{y})$ for each $\bar{y}$, the estimate $\hat{y}$ can be found which minimizes the average loss.

The mathematical form of the Bayes' estimate is obtained as follows. If $\bar{y}$ is the true state of nature and we observe z, then we lose an amount $c[\bar{y}, \hat{y}(z)]$ by using the estimate $\hat{y}$. When $\bar{y}$ is the true state of nature the risk is the average of this loss function over all possible outcomes of the experiment. That is,

$$\rho(\bar{y}, \hat{y}) = \int_{-\infty}^{\infty}\!\!\cdots\!\int c[\bar{y}, \hat{y}(z)]\, p(z/\bar{y})\, dz. \qquad (1.3)$$

The risk depends on both the state of nature $\bar{y}$ and on the estimate $\hat{y}$.

The average risk is the average of $\rho(\bar{y}, \hat{y})$ over all possible states of nature. That is,

$$\rho(\hat{y}) = \int_{-\infty}^{\infty}\!\!\cdots\!\int \rho(\bar{y}, \hat{y}) f(\bar{y})\, d\bar{y} = \int_{-\infty}^{\infty}\!\!\cdots\!\int c[\bar{y}, \hat{y}(z)]\, p(z/\bar{y}) f(\bar{y})\, dz\, d\bar{y}. \qquad ([illegible].4)$$

Using Bayes' Theorem we obtain

$$\rho(\hat{\bar{y}}) = \int\cdots\int_{-\infty}^{\infty} c[\bar{y}, \hat{\bar{y}}(z)] f(\bar{y}/z) p(z) d\bar{y} dz$$
$$= \int\cdots\int_{-\infty}^{\infty} p(z) \rho_z (\hat{\bar{y}}) dz \qquad (115)$$

where $\rho_z(\hat{\bar{y}}) = \int\cdots\int_{-\infty}^{\infty} c[\bar{y}, \hat{\bar{y}}(z)] f(\bar{y}/z) d\bar{y}$ is the conditional risk. Since $p(z)$ is non-negative and independent of $\hat{\bar{y}}$, we need only minimize $\rho_z(\hat{\bar{y}})$ for each z in order to minimize $\rho(\hat{\bar{y}})$.

It is now necessary to specify a loss function $c(\bar{y}, \hat{\bar{y}})$. We will consider the quadratic loss function $K(\bar{y}-\hat{\bar{y}})^2$ where K is a positive constant. Using this loss function our problem reduces to minimizing

$$\rho_z(\hat{\bar{y}}) = K \int\cdots\int_{-\infty}^{\infty} (\bar{y}-\hat{\bar{y}})^2 f(\bar{y}/z) d\bar{y} \qquad (116)$$

for each z by choosing the appropriate estimate $\hat{\bar{y}}$. This minimization is accomplished by differentiating $\rho_z(\hat{\bar{y}})$ with respect to $\hat{\bar{y}}$ and equating the result to zero. That is,

$$\frac{\partial}{\partial \hat{\bar{y}}} \rho_z(\hat{\bar{y}}) = -2K \int\cdots\int_{-\infty}^{\infty} (\bar{y}-\hat{\bar{y}}) f(\bar{y}/z) d\bar{y} = 0. \qquad (117)$$

The solution of this equation for $\hat{\bar{y}}$ yields the Bayes' estimate

$$\hat{\bar{y}} = \int\cdots\int_{-\infty}^{\infty} \bar{y} f(\bar{y}/z) d\bar{y}. \qquad (118)$$

Hence, for a quadratic loss function the Bayes' estimate $\hat{\bar{y}}$ is the mean of the a posteriori distribution $f(\bar{y}/z)$.

We will now find the Bayes' estimate of $\bar{y}_i$ which is the average number of signal-plus-noise photons which are incident upon the i^{th} cell of the image plane per unit time. Assume that the a priori density function of the mean rate of signal-plus-noise photons $\bar{y}$ is

$$f(\bar{y}) = \prod_{i=1}^{m} f(\bar{y}_i) = \prod_{i=1}^{m} \frac{\alpha_i(\alpha_i \bar{y}_i)^{u_i - 1} e^{-\alpha_i \bar{y}_i}}{\Gamma(u_i)}, \quad \bar{y}_i \geq 0 \qquad (119)$$

$$= 0, \qquad \text{otherwise}$$

where m is the number of cells in the image plane. The conditional Poisson distribution of the output photoelectrons z is

$$p(z/\bar{y}) = \prod_{i=1}^{m} p(z_i/\bar{y}) = \prod_{i=1}^{m} \frac{e^{-\eta\tau \bar{y}_i}(\eta\tau \bar{y}_i)^{z_i}}{z_i!} \qquad (120)$$

where z_i is the number of photoelectrons or "counts" emitted from the i^{th} cell during time τ. The a posteriori density function of $\bar{y}$, $f(\bar{y}/z)$, is given by Bayes' formula

$$f(\bar{y}/z) = \frac{p(z/\bar{y})f(\bar{y})}{p(z)} \qquad (121)$$

where

$$p(z) = \underbrace{\int \cdots \int}_{m} {}_{-\infty}^{\infty} p(z/\bar{y})f(\bar{y})d\bar{y}$$

$$= \prod_{i=1}^{m} \frac{\alpha_i^{u_i}(\eta\tau)^{z_i}\, \Gamma(z_i+u_i)}{z_i!\; \Gamma(u_i)(\alpha_i+\eta\tau)^{z_i+u_i}} \quad . \qquad (122)$$

The a posteriori density function becomes

$$f(\bar{y}/z) = \prod_{i=1}^{m} \frac{(\alpha_i+\eta\tau)[\bar{y}_i(\alpha_i+\eta\tau)]^{z_i+u_i-1} e^{-(\alpha_i+\eta\tau)\bar{y}_i}}{\Gamma(z_i+u_i)} \quad . \qquad (123)$$

The Bayes' estimate (conditional mean of $\bar{y}_i$) is then

$$\hat{\bar{y}}_i = \frac{z_i + u_i}{\alpha_i + \eta\tau} \quad . \tag{124}$$

For this case the conditional variance of $\bar{y}_i$ is

$$\text{var}\,(\bar{y}_i/z) = \frac{z_i + u_i}{(\alpha_i + \eta\tau)^2} \quad . \tag{125}$$

As τ becomes large, $E(\bar{y}_i/z)$ approaches $z_i/\eta\tau$ and var $(\bar{y}_i/z)$ approaches zero. Hence, as τ becomes large the estimate $\hat{\bar{y}}_i$ approaches the true value of $\bar{y}_i$ since the variance approaches zero. For multiple sampling the Bayes' estimate becomes

$$\hat{\bar{y}}_i(k) = \mu_i^k = \frac{\sum_{j=1}^{k} z_i^j + u_i}{(\alpha_i + k\eta\tau)} \tag{126}$$

where the superscript k represents the number of samples.

Now assume that the a priori density function of the mean rate of signal photons $\bar{s}$ ($\bar{s} = A\bar{x}$) is

$$f(\bar{s}) = \prod_{i=1}^{m} f(\bar{s}_i) = \prod_{i=1}^{m} \frac{\beta_i(\beta_i \bar{s}_i)^{u_i - 1} e^{-\beta_i \bar{s}_i}}{\Gamma(u_i)}, \quad \bar{s}_i \geq 0$$

$$= 0, \qquad \text{otherwise.} \tag{127}$$

The mean rate of noise photons are now assumed to be fixed and known. The distribution of the output photoelectrons z_i is

$$p(z_i) = \int_0^{\infty} p(z_i/\bar{s})\, f(\bar{s})\, d\bar{s}$$

$$= \frac{e^{-\bar{n}_i \eta\tau} (\eta\tau)^{z_i} \beta_i^{u_i}}{z_i!\ \Gamma(u_i)} \sum_{\ell=0}^{z_i} \binom{z_i}{\ell} \frac{\bar{n}_i^{-\ell}\ \Gamma(z_i + u_i - \ell)}{(\eta\tau + \beta_i)^{z_i + u_i - \ell}} \qquad (128)$$

where

$$\binom{z_i}{\ell} = z_i! / (z_i - \ell)!\ \ell!\ . \qquad (129)$$

Using the Bayes' formula and (120), (127), and (128) yields

$$f(\bar{s}_i/z) = \frac{e^{-(\beta_i + \eta\tau)\bar{s}_i} (\bar{s}_i + \bar{n}_i)^{z_i}\ \bar{s}_i^{u_i - 1} (\eta\tau + \beta_i)^{z_i + u_i}}{\sum_{\ell=0}^{z_i} \binom{z_i}{\ell} [\bar{n}_i (\eta\tau + \beta_i)]^{\ell}\, \Gamma(z_i + u_i - \ell)}\ . \qquad (130)$$

The Bayes' estimate of $\bar{s}_i$ is then

$$\hat{\bar{s}}_i = E(\bar{s}_i/z) = \frac{1}{(\beta_i + \eta\tau)} \frac{\sum_{j=0}^{z_i} \binom{z_i}{j} [\bar{n}_i(\beta_i + \eta\tau)]^{j}\, \Gamma(z_i + u_i + 1 - j)}{\sum_{\ell=0}^{z_i} \binom{z_i}{\ell} [\bar{n}_i(\beta_i + \eta\tau)]^{\ell}\, \Gamma(z_i + u_i - \ell)}\ . \qquad (131)$$

For $\bar{n}_i = 0$ this estimate becomes $\hat{\bar{s}}_i = (z_i + u_i)/(\beta_i + \eta\tau)$ which is the same as (124).

The estimate of the object $\bar{x}$ using the estimate of the objects' image $\bar{s}$ is

$$\hat{\bar{x}} = (A'A)^{-1} A' \hat{\bar{s}} \qquad (132)$$

where $\hat{\bar{s}} = \begin{pmatrix} \hat{\bar{s}}_1 \\ \vdots \\ \hat{\bar{s}}_m \end{pmatrix}$

and $(A'A)^{-1}A'$ is the pseudoinverse of the matrix A (Deutsch, 1965). The direct derivation of the Bayes' estimate of the mean rate of photons $\bar{x}$ emitted from the object is very difficult and will not be considered.

Maximum A Posteriori Estimate

When no costs are specified in an estimation problem, a reasonable estimation procedure is to maximize the a posteriori density function $f(\bar{y}/z) = f(\bar{y})\ p\ (z/\bar{y})/p(z)$. The maximum a posteriori estimate $\hat{\bar{y}}$ is defined as the value of $\bar{y}$ that maximizes $f(\bar{y}/z)$.

The maximum a posteriori estimate of $\bar{y}$ is found by solving the m simultaneous equations.

$$\frac{\partial}{\partial \bar{y}_i} f(\bar{y}/z) = 0 \qquad (i = 1, 2, \ldots, m). \tag{133}$$

Since $f(\bar{y}/z)$ is a monotonic function, $\ell n f(\bar{y}/z)$ has its maximum for the same values of $\bar{y}$ that maximizes $f(\bar{y}/z)$. Hence, we can solve the equivalent m simultaneous equations.

$$\frac{\partial}{\partial \bar{y}_i} \ell n\ f(\bar{y}/z) = 0 \qquad (i = 1, 2, \ldots, m). \tag{134}$$

There may be several roots of these equations in which case the solution $\hat{\bar{y}}$ that yields the highest peak of the function $f(\bar{y}/z)$ must be chosen. Since the denominator of $f(\bar{y})p(z/\bar{y})/p(z)$ does not depend on $\bar{y}$, maximizing $f(\bar{y}/z)$ is equivalent to maximizing $f(\bar{y})p(z/\bar{y})$.

We will now find the maximum a posteriori estimate of $\bar{y}_i$ which is the average number of photons that are incident upon the i^{th} cell or region of the image plane per unit time. Assume that the a priori density function of the mean rate of signal-plus-noise photons $\bar{y}$ is

$$f(\bar{y}) = \prod_{i=1}^{m} f(\bar{y}_i) = \prod_{i=1}^{m} \frac{\alpha_i(\alpha_i \bar{y}_i)^{u_i-1} e^{-\alpha_i \bar{y}_i}}{\Gamma(u_i)}, \quad \bar{y}_i \geq 0 \tag{135}$$

$$= 0, \qquad \text{otherwise.}$$

The conditional Poisson distribution of the output photoelectrons z is

$$p(z/\bar{y}) = \prod_{i=1}^{m} p(z_i/\bar{y}) = \prod_{i=1}^{m} \frac{e^{-\eta\tau \bar{y}_i} (\eta\tau \bar{y}_i)^{z_i}}{z_i!} . \tag{136}$$

Now maximize $p(z/\bar{y})\, f(\bar{y})$ with respect to $\bar{y}$. That is,

$$\frac{\partial}{\partial \bar{y}_j} \ln p(z/\bar{y}) f(\bar{y}) = \frac{\partial}{\partial \bar{y}_j} \sum_{i=1}^{m} \left[\ln f(\bar{y}_i) + \ln p(z_i/\bar{y}) \right] \tag{137}$$

$$= \frac{(u_i - 1)}{\bar{y}_j} - \alpha_j + \frac{z_j}{\bar{y}_j} - \eta\tau = 0.$$

Hence, the maximum a posteriori estimate of $\bar{y}_i$ is

$$\hat{\bar{y}}_i = \frac{z_i + u_i - 1}{\alpha_i + \eta\tau} . \tag{138}$$

Now assume that the a priori density function of the mean rate of signal photons is

$$f(\bar{s}) = \prod_{i=1}^{m} \frac{\beta_i(\beta_i \bar{s}_i)^{u_i-1} e^{-\beta_i \bar{s}_i}}{\Gamma(u_i)}, \qquad \bar{s}_i \geq 0$$

$$= 0, \qquad \text{otherwise.} \tag{139}$$

Assume that the noise is fixed and known. Maximize $p(z/\bar{s})\, f(\bar{s})$, where $p(z/\bar{s})$ is given in (136), with respect to $\bar{s}$. That is,

$$\frac{\partial}{\partial \bar{s}_j} \ln p(z/\bar{s}) f(\bar{s}) = \frac{\partial}{\partial \bar{s}_j} \sum_{i=1}^{m} [\, \ln f(\bar{s}) + \ln p(z_i/\bar{s})]$$

$$= \frac{u_j - 1}{\bar{s}_j} - \beta_j - \eta\tau + \frac{z_j}{\bar{s}_j + \bar{n}_j} = 0. \tag{140}$$

Hence, the maximum a posteriori estimate of $\bar{s}_i$ is

$$\hat{\bar{s}}_i = \frac{z_i + u_i - 1 - \bar{n}_i(\eta\tau + \beta_i) + \sqrt{[\bar{n}_i(\eta\tau + \beta_i) - z_i + u_i - 1]^2 + 4z_i(u_i - 1)}}{2(\eta\tau + \beta_i)}. \tag{141}$$

This solution becomes the same as (138), as it should, when the mean rate of noise photons n_i is equal to zero. Again the estimate of the object $\bar{x}$ using the estimate of the objects' image $\bar{s}$ is that of (132).

Consider finding the estimate of $\bar{x}$ directly when the a priori density function of $\bar{s}$ is that of (139). To find the maximum a posteriori estimate of $\bar{x}$ we must find the value of $\bar{x}$ that maximizes $p(z/\bar{s})f(\bar{s})$. This value of $\bar{x}$ is found by taking a derivative of $p(z/\bar{s})f(\bar{s})$ with respect to $\bar{x}$ and setting the result equal to zero. When this is done we obtain

$$\sum_{i=1}^{m} \left[\frac{(u_i - 1)}{a_i \bar{x}} - \beta_i + \eta\tau + \frac{z_i}{a_i \bar{x} + \bar{n}_i} \right] a_{ij} = \sum_{i=1}^{m} Y_i a_{ij} = 0 \qquad (142)$$

where $Y_i = [(u_i - 1)/ a_i\bar{x} - \beta_i + \eta\tau + z_i/(a_i\bar{x} + \bar{n}_i)]$. In matrix notation (142) can be written as

$$Y' \alpha_j = 0 \qquad (j = 1, 2, \ldots, m) \qquad (143)$$

where α_j is the j^{th} column vector of the system matrix A. Since (142) must hold for all $j = 1, \ldots, m$, it constitutes a set of m equations with m unknowns, $\bar{x}_1, \ldots, \bar{x}_m$, which are to be estimated. For the set of m equations we have

$$A' Y = \underline{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \qquad (144)$$

which says that Ymust be contained in the null space of A'. If A has an inverse, the null space of A' contains only the zero vector. In this case we have the solution $Y = \underline{0}$ which is a unique solution. $Y = \underline{0}$ is always a solution of $A'Y = \underline{0}$, but if A is not square there will be other solutions (no unique solution). For $Y'A = \underline{0}$, Y must be orthogonal to all the columns of A.

We will consider the solution $Y = \underline{0}$ which implies that

$$a_i\bar{x} = z_i + u_i - 1 - \bar{n}_i(\eta\tau + \beta_i) + \sqrt{[\bar{n}_i(\eta\tau + \beta_i) - z_i + u_i - 1]^2 + 4z_i(u_i - 1)} \qquad (145)$$

$$= \hat{s}_i .$$

In matrix notation this becomes $A\hat{x} = \hat{s}$ or

$$\hat{\bar{x}} = (A'A)^{-1}A'\,\hat{\bar{s}}\,. \tag{146}$$

This particular solution gives the same result that was obtained by first finding the estimate of $\bar{s}$ and then using it to find the estimate of $\bar{x}$. We must keep in mind that for the present case, (144) has other solutions besides (145) and that there may exist several relative maxima of $\ell n\, p(z/\bar{s})f(\bar{s})$, from which we can have only one maximal value. It is quite possible that the solution of (145) does not give rise to the absolute maximum of $p(z/\bar{s})f(\bar{s})$ in which case (145) would not be the maximum a posteriori estimate.

Maximum Likelihood Estimate

The maximum a posteriori estimate depends on the a priori probabiltiy density function, but for some situations no such a priori information may be available. Under these circumstances we need to consider the maximum likelihood estimate. Also, if the a priori density function is broad and flat and relatively independent of $\bar{y}$ over the region where $p(z/\bar{y})$ is significant (i.e., initial knowledge of $\bar{y}$ is very small) then maximizing $p(z/\bar{y})$ is nearly equivalent to maximizing $f(\bar{y}/z)$. The value of $\bar{y}$ which maximizes $p(z/\bar{y})$ is defined as the maximum likelihood estimate of $\bar{y}$.

Since $p(z/\bar{y})$ is a monotonic function we can solve either of the two equations $\partial p(z/\bar{y})/\partial \bar{y}_j = 0$ or $\partial\, \ell n p(\bar{z}/\bar{y})/\partial \bar{y}_j = 0$ for the estimate of $\bar{y}$. Thus we have

$$\frac{\partial}{\partial \bar{y}_j} \ln p(z/\bar{y}) = \frac{\partial}{\partial \bar{y}_j} \sum_{i=1}^{m} [z_i \ln(\eta\tau \bar{y}_i) - \eta\tau \bar{y}_i - \ln(z_i!)] = \frac{z_j}{y_j} - \eta\tau . \quad (147)$$

Hence, the maximum likelihood estimate of $\bar{y}_i$ is

$$\hat{\bar{y}}_i = \frac{z_i}{\eta\,\tau} . \quad (148)$$

The maximum likelihood estimate of $\bar{s}$ is found in the same manner as above. That is,

$$\frac{\partial}{\partial \bar{s}_j} \ln p(z/\bar{y}) = \frac{z_j}{\bar{s}_j + \bar{n}_j} - \eta\tau = 0. \quad (149)$$

Hence, the maximum likelihood estimate of $\bar{s}_i$ is

$$\hat{\bar{s}}_i = \frac{z_i}{\eta\,\tau} - \bar{n}_i . \quad (150)$$

The estimate of the object $\bar{x}$ using the estimate of the object's image $\bar{s}$ is the same as in (132).

Now consider finding the maximum likelihood estimate of the object $\bar{x}$. To find this estimate we need to find the value of $\bar{x}$ that maximizes $p(z/\bar{x})$ or alternately $\ln p(z/\bar{x})$. Taking the derivative of $\ln p(z/\bar{x})$ with respect to $\bar{x}$ and setting the result equal to zero yields

$$\frac{\partial}{\partial \bar{x}_j} \ln p(z/\bar{x}) = \sum_{i=1}^{m} \left[\frac{z_i}{a_i \bar{x} + \bar{n}_i} - \eta\tau\right] a_{ij} = \sum_{i=1}^{m} Y_i a_{ij} = 0 \quad (151)$$

where $Y_i = z_i/(a_i\bar{x} + \bar{n}_i) - \eta\tau$. In matrix notation this is written as

$$Y' \alpha_j = 0 \qquad (j = 1, \ldots, m) \tag{152}$$

where α_j is the j^{th} column vector of the system matrix A. Equation (151) holds for j = 1, 2, ..., m; hence, it constitutes m equations in the m unknowns $\bar{x}_1, \ldots, \bar{x}_m$ which we want to estimate. Using matrix notation this set of m equations can be written as

$$A'Y = \underline{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{153}$$

which means that Y must be contained within the null space of A'. The discussion of the solutions of (144) for the maximum a posteriori estimate of $\bar{x}$ also applies to (152). We will consider the solution $Y = \underline{0}$ which implies that

$$a_i\bar{x} = \frac{z_i}{\eta\tau} - \bar{n}_i \, . \tag{154}$$

For general A the maximum likelihood estimate of $\bar{x}$ is

$$\hat{\bar{x}} = (A'A)^{-1}A'(\frac{z}{\eta\tau} - \bar{n}) = (A'A)^{-1}A'(\hat{\bar{y}} - \bar{n}) = (A'A)^{-1}A'(\hat{\bar{s}}). \tag{155}$$

This is the same estimate that is obtained when we estimate $\bar{s}$ first and then use this estimate to estimate $\bar{x}$. It should again be pointed out that (153) has other solutions besides (154), one or more of which may give rise to a maximum value of $p(z/\bar{x})$, which is larger than the value due to (154), in which case (154) would not be the maximum likelihood estimate.

Discussion of Estimates

The Bayes' estimate and the maximum a posteriori estimate have received some criticism. The basic argument against them is based upon the requirement of a priori probability density functions for the random variables to be observed during an experiment. The maximum likelihood estimate has objectionable small-sample-size properties (Deutsch, 1965).

It should be pointed out that all the estimates of $\bar{y}_i$ of this section, given that $\bar{y}_i$ has a gamma distribution, are linear estimates (linear with respect to the observable z) as is the minimum MSE estimate. However, the Bayes' and maximum a posteriori estimates of $\bar{y}_i$ and $\bar{s}_i$, given that $\bar{s}_i$ has a gamma distribution, are not linear estimates. The Bayes' estimate of $\bar{y}_i$ is the same as the minimum MSE estimate of $\bar{y}_i$, given that $\bar{y}_i$ has a gamma distribution. The maximum a posteriori estimate of $\bar{y}_i$ differs from the minimum MSE estimates only by a minus ... e in the numerator. The maximum a posteriori estimate becomes the same as the Bayes' and minimum MSE estimates for large values of u_i.

FIXED-SAMPLE SIGNAL DETECTION

Introduction

For the fixed-sample detection procedure, we will consider the Bayes' decision rule. By definition, Bayes' decision rule is the decision rule that minimizes the average loss.

We will consider only situations in which there are two possible states of nature, ω_1 and ω_2. We will assume that the a priori probabilities $p(\omega_1)$ and $p(\omega_2)$ are known. Also, we will assume the probabilities $p(z/\omega_i)$ $(i = 1, 2)$ are known.

We observe the outcome of the experiment and decide which state of nature is present. If we choose ω_j as the state of nature when ω_i is the true state of nature, we lose an amount $c(\omega_i, \omega_j) = c_{ij}$. We want to find the decision rule that minimizes the average loss.

To determine the form of Bayes' decision rule, first calculate the average loss resulting when decision rule $d(\cdot)$ is used. If we observe z when ω_i is the true state of nature we lose amount $c[\omega_i, d(z)]$. Also, when ω_i is true, z occurs with probability $p(z/\omega_i)$. For the average loss or risk we can then write

$$\rho(\omega_i, d) = \sum_{z=0}^{\infty} c[\omega_i, d(z)]\, p(z/\omega_i). \qquad (156)$$

Since ω_i occurs with probability $p(\omega_i)$, the average loss or risk resulting when the decision rule d is used is

$$\rho(d) = \sum_{i=1}^{2} \rho(\omega_i, d)\, p(\omega_i). \tag{157}$$

By definition Bayes' decision rule d* is that rule which minimizes the average risk. That is, we require that $\rho(d^*) \leq \rho(d)$ for all possible decision rules d.

The derivation of this rule follows.

$$\begin{aligned} \rho(d) &= \sum_{i=1}^{2} p(\omega_i) \sum_{z=0}^{\infty} c[\omega_i, d(z)]\, p(z/\omega_i) \\ &= \sum_{z=0}^{\infty} \sum_{i=1}^{2} c[\omega_i, d(z)]\, p(\omega_i/z) p(z) \\ &= \sum_{z=0}^{\infty} p(z) \rho_z(d). \end{aligned} \tag{158}$$

The quantity $\rho_z(d) = \sum_{i=1}^{2} c[\omega_i, d(z)]\, p(\omega_i/z)$ is called the conditional risk.

Since p(z) is non-negative and does not depend on our decision rule, we can minimize $\rho(d)$ by choosing the decision rule d that minimizes $\rho_z(d)$ for each z. The conditional risk can be written as

$$\rho_z(d) = c[\omega_1, d(z)]\, p(\omega_1/z) + c[\omega_2, d(z)]\, p(\omega_2/z). \tag{159}$$

In accordance with the decision rule we must choose either $d(z) = \omega_1$ or $d(z) = \omega_2$. If we choose $d(z) = \omega_1$ then $\rho_z(d) = c_{11} p(\omega_1/z) + c_{21} p(\omega_2/z)$. Similarly, if we choose $d(z) = \omega_2$ then $\rho_z(d) =$

$c_{12}p(\omega_1/z) + c_{22}p(\omega_2/z)$. Bayes' decision rule becomes: choose ω_1 if $c_{11}p(\omega_1/z) + c_{21}p(\omega_2/z) < c_{12}p(\omega_1/z) + c_{22}p(\omega_2/z)$ or by using Bayes' formula we have:

$$\text{choose } \omega_1 \text{ if } \ell(z) > \delta, \qquad \text{choose } \omega_2 \text{ if } \ell(z) \leq \delta, \tag{160}$$

where

$$\delta \triangleq \frac{p(\omega_2)}{p(\omega_1)}\left(\frac{c_{21}-c_{22}}{c_{12}-c_{11}}\right) \tag{161}$$

and

$$\ell(z) = \ell(z_1, z_2, \ldots, z_m) = \frac{p(z_1, z_2, \ldots, z_m/\omega_1)}{p(z_1, z_2, \ldots, z_m/\omega_2)} \tag{162}$$

is the likelihood ratio of the observation and $p(z_1, \ldots z_m/\omega_i)$ is the conditional probability of observing "counts" $z_1, \ldots, z_m$ when ω_i is the true state of nature. We have assumed that these "counts" are independent; hence, we have

$$\ell(z) = \prod_{i=1}^{m} \frac{p(z_i/\omega_1)}{p(z_i/\omega_2)} . \tag{163}$$

We will assume throughout the rest of this paper that $c_{22} = c_{11} = 0$, $c_{12} = c_{21}$, and that $p(\omega_1) = p(\omega_2) = 1/2$; hence, $\delta = 1$.

Since the natural logarithm is a monotonically increasing function of its argument, the logarithm can be taken of both sides of the inequalities in (160) to obtain an equivalent decision procedure. Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } L(z) \triangleq \ln \ell(z) > 0,$$
$$\text{choose } \omega_2 \text{ if } L(z) \leq 0. \qquad (164)$$

Our problem is to observe the "counts" at the output of the detector and from this decide which of the two signals gave rise to this output. Figure 12 illustrates the problem.

We will be interested in assessing the error associated with Bayes' decision rule. When analyzing the error probability, we will consider throughout the remainder of this paper only one region or cell in the image plane where measurements will be made. This assumption is for mathematical convenience and does not introduce a serious loss of generality. Hence, for error analysis our problem is reduced to a case of a scalar signal. At the end of this section a comparison will be made of the error probabilities for the various cases that will be considered.

Two Known Signals

Assume that at the input of the detector we have one of two possible signals with known mean rates $\bar{y}_1$ and $\bar{y}_2$ both of which contain any constant background noise that may be present. For convenience in discussion, we will refer to $\bar{y}_1$ and $\bar{y}_2$ as the signals. For the present case, we have two possible states of nature:

$$\omega_1 : \bar{y} = \bar{y}_1 \quad \text{(known)},$$
$$\omega_2 : \bar{y} = \bar{y}_2 \quad \text{(known)},$$

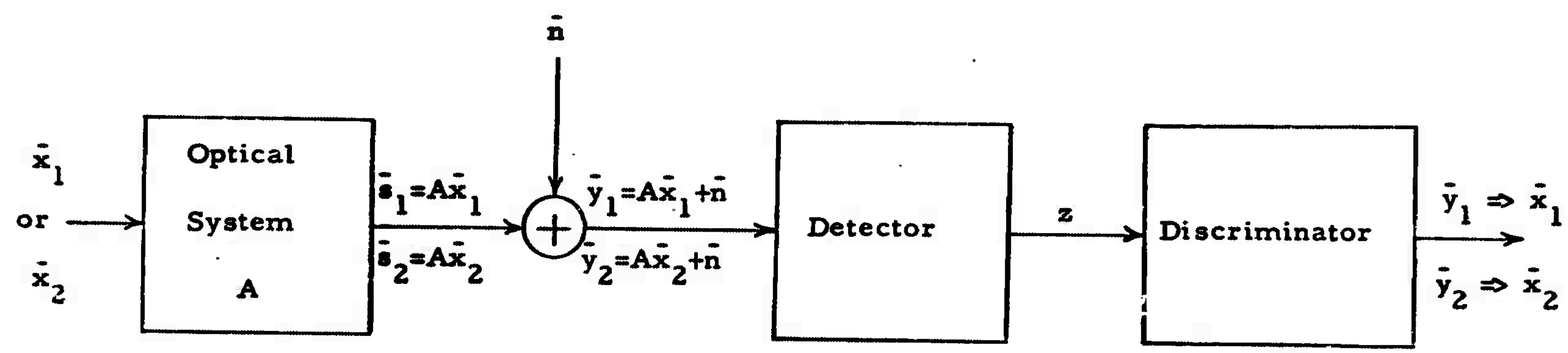

Figure 12. Diagram of the optical system, and detector, and discriminator.

where $\bar{y}_1 > \bar{y}_2$. Knowing that signal $\bar{y}_i$ is present is equivalent to knowing that $\bar{x}_i$ is present. The objective of our discrimination precedure is to decide which of the two states of nature, ω_1 or ω_2, is the true state of nature.

The probability of z_i photoelectrons being emitted in time τ from the i^{th} cell of the image plane, given that ω_j is the true state of nature, is

$$p(z_i/\omega_j) = \frac{(\eta\tau\bar{y}_{ji})^{z_i} e^{-\eta\tau\bar{y}_{ji}}}{z_i!} \quad (j=1,2) \tag{165}$$

where $\bar{y}_{ji}$ is the i^{th} component of signal $\bar{y}_j$(i.e., $\bar{y}_j = (\bar{y}_{j1}, \ldots, \bar{y}_{ji}, \ldots, \bar{y}_{jm})$). Hence, the likelihood ratio is

$$\ell(z) = \frac{p(z/\omega_1)}{p(z/\omega_2)} = \prod_{i=1}^{m} \left(\frac{\bar{y}_{1i}}{\bar{y}_{2i}}\right)^{z_i} e^{-\eta\tau(\bar{y}_{1i}-\bar{y}_{2i})} \tag{166}$$

where m is the number of cell or measurement positions in the image plane. A more convenient form is

$$L(z) \triangleq \ln \ell(z) = \sum_{i=1}^{m} [z_i \ln(\bar{y}_{1i}/\bar{y}_{2i}) - \eta\tau(\bar{y}_{1i}-\bar{y}_{2i})]. \tag{167}$$

Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} z_i \ln(\bar{y}_{1i}/\bar{y}_{2i}) > \eta\tau \sum_{i=1}^{m} (\bar{y}_{1i}-\bar{y}_{2i}), \tag{168}$$

choose ω_2 otherwise.

This detector has the form of a digital matched filter where the filter is matched to $\ell n(\bar{y}_{1i}/\bar{y}_{2i})$. This detector is illustrated in Figure 13. The above case may, as a special case, be considered as two known signals $\bar{s}_1$ and $\bar{s}_2$ ($\bar{s}_2$ may or may not be zero but $\bar{s}_1 > \bar{s}_2$) imbedded in known noise $\bar{n}$. The two possible states of nature for this case are:

$$\omega_1 : \bar{y} = \bar{s}_1 + \bar{n} \text{ (known signal plus noise)},$$
$$\omega_2 : \bar{y} = \bar{s}_2 + \bar{n} \text{ (known signal plus noise)}.$$

Bayes' decision rule for this case is:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} z_i \, \ell n\left(\frac{\bar{s}_{1i}+\bar{n}_i}{\bar{s}_{2i}+\bar{n}_i}\right) > \eta\tau \sum_{i=1}^{m} (\bar{s}_{1i}-\bar{s}_{2i}),$$
$$\text{choose } \omega_2 \text{ otherwise,} \tag{169}$$

where $\bar{s}_{ji}$ is the i^{th} component of the vector signal $\bar{s}_j$. For the case where $\bar{s}_2 = 0$, the noise is uniform (i.e., $\bar{n}_i = \bar{n}_o$ for all i), and we have a small signal-to-noise ratio $\bar{s}_{1i}/\bar{n}_i$ for all i, Bayes' decision rule reduces to:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} z_i s_{1i} > \eta\tau\bar{n}_o \sum_{i=1}^{m} \bar{s}_{1i},$$
$$\text{choose } \omega_2 \text{ otherwise.} \tag{170}$$

This detector has the form of a digital matched filter where the filter is matched to the signal $\bar{s}_1$. This detector is illustrated in Figure 14.

We will now determine the error probability associated with the general case of two possible known signals $\bar{y}_1$ and $\bar{y}_2$. As mentioned earlier, only a single cell of the image plane will be used in our error

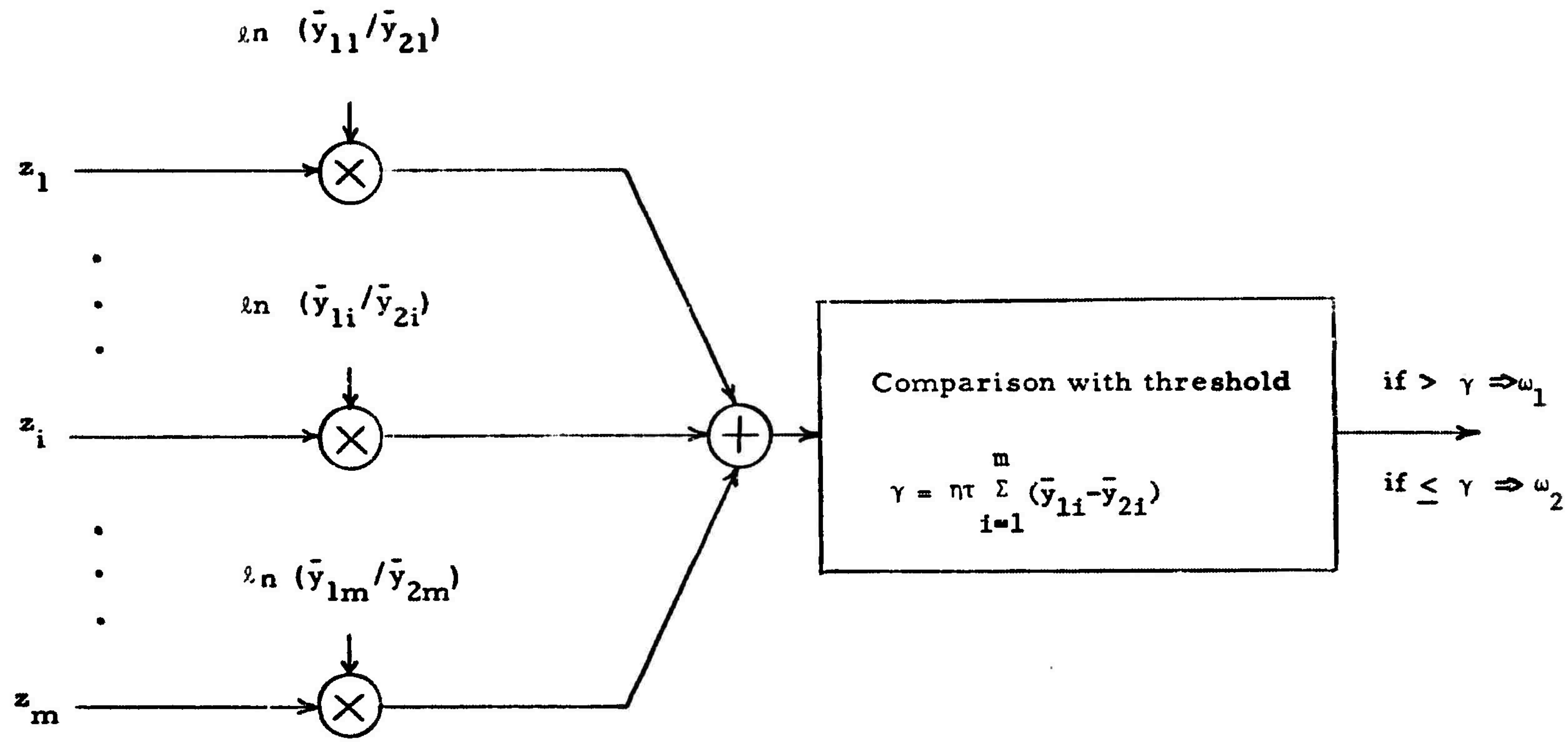

Figure 13. Diagram of the general multiple cell threshold detector designed for the two known signals $\bar{y}_1$ and $\bar{y}_2$.

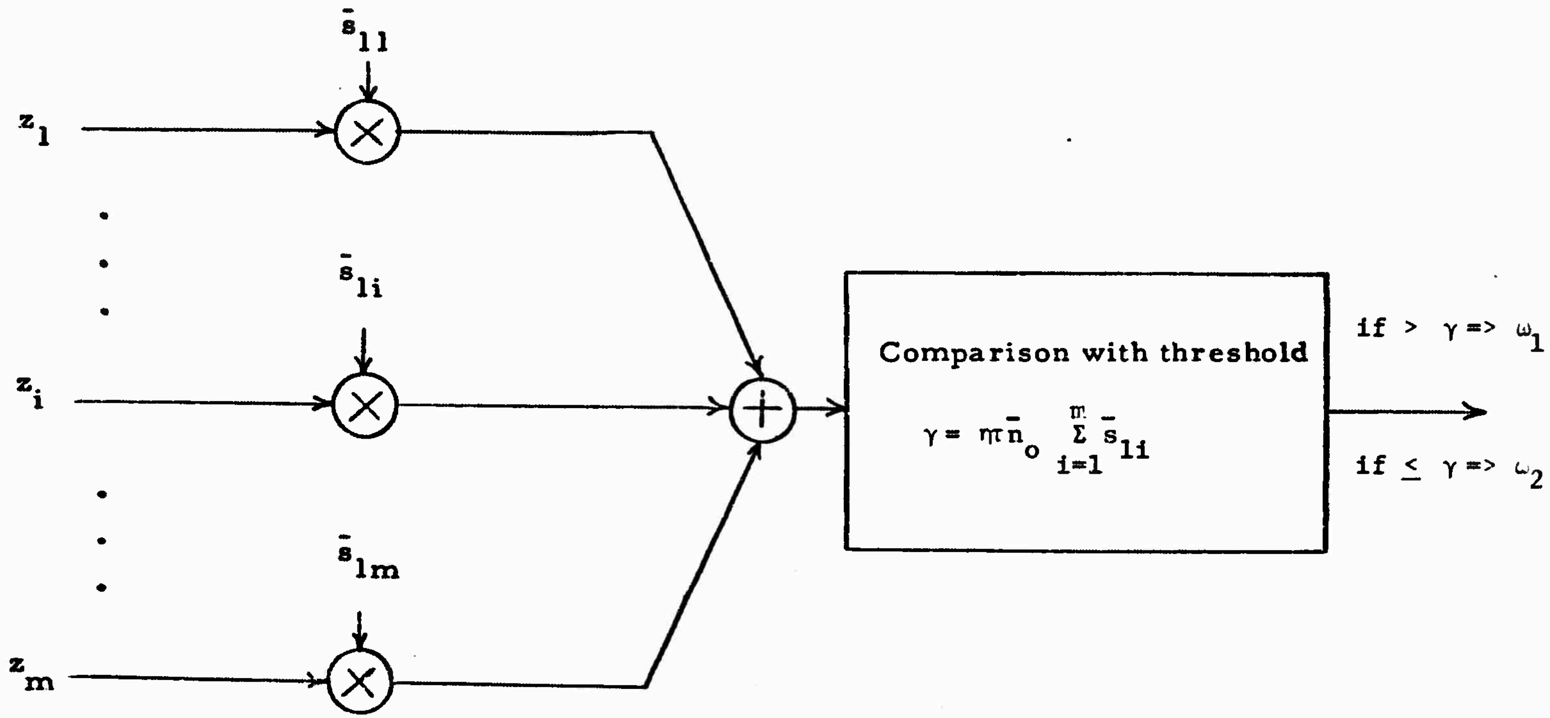

Figure 14. Diagram of the multiple cell digital filter matched to the signal $\bar{s}_1$.

analysis, thus reducing the problem to one of a scalar signal. Bayes' decision rule now becomes:

$$\text{choose } \omega_1 \text{ if } z > \frac{\eta\tau(\bar{y}_1 - \bar{y}_2)}{\ln(\bar{y}_1/\bar{y}_2)} = \gamma, \tag{171}$$

choose ω_2 otherwise,

where z, $\bar{y}_1$, and $\bar{y}_2$ are all scalars.

We will now determine the error probability for this procedure. The general expression for the error probability is

$$P_e = p(\omega_1)P(FD) + p(\omega_2)P(FA) \tag{172}$$

where P(FA) is the probability of saying $\bar{y}_1$ is present when $\bar{y}_2$ is present (probability of false alarm), and P(FD) is the probability of saying $\bar{y}_2$ is present when $\bar{y}_1$ is present (probability of false dismissal). Assume that $p(\omega_2) = p(\omega_1) = 1/2$. Thus

$$P_e = \tfrac{1}{2}P[z > \gamma/\omega_2] + \tfrac{1}{2}P[z < \gamma/\omega_1]$$

$$= \tfrac{1}{2}\left[1 - \sum_{z=0}^{\gamma} \frac{(\eta\tau\bar{y}_2)^z e^{-\eta\tau\bar{y}_2}}{z!} + \sum_{z=0}^{\gamma} \frac{(\eta\tau\bar{y}_1)^z e^{-\eta\tau\bar{y}_1}}{z!}\right]. \tag{173}$$

For large values of z the Central Limit Theorem applies and hence z becomes approximately Gaussian. In order to completely specify the Gaussian densities $f(z/\omega_2)$ and $f(z/\omega_1)$ we need to find the conditional means, $E(z/\omega_1)$ and $E(z/\omega_2)$, and the conditional variances, var (z/ω_1) and var (z/ω_2). These are given as follows:

$$E(z/\omega_1) = \mu_1 = \eta T \bar{y}_1, \tag{174}$$

$$E(z/\omega_2) = \mu_2 = \eta T \bar{y}_2, \tag{175}$$

$$\text{Var}\,(z/\omega_1) = \sigma_1^2 = \eta T \bar{y}_1, \tag{176}$$

$$\text{Var}\,(z/\omega_2) = \sigma_2^2 = \eta T \bar{y}_2. \tag{177}$$

Hence, using (174) - (177) and assuming a Gaussian probability density function approximation we have

$$\begin{aligned} P_e &\approx \tfrac{1}{2}\left[\int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi\sigma_2^2}}\, e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}\, dx + \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi\sigma_1^2}}\, e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\, dx\right] \\ &= \tfrac{1}{2}\left[\int_{(\gamma-\mu_2)/\sigma_2}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx + \int_{(-\gamma+\mu_1)/\sigma_1}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx\right]. \end{aligned} \tag{178}$$

Under conditions where the Gaussian approximation holds, this form of the error probability would be very useful in analyzing the error probability for the general case of m cells or measurement positions in the image plane (see (168)).

It is desirable to compute the error probabilities of these two methods to determine how good the Gaussian approximation is. The Gaussian approximation turns out to be good even for very small values of z. The approximation is good for small values of z because we are working with the "tails" of two distributions and when they are added the approximation errors of the distributions compensate. This does not hold in general for $p(\omega_1) \neq 1/2$. For an indication of the validity of this approximation, see Figure 15 and Table 1. Figure 15 compares the ratio of the Poisson error probability and the Gaussian-approximation error probability for different ratios of the signals $\bar{y}_1$ and $\bar{y}_2$. The Poisson error probabilities and the approximate error probabilities have a significant difference only for situations where the error probabilities become small (e.g., $P_e < 10^{-3}$). For the case of very small error probabilities, we are far out on the "tails" of the distributions and our approximation breaks down. The approximation breaks down because the Central Limit Theorem does not coverge linearly and hence the Gaussian approximation holds only for the center portion of the distribution.

Two Unknown Signals

Now assume that at the input of the detector we have one of two possible signals $\bar{y}_1$ and $\bar{y}_2$, both of which contain any background noise that may be present. These signals are assumed fixed but unknown and having each of their elements taken from statistically independent gamma distributions with known parameters:

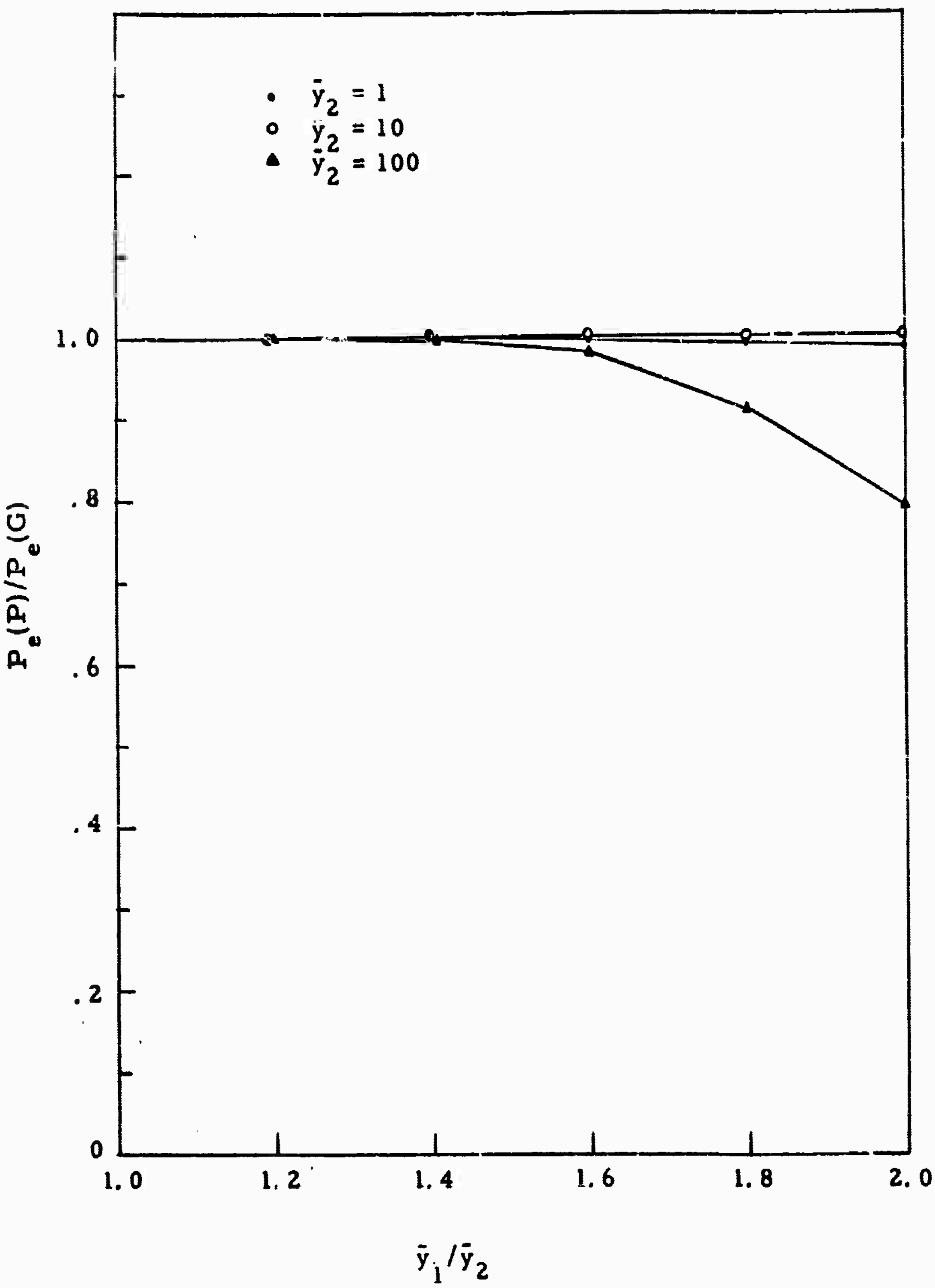

Figure 15. Ratio of Poisson error probability and Gaussian-approximation error probability versus ratio of $\bar{y}_1$ and $\bar{y}_2$ for the case of discriminating between two known signals $\bar{y}_1$ and $\bar{y}_2$ where $\eta\tau = 1$.

Table 1. Poisson error probabilities and Gaussian-approximation error probabilities for the cases of discriminating between two known signals and discriminating between two unknown signals.

	Two known Signals			Two Unknown Signals								
				$\bar{\bar{y}}_2 = 1$			$\bar{\bar{y}}_2 = 10$			$\bar{\bar{y}}_2 = 100$		
$\bar{\bar{y}}_1/\bar{\bar{y}}_2$	$\bar{y}_2 = 1$	$\bar{y}_2 = 10$	$\bar{y}_2 = 100$	u = 1	u = 10	u = 100	u = 1	u = 10	u = 100	u = 1	u = 10	u = 100
P_e (Poisson)												
1.2	.4634	.3821	.1699	.4634	.4634	.4634	.3821	.3821	.3821	.1699	.1699	.1699
1.4	.4280	.2816	.0334	.4280	.4280	.4280	.2816	.2816	.2816	.0345	.0345	.0334
1.6	.3946	.2008	$.4005 \times 10^{-2}$	.3946	.3946	.3946	.2008	.2008	.2008	$.4594 \times 10^{-2}$	$.4594 \times 10^{-2}$	$.4169 \times 10^{-2}$
1.8	.3635	.1391	$.3149 \times 10^{-3}$	.3635	.3635	.3635	.1391	.1391	.1391	$.5220 \times 10^{-3}$	$.5220 \times 10^{-3}$	$.4171 \times 10^{-3}$
2.0	.3351	.0942	$.1804 \times 10^{-4}$	.3351	.3351	.3351	.1008	.0942	.0942	$.6754 \times 10^{-4}$	$.6754 \times 10^{-4}$	$.3566 \times 10^{-4}$
P_e (Gaussian approximation)												
1.2	.4619	.3813	.1698	.4737	.4639	.4622	.4652	.4174	.3872	.4638	.3913	.2549
1.4	.4271	.2807	.0333	.4507	.4311	.4276	.4359	.3477	.2917	.4335	.3048	.1081
1.6	.3951	.2001	$.4058 \times 10^{-2}$	.4303	.4012	.3957	.4105	.2894	.2145	.4075	.2376	.0403
1.8	.3654	.1386	$.3435 \times 10^{-3}$	.4120	.3737	.3648	.3885	.2411	.1545	.3849	.1859	.01414
2.0	.3379	.0936	$.2267 \times 10^{-4}$	.3954	.3484	.3390	.3690	.2014	.1096	.3650	.1465	.004925

$$f(\bar{y}_i) = \prod_{i=1}^{m} f(\bar{y}_{ji}) = \prod_{i=1}^{m} \frac{\alpha_{ji}(\alpha_{ji}\bar{y}_{ji})^{u_{ji}-1} e^{-\alpha_{ji}\bar{y}_{ji}}}{\Gamma(u_{ji})} \quad (j=1, 2), \; \bar{y}_i \geq 0$$

$$= 0, \qquad \text{otherwise.} \tag{179}$$

Our two possible states of nature are:

ω_1: $\bar{y} = \bar{y}_1$ (unknown),

ω_2: $\bar{y} = \bar{y}_2$ (unknown),

where $\bar{\bar{y}}_1 > \bar{\bar{y}}_2$. (The quantity $\bar{\bar{y}}$ is the expected value of $\bar{y}$.) We want to decide which of these two states of nature is present.

The likelihood ratio in this case is

$$\ell(z) = \frac{p(z/\omega_1)}{p(z/\omega_2)} = \frac{\int_0^\infty p(z/\omega_1, \bar{y}_1) f(\bar{y}_1) d\bar{y}_1}{\int_0^\infty p(z/\omega_2, \bar{y}_2) f(\bar{y}_2) d\bar{y}_2} \tag{180}$$

where $$p(z/\omega_j, \bar{y}_j) = \prod_{i=1}^{m} \frac{(\eta\tau\bar{y}_{ji})^{z_i} e^{-\eta\tau\bar{y}_{ji}}}{z_i!} \tag{181}$$

and $$p(z/\omega_j) = \int_0^\infty p(z/\omega_j, \bar{y}_j) f(\bar{y}_j) d\bar{y}_j$$

$$= \prod_{i=1}^{m} \left(\frac{\alpha_{ji}}{\alpha_{ji}+\eta\tau}\right)^{u_{ji}} \frac{(\eta\tau)^{z_i}\Gamma(z_i+u_{ji})}{z_i!(\eta\tau+\alpha_{ji})^{z_i}\Gamma(u_{ji})} \quad (j=1, 2). \tag{182}$$

Hence, the likelihood ratio becomes

$$\ell(z) = \prod_{i=1}^{m} \frac{\alpha_{1i}^{u_{1i}} (\alpha_{2i}+\eta\tau)^{u_{2i}} \Gamma(z_i+u_{1i}) \Gamma(u_{2i})}{\alpha_{2i}^{u_{2i}} (\alpha_{1i}+\eta\tau)^{u_{1i}} \Gamma(z_i+u_{2i}) \Gamma(u_{1i})} \left(\frac{\eta\tau+\alpha_{2i}}{\eta\tau+\alpha_{1i}}\right)^{z_i} . \tag{183}$$

Assume for convenience that $u_{1i} = u_{2i} = u_i$ for all i, then

$$\ell(z) = \prod_{i=1}^{m} \left(\frac{\alpha_{1i}}{\alpha_{2i}}\right)^{u_i} \left(\frac{\alpha_{2i}+\eta\tau}{\alpha_{1i}+\eta\tau}\right)^{z_i+u_i} \tag{184}$$

or

$$L(z) = \ln \ell(z) = \sum_{i=1}^{m} \left[z_i \ln\left(\frac{\alpha_{2i}+\eta\tau}{\alpha_{1i}+\eta\tau}\right) + u_i \left[\ln\left(\frac{\alpha_{2i}+\eta\tau}{\alpha_{1i}+\eta\tau}\right) - \ln\left(\frac{\alpha_{2i}}{\alpha_{1i}}\right)\right]\right] . \tag{185}$$

Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} z_i \ln\left(\frac{\alpha_{2i}+\eta\tau}{\alpha_{1i}+\eta\tau}\right) > \sum_{i=1}^{m} u_i \left[\ln\left(\frac{\alpha_{2i}}{\alpha_{1i}}\right) - \ln\left(\frac{\alpha_{2i}+\eta\tau}{\alpha_{1i}+\eta\tau}\right)\right] , \tag{186}$$

choose ω_2 otherwise.

For this case the detector has the form of a digital matched filter where the filter is matched to $\ell n[(\alpha_{2i} + \eta\tau)/(\alpha_{1i} + \eta\tau)]$. This detector is illustrated in Figure 16.

For an error analysis, we again consider the single cell case. Bayes' decision rule for this case becomes:

$$\text{choose } \omega_1 \text{ if } z > u\left[\frac{\ell n(\alpha_2/\alpha_1)}{\ell n\left(\frac{\alpha_2+\eta\tau}{\alpha_1+\eta\tau}\right)} - 1\right] = \gamma, \quad (187)$$

choose ω_2 otherwise.

The error probability is

$$P_e = \tfrac{1}{2}\left[1 - \sum_{z=0}^{\gamma} \frac{(\eta\tau u/\alpha_2)^z e^{-\eta\tau u/\alpha_2}}{z!} + \sum_{z=0}^{\gamma} \frac{(\eta\tau u/\alpha_1)^z e^{-\eta\tau u/\alpha_1}}{z!}\right]. \quad (188)$$

Again assume a Gaussian approximation for z. The conditional means and variances necessary to specify the Gaussian densities $f(z/\omega_2)$ and $f(z/\omega_1)$ are as follows:

$$E(z/\omega_i) = E[E(z/\omega_i, \bar{y}_i)] = E(\eta\tau\bar{y}_i) = \eta\tau u/\alpha_i, \quad ([illegible])$$

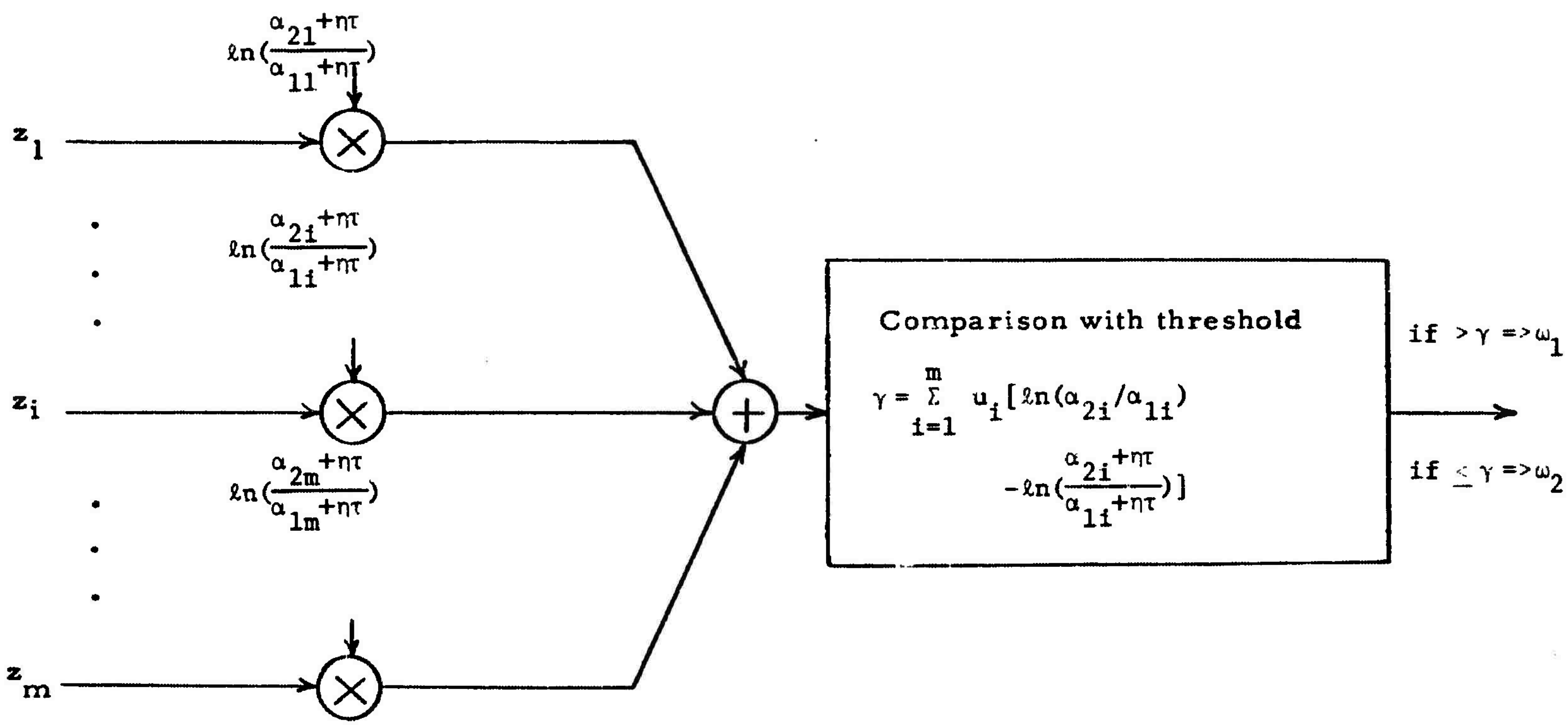

Figure 16. Diagram of the multiple cell threshold detector designed for the two unknown signals $\bar{y}_1$ and $\bar{y}_2$.

$$E(z/\omega_1) = \mu_1 = n\tau u/\alpha_1, \tag{190}$$

$$E(z/\omega_2) = \mu_2 = n\tau u/\alpha_2, \tag{191}$$

$$\mathrm{Var}\ (z/\omega_i) = E[\mathrm{Var}(z/\omega_i,\bar{y}_i)] + \mathrm{Var}[E(z/\omega_i,\bar{y}_i)], \tag{192}$$

$$\mathrm{Var}(z/\omega_1) = \sigma_1^{\,2} = n\tau u/\alpha_1 + (n\tau/\alpha_1)^2 u, \tag{193}$$

$$\mathrm{Var}(z/\omega_2) = \sigma_2^{\,2} = n\tau u/\alpha_2 + (n\tau/\alpha_2)^2 u. \tag{194}$$

Hence, using (189)-(194) and a Gaussian probability density function approximation we have

$$P_e \approx \tfrac{1}{2}\left[\int_{\frac{\gamma-\mu_2}{\sigma_2}}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx + \int_{\frac{-\gamma+\mu_1}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx\right]. \tag{195}$$

We need to compute the error probabilities for these two cases (Poisson and Gaussian-approximation) and compar them to determine how good the Gaussian approximation is for the two-unknown-signals case.

See Figures 17 and 18 and Table 1 for a comparison of the Poisson error probabilities with the Gaussian-approximation error probabilities. Figure 17 shows the Poisson error probabilities and the Gaussian-approximation error probabilities for the cases where both signals are known and b signals are unknown. The Poisson error probabilities, known and unknown signals, coincides with the Gaussian-approximation error probabilities for known signals. The Gaussian approximation for unknown signals is reasonably accurate only for cases where the variances of the unknown signals are smaller than about 0.1. Figure 18 compares the ratio of the Poisson error probability and the Gaussian-approximation error probability for different ratios of the expected values of the signals (i.e., $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$) and different values of u (note that var $\bar{y}_i = \bar{\bar{y}}_i^{\,2} / u$ (i = 1, 2)).

One Unknown Signal and One Known Signal

Consider the case of having present at the input of the detector either the known signal $\bar{y}_2$ or the unknown signal $\bar{y}_1$. For this case we have the two possible states of nature:

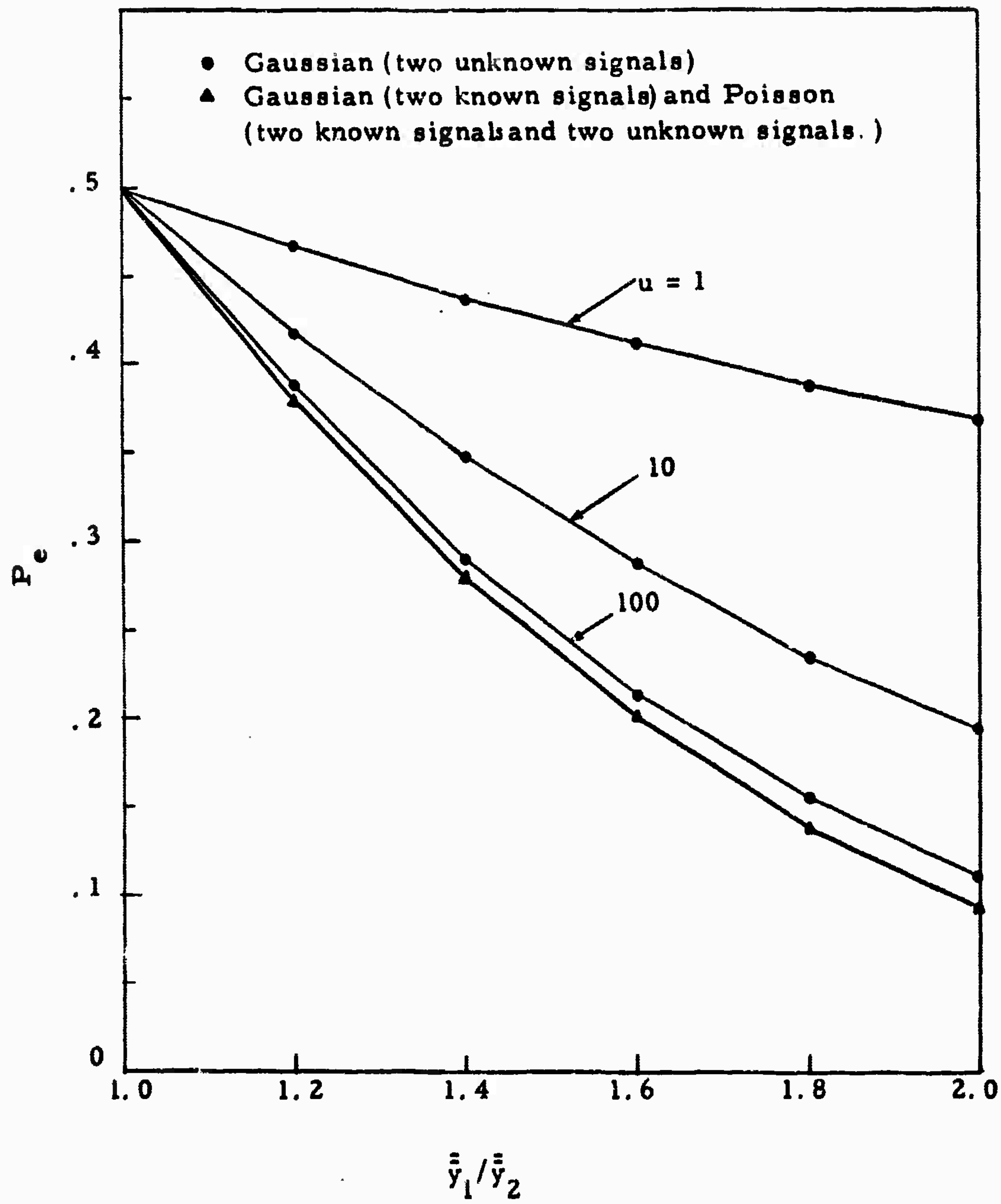

Figure 17. Poisson error probabilities and Gaussian-approximation error probabilities versus the ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for the case of two known signals and the case of two unknown signals where $\bar{\bar{y}}_2 = 10$ and $\eta\tau = 1$.

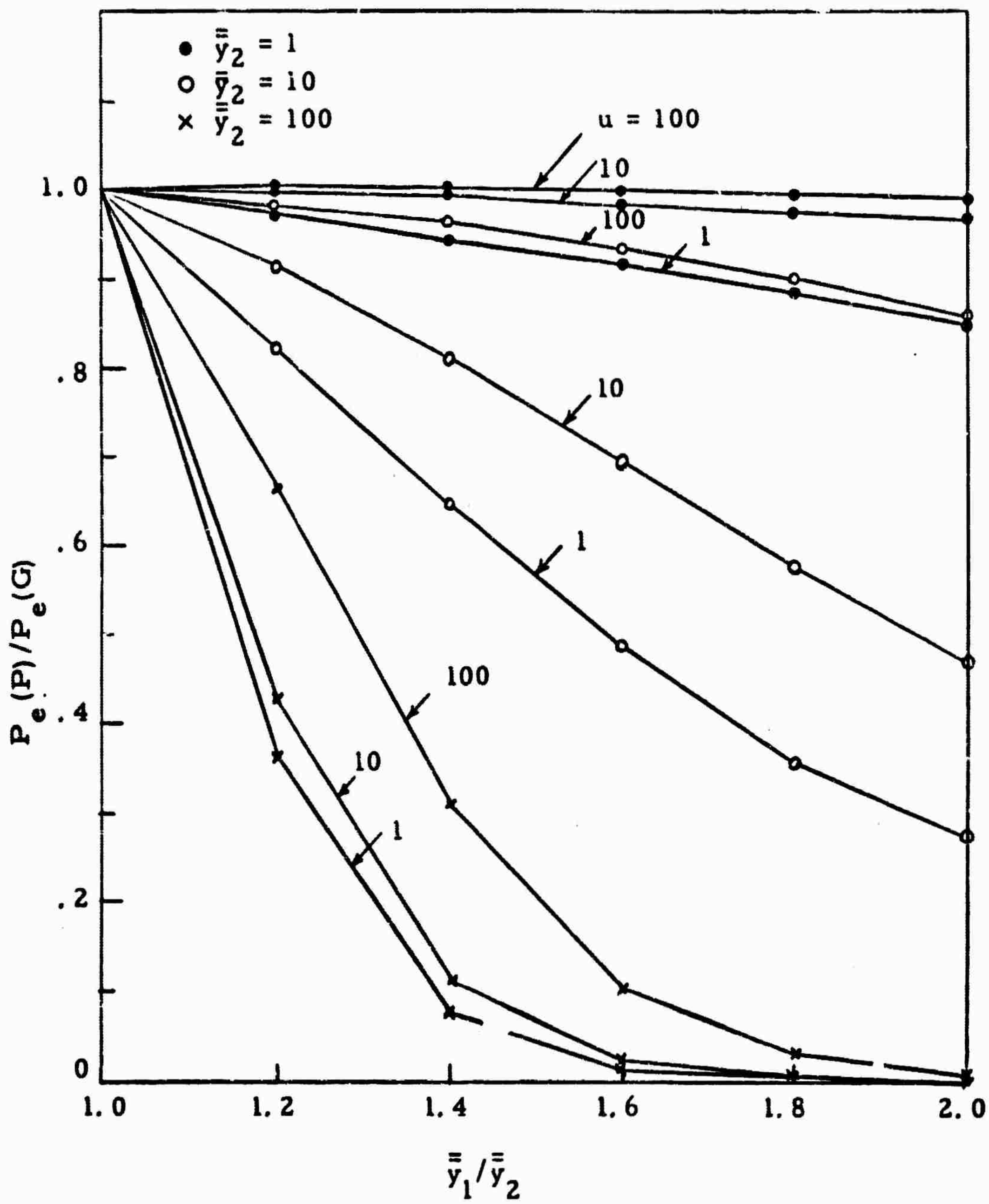

Figure 18. Ratio of Poisson error probability and Gaussian-approximation error probability versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for the case of discriminating between two unknown signals $\bar{y}_1$ and $\bar{y}_2$ where $\eta\tau = 1$.

ω_1: $\bar{y} = \bar{y}_1$ (unknown),

ω_2: $\bar{y} = \bar{y}_2$ (known),

where $\bar{y}_1 > \bar{y}_2$. We consider this case because of its advantages in analyzing the detection error from which we can gain additional insight to the detection process. This case corresponds to physical cases of large signal-to-noise ratio in which we neglect the noise and discriminate between two signals. The unknown signal $\bar{y}_1$ is assumed fixed but unknown and having each of its elements taken from statistically independent gamma distributions with known parameters:

$$f(\bar{y}_1) = \prod_{i=1}^{m} f(\bar{y}_{1i}) = \prod_{i=1}^{m} \frac{\alpha_i(\alpha_i \bar{y}_{1i})^{u_i - 1} e^{-\alpha_i \bar{y}_i}}{\Gamma(u_i)}, \quad \bar{y}_i \geq 0 \tag{196}$$

$$= 0, \qquad \text{otherwise.}$$

Bayes' decision rule is:

$$\text{choose } \omega_1 \text{ if } \ell(z) = \frac{p(z/\omega_1)}{p(z/\omega_2)} > 1, \tag{197}$$

$$\text{choose } \omega_2 \text{ otherwise,}$$

where

$$p(z/\omega_2) = \prod_{i=1}^{m} \frac{(\eta\tau \bar{y}_{2i})^{z_i} e^{-\eta\tau \bar{y}_{2i}}}{z_i!} \tag{198}$$

and

$$p(z/\omega_1) = \int_0^\infty p(z/\omega_1,\bar{y}_1)f(\bar{y}_1)d\bar{y}_1 = \prod_{i=1}^{m} \left(\frac{\alpha_i}{\alpha_i+n\tau}\right)^{u_i} \frac{(n\tau)^{z_i}\Gamma(z_i+u_i)}{z_i!(n\tau+\alpha_i)^{z_i}\Gamma(u_i)} \quad . \tag{199}$$

The likelihood ratio becomes

$$\ell(z) = \prod_{i=1}^{m} \left(\frac{\alpha_i}{\alpha_i+n\tau}\right)^{u_i} \frac{\Gamma(z_i+u_i)\, e^{n\tau\bar{y}_{2i}}}{[\bar{y}_{2i}(n\tau+\alpha_i)]^{z_i}\Gamma(u_i)} \quad . \tag{200}$$

Taking the logarithm of the likelihood ratio yields

$$L(z) = \ln\ell(z) = \sum_{i=1}^{m}\Big[u_i\ln\left(\frac{\alpha_i}{\alpha_i+n\tau}\right) - \ln\Gamma(u_i) + \ln\Gamma(z_i+u_i) - z_i\ln[\bar{y}_{2i}(n\tau+\alpha_i)] + n\tau\bar{y}_{2i}\Big]. \tag{201}$$

Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m}\Big[\ln\Gamma(z_i+u_i) - z_i\ln[\bar{y}_{2i}(n\tau+\alpha_i)]\Big] > \sum_{i=1}^{m}\Big[\ln\Gamma(u_i) + u_i\ln\left(\frac{\alpha_i+n\tau}{\alpha_i}\right) - n\tau\bar{y}_{2i}\Big], \tag{202}$$

choose ω_2 otherwise.

Using the approximation that

$$\ln\Gamma(x+1) \simeq \tfrac{1}{2}\ln 2\pi - x + (x+\tfrac{1}{2})\ln x \tag{203}$$

we can write for Bayes' decision rule:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m}\left[(z_i+u_i-\tfrac{1}{2})\ \ln(z_i+u_i-1)-z_i\ \ln[\bar{y}_{2i}(n\tau+\alpha_i)]-z_i\right] \tag{204}$$

$$> \sum_{i=1}^{m}\left[(u_i-\tfrac{1}{2})\ln(u_i-1)\ +\ u_i\ln\ (\frac{\alpha_i+n\tau}{\alpha_i})-n\tau\bar{y}_{2i}\right],$$

choose ω_2 otherwise.

This detector is illustrated in Figure 19. For the single cell case, Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } (z+u-\tfrac{1}{2})\ \ln(z+u-1)-z\ \ln[\bar{y}_2(n\tau+\alpha)]-z$$

$$>(u-\tfrac{1}{2})\ \ln(u-1)\ +\ u\ln(\frac{\alpha+n\tau}{\alpha})-n\tau\bar{y}_2, \tag{205}$$

choose ω_2 otherwise.

To analytically find the error probability of this decision rule as was done in the previous cases appears very difficult. For this case, computer simulation (Monte Carlo method) was used to analyze the error probability. The results are discussed later.

Unknown Signal Imbedded in Known Noise

Consider the case where we have one unknown signal s which if present at the input of the detector is imbedded in known noise $\bar{n}$. The two possible states of nature are:

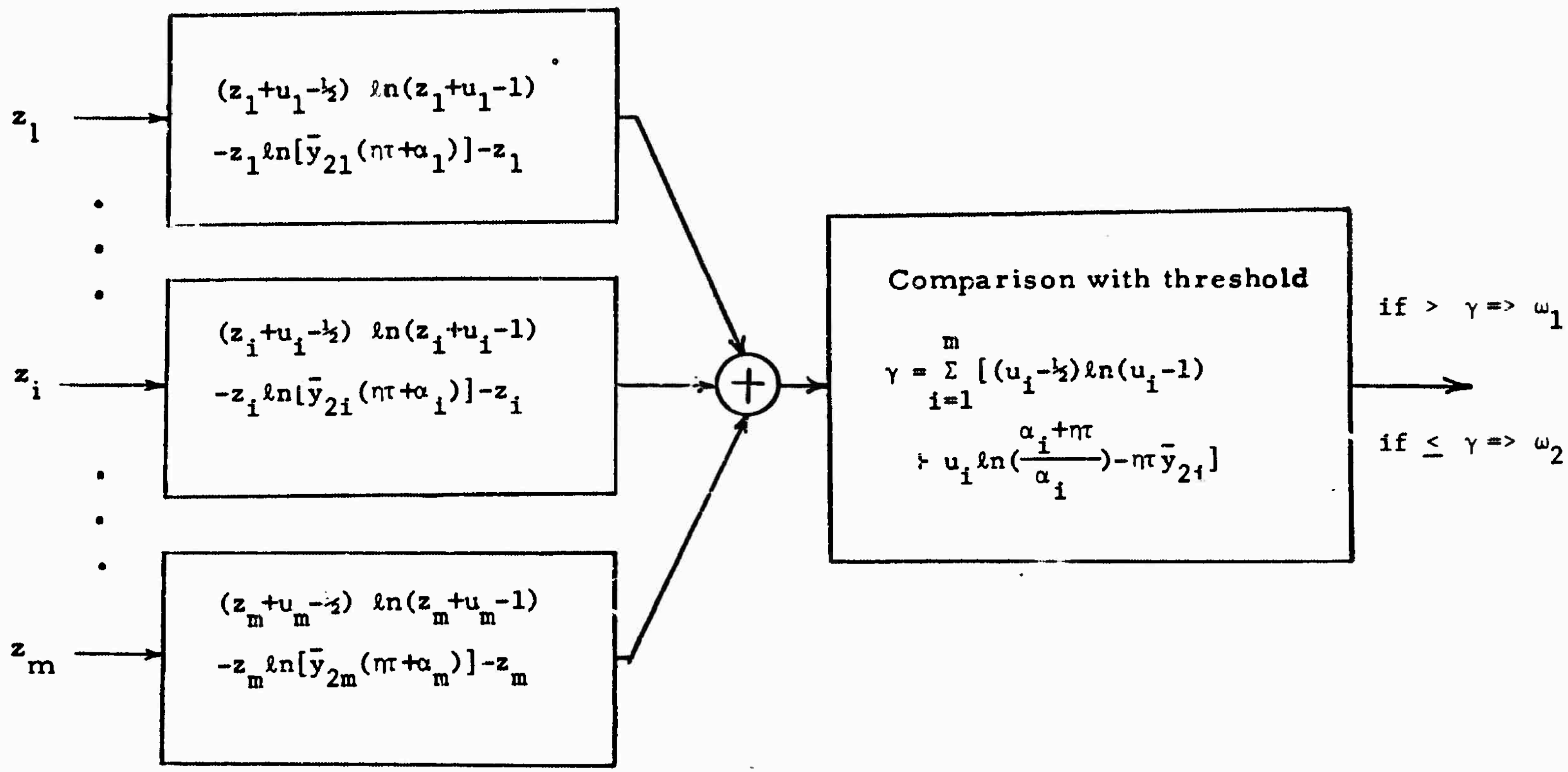

Figure 19. Diagram of the multiple cell threshold detector designed for one known signal $\bar{y}_2$ and one unknown signal $\bar{y}_1$.

ω_1: $\bar{y} = \bar{s} + \bar{n}$ (unknown signal plus noise),

ω_2: $\bar{y} = \bar{n}$ (noise alone).

The signal $\bar{s}$ is assumed fixed but unknown and having each of its elements taken from statistically independent gamma distributions with known parameters:

$$f(\bar{s}) = \prod_{i=1}^{m} f(\bar{s}_i) = \prod_{i=1}^{m} \frac{\beta_i(\beta_i\bar{s}_i)^{u_i-1} e^{-\beta_i\bar{s}_i}}{\Gamma(u_i)}, \quad \bar{s}_i \geq 0 \tag{206}$$

$$= 0, \qquad \text{otherwise.}$$

For this case

$$p(z/\omega_2) = \prod_{i=1}^{m} \frac{(\eta\tau\bar{n}_i)^{z_i} e^{-\eta\tau\bar{n}_i}}{z_i!} \tag{207}$$

and

$$p(z/\omega_2,\bar{s}) = \prod_{i=1}^{m} \frac{[\eta\tau(\bar{n}_i+\bar{s}_i)]^{z_i} e^{-\eta\tau(\bar{s}_i+\bar{n}_i)}}{z_i!} . \tag{208}$$

Hence,

$$\ell(z/\bar{s}) = \frac{p(z/\omega_1,\bar{s})}{p(z/\omega_2,\bar{s})} = \prod_{i=1}^{m} \left(\frac{\bar{s}_i+\bar{n}_i}{\bar{n}_i}\right)^{z_i} e^{-\eta\tau\bar{s}_i}. \tag{209}$$

Our likelihood ratio then becomes

$$\ell(z) = \int_0^\infty \ell(z/\bar{s}) f(\bar{s}) d\bar{s} = \prod_{i=1}^{m} \left(\frac{\beta_i}{\beta_i + n\tau}\right)^{u_i} \frac{\sum_{j=0}^{z_i} \binom{z_i}{j} [\bar{n}_i(\beta_i + n\tau)]^j \, \Gamma(z_i + u_i - j)}{[\bar{n}_i(\beta_i + n\tau)]^{z_i} \, \Gamma(u_i)} \quad (210)$$

Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} \left[\ln\left[\sum_{j=0}^{z_i} \binom{z_i}{j} [\bar{n}_i(\beta_i + n\tau)]^j \, \Gamma(z_i + u_i - j) \right] \right. \quad (211)$$

$$\left. - z_i \ln[\bar{n}_i(\beta_i + n\tau)] \right] > \sum_{i=1}^{m} \left[u_i \ln\left(\frac{\beta_i + n\tau}{\beta_i}\right) + \ln\Gamma(u_i) \right],$$

choose $_2$ otherwise.

Using the approximation of (203) Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \sum_{i=1}^{m} \left[\ln\left[\sum_{j=0}^{z_i} \frac{[\bar{n}_i(\beta_i + n\tau)]^j \Gamma(z_i + u_i - j)}{(z_i - j)! \, j!} \right] \right.$$

$$\left. + z_i [\ln z_i - 1 - \ln[\bar{n}_i(\beta_i + n\tau)] \right] + \tfrac{1}{2} \ln z_i \quad (212)$$

$$> \sum_{i=1}^{m} [(u_i - \tfrac{1}{2}) \ln(u_i - 1) - u_i + 1],$$

choose ω_2 otherwise.

This detector is illustrated in Figure 20.

For the single cell case Bayes' decision rule becomes:

$$\text{choose } \omega_1 \text{ if } \ln\left[\sum_{j=0}^{z_i} \frac{[\bar{n}(\beta+\eta\tau)]^j\,\Gamma(z+u-j)}{(z-j)!\;j!}\right]$$

$$+ z\left[\ln z - 1 - \ln[\bar{n}(\beta+\eta\tau)]\right] + \tfrac{1}{2}\ln z \qquad (213)$$

$$> (u-\tfrac{1}{2})\ln(u-1) - u + 1,$$

choose ω_2 otherwise.

To find the error probability of this decision rule analytically appears extremely difficult, if not impossible. For this case it is hard to use appropriate approximations without approximating the problem away. About the only course left open is to simulate the problem.

Estimator-Correlator

We have been considering the optimum detection (Bayes' decision rule) of Poisson signals with unknown parameters (mean rates) where the a priori probability distributions of the unknown parameters are available at the receiver. It follows from the optimality of the solution that it is not possible to improve the detection performance by estimating the unknown signals first and then using these estimates in a detector as if

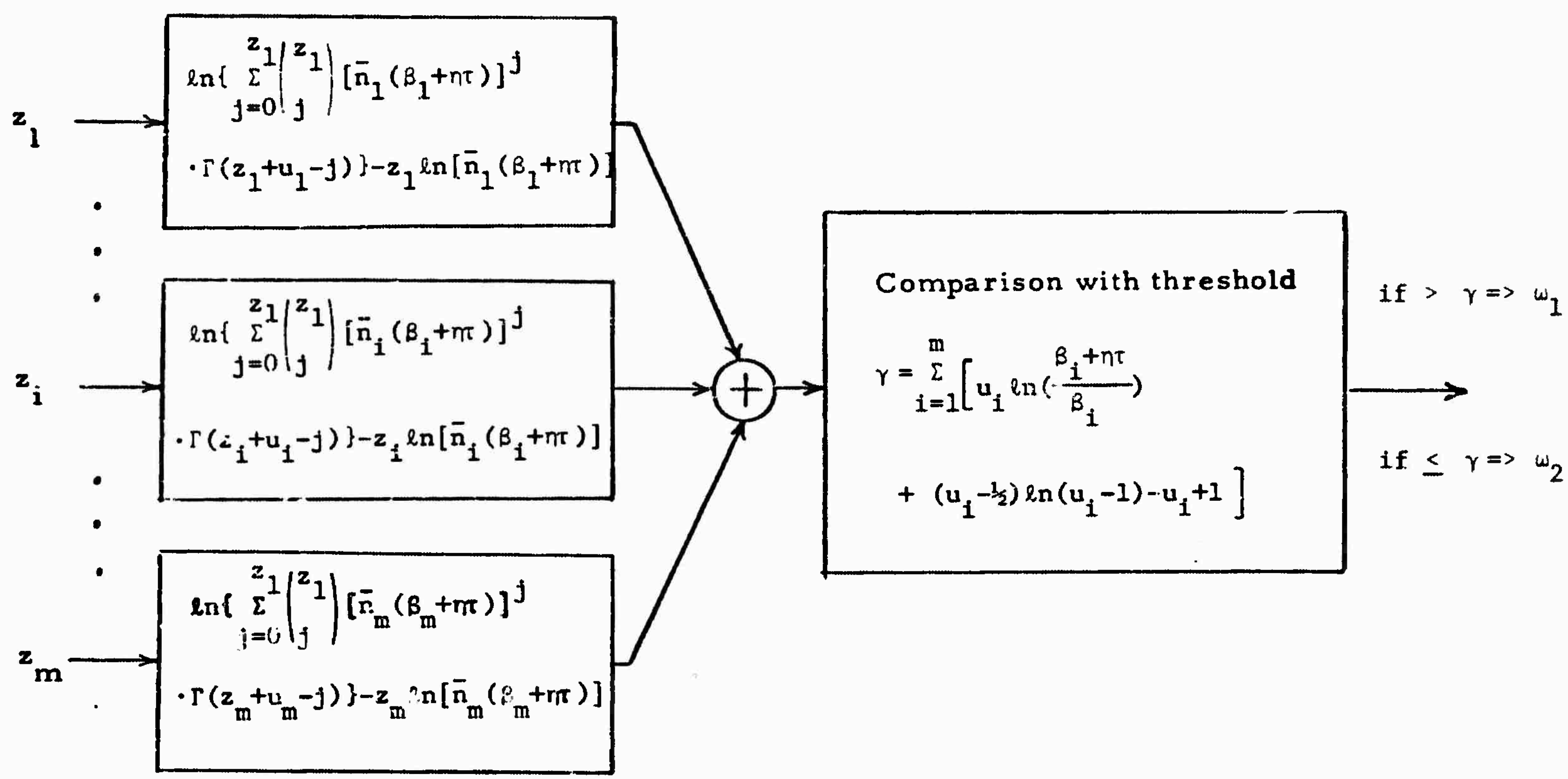

Figure 20. Diagram of the multiple cell threshold detector designed for an unknown signal $\bar{s}$ which if present is imbedded in known noise $\bar{n}$.

they were the true input signals. Kailath (1963) has shown that for the special case of Gaussian signals with unknown parameters the optimum detector can be interpreted as a minimum mean-square-error estimate of the signal followed by a detector that treats the estimate as the true value of the input signal. We will refer to this type of detector as an estimator-correlator. The results of the Gaussian signal case do not apply in general to other signal distributions.

The question arises as to the difference between the Bayes' decision rule and the decision rule associated with an estimator-correlator when the unknown signals have conditional Poisson distributions with unknown mean rates which have known gamma distributions.

Consider the case of receiving at the input of the detector one of two possible signals with conditional Poisson distributions and fixed but unknown mean rates $\bar{y}_1$ and $\bar{y}_2$. Each of th[illegible]ates are taken from gamma distributions with known parameters. For this case Bayes' decision rule (scalar case) was found to be:

$$\text{choose } \omega_1 \text{ if } z > u\left[\frac{\ln(\alpha_2/\alpha_1)}{\ln\left(\frac{\alpha_2+\eta\tau}{\alpha_1+\eta\tau}\right)} - 1\right] = \gamma, \tag{214}$$

choose ω_2 otherwise.

For the known signal case, Bayes' decision rule is:

$$\text{choose } \omega_1 \text{ if } z > \frac{\eta\tau(\bar{y}_1 - \bar{y}_2)}{\ln(\bar{y}_1/\bar{y}_2)}, \tag{215}$$

choose ω_2 otherwise.

Hence, for the estimator-correlator with which we make a comparison we have:

$$\text{choose } \omega_1 \text{ if } z > \frac{\eta\tau(\hat{\bar{y}}_1 - \hat{\bar{y}}_2)}{\ln(\hat{\bar{y}}_1/\hat{\bar{y}}_2)}, \tag{216}$$

choose ω_2 otherwise.

where $\hat{\bar{y}}_1$ and $\hat{\bar{y}}_2$ are estimates of the unknown mean rates, and here we consider these estimates to be the true mean rates.

Using the assumption that $\bar{y}_1$ and $\bar{y}_2$ are taken from known gamma distributions our estimates of them are as follows:

1) Maximum a posteriori estimate

$$\hat{\bar{y}}_i = \frac{z + u_i - 1}{\alpha_i + \eta\tau} \quad (i = 1, 2), \tag{217}$$

2) Bayes' estimate

$$\hat{\hat{y}}_i = \frac{z + u_i}{\alpha_i + \eta\tau} \quad (i = 1, 2), \tag{218}$$

3) Minimum mean-square error estimate

$$\hat{\hat{y}}_i = \frac{z + u_i}{\alpha_i + \eta\tau} \quad (i = 1, 2), \tag{219}$$

4) Maximum likelihood estimate

$$\hat{\hat{y}}_i = \frac{z}{\eta\tau} \quad (i = 1, 2) \; . \tag{220}$$

It is observed that the Bayes' estimate and the minimum MSE estimate are equal. The maximum a posteriori estimate differs from these two estimates by a minus one in the numerator and for u >> 1 all three estimates become approximately equal.

Consider the maximum a posteriori estimate in the estimator-correlator. The decision rule becomes:

$$\text{choose } \omega_1 \text{ if } z > \frac{(u-1)\left(\frac{1}{\alpha_1+\eta\tau} - \frac{1}{\alpha_2+\eta\tau}\right)}{\ln\left(\frac{\alpha_2+\eta\tau}{\alpha_1+\eta\tau}\right) - \left(\frac{1}{\alpha_1+\eta\tau} - \frac{1}{\alpha_2+\eta\tau}\right)}, \tag{221}$$

choose ω_2 otherwise.

When the Bayes' estimate and the minimum MSE estimate are used in the estimator-correlator the decision rule becomes:

$$\text{choose } \omega_1 \text{ if } z > \frac{u\left(\frac{1}{\alpha_1+\eta\tau} - \frac{1}{\alpha_2+\eta\tau}\right)}{\ln\left(\frac{\alpha_2+\eta\tau}{\alpha_1+\eta\tau}\right) - \left(\frac{1}{\alpha_1+\eta\tau} - \frac{1}{\alpha_2+\eta\tau}\right)}, \tag{222}$$

choose ω_2 otherwise.

The thresholds for the above two decision rules differ only by the terms u and u-1. For u >> 1 the two threshold become approximately equal, and hence the decision rules become approximately the same. The maximum likelihood estimate does not work in this case because there is no a priori information contained in it and hence there is no way to distinguish between $\bar{y}_1$ and $\bar{y}_2$. A comparison of the error probabilities resulting from the decision rules of the estimator-correlators with the error probabilities resulting from Bayes' decision rule will follow.

Comparison of Error Probabilities

In this section comparisons will be made of calculated and simulated error probabilities resulting from the various fixed-sample detectors that we have considered.

Figures 21, 22, and 23 show the calculated error probabilities for the two-known-signals case, the two-unknown-signals case, and both cases of the estimator-correlator for various ratios of the signals $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ and three values of u. As u increases (var $\bar{y}_i$ decreases) the error probabilities of the two-unknown-signals case and the two cases of estimator-correlators approach the error probabilities of the two-known-signals case, and in fact they all coincide for u = 100. The Bayes' and minimum MSE estimator-correlator appears to be superior to the maximum a posteriori estimator-correlator for small values of u. Figures 24 and 25 correspond to Figure 22 where the measuring interval ($\eta\tau$) has been increased to 2 and 3 respectively. Figures 26, 27 and 28 are the computer-simulated error probabilities corresponding to Figures 22, 24, and 25. The simulated results compare very favorably with the calculated results. For both the calculated and simulated results, the Bayes' and minimum MSE estimator-correlator is consistently better than the maximum a posteriori estimator-correlator. Figure 29 shows the error probabilities for the case of one unknown signal in known noise and the case of one unknown signal and one known signal. Their error probabilities are compared with the simulated error probabilities

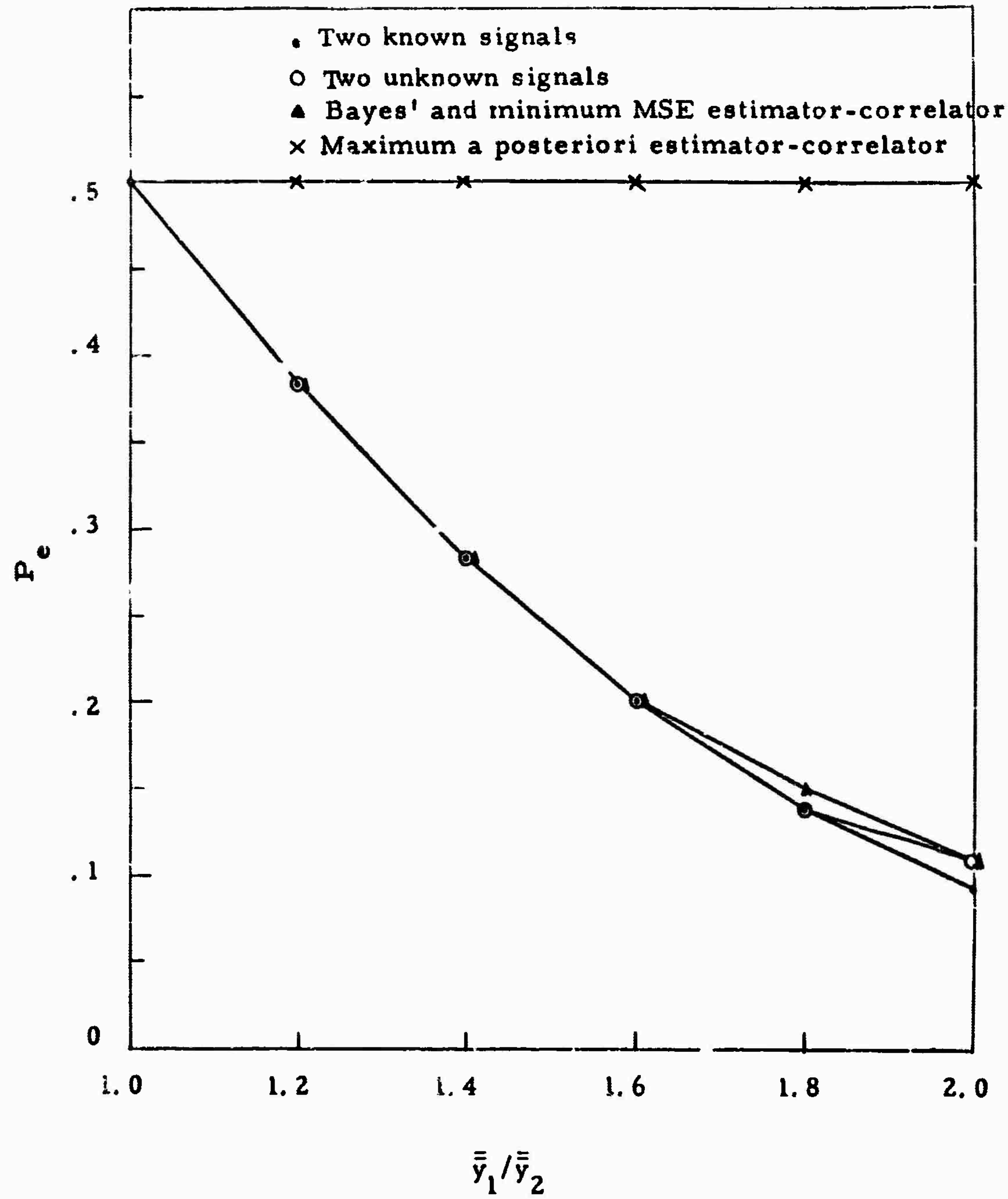

Figure 21. Calculated error probabilities versus ratio of $\bar{y}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 1, y_2 = 10, and $\eta^1\tau$ =1.

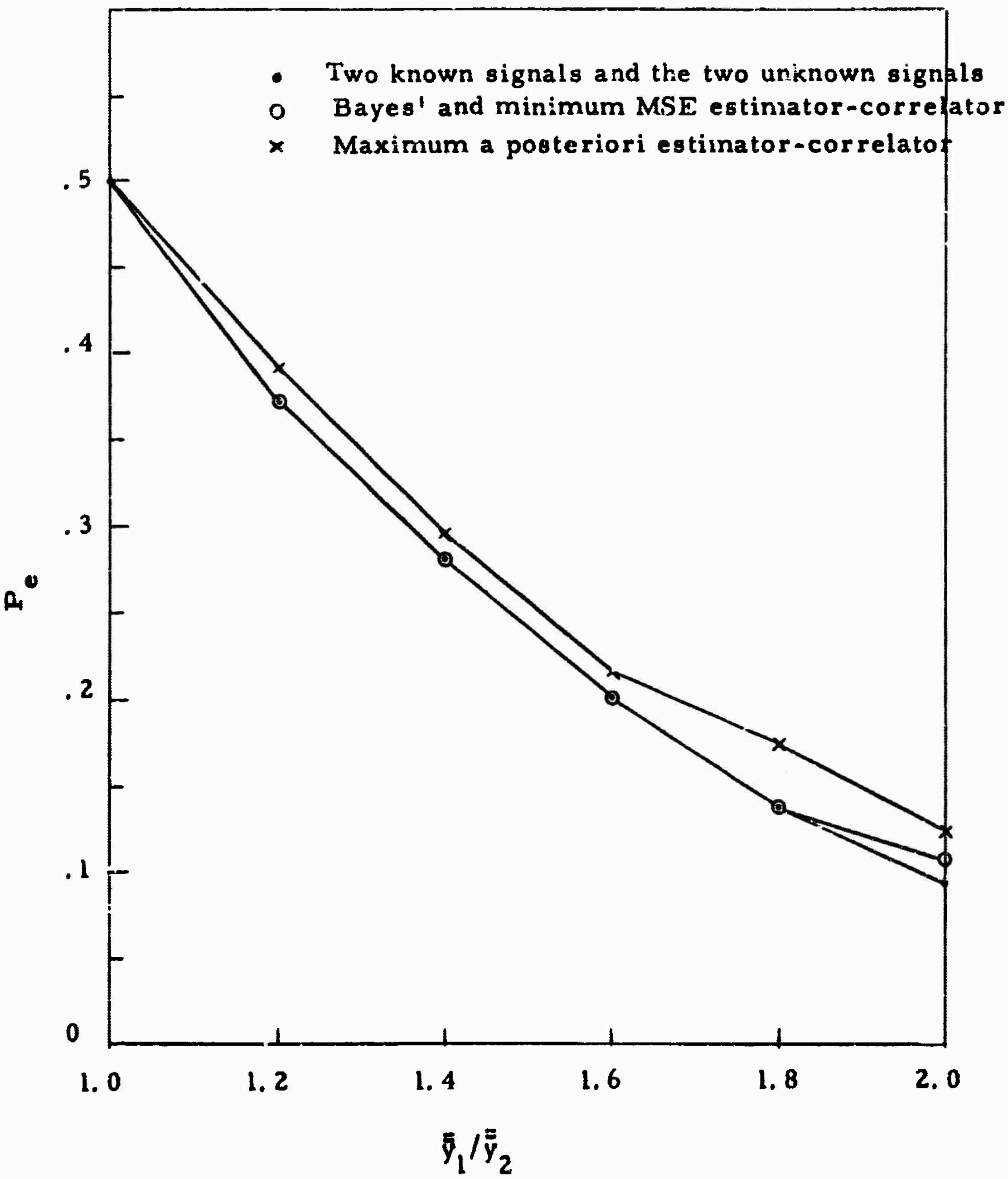

Figure 22. Calculated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 10, $\bar{\bar{y}}_2$ = 10, and $\eta\tau$ = 1.

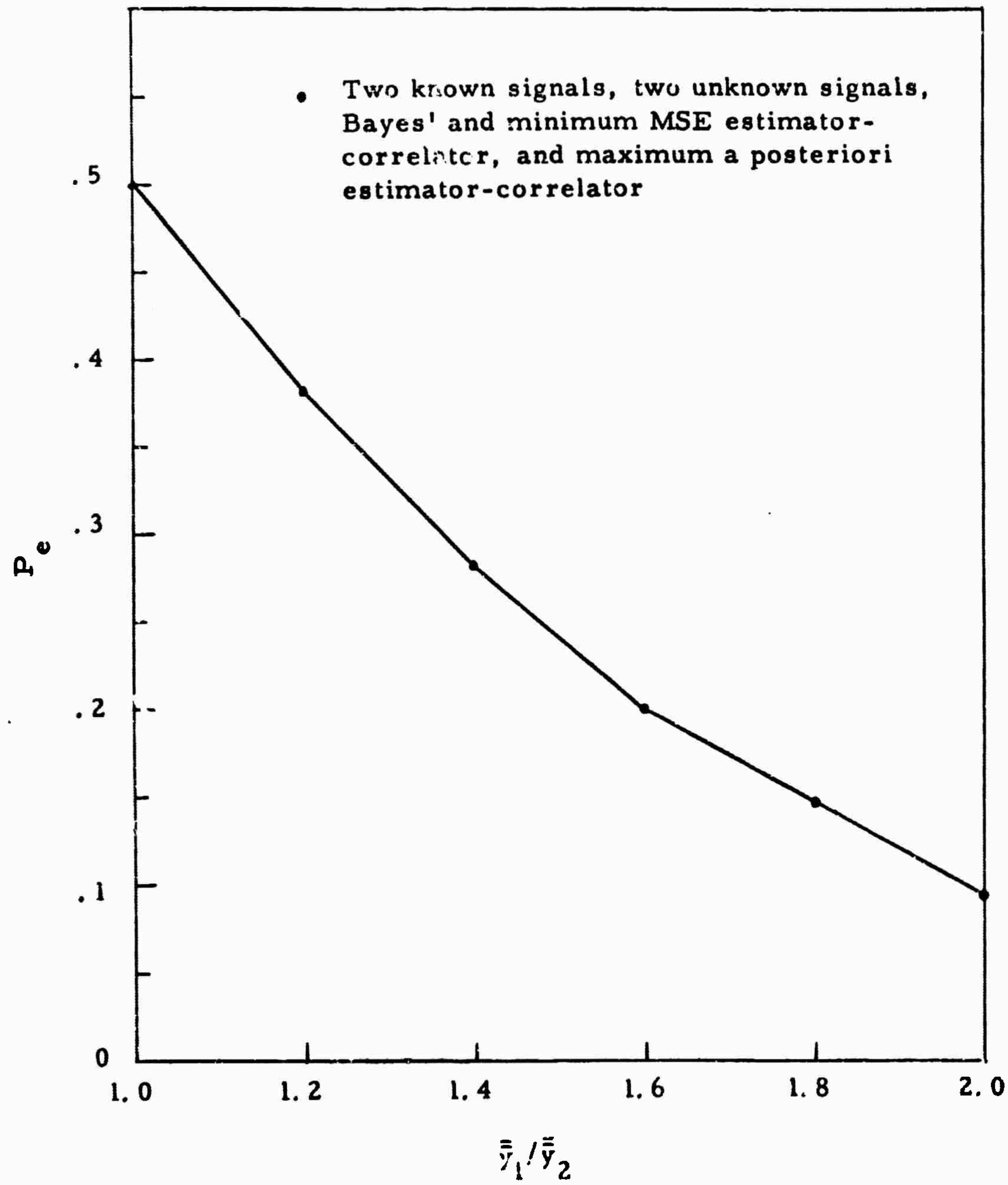

Figure 23. Calculated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 100, $\bar{\bar{y}}_2$ = 10, and $\eta\tau$ = 1.

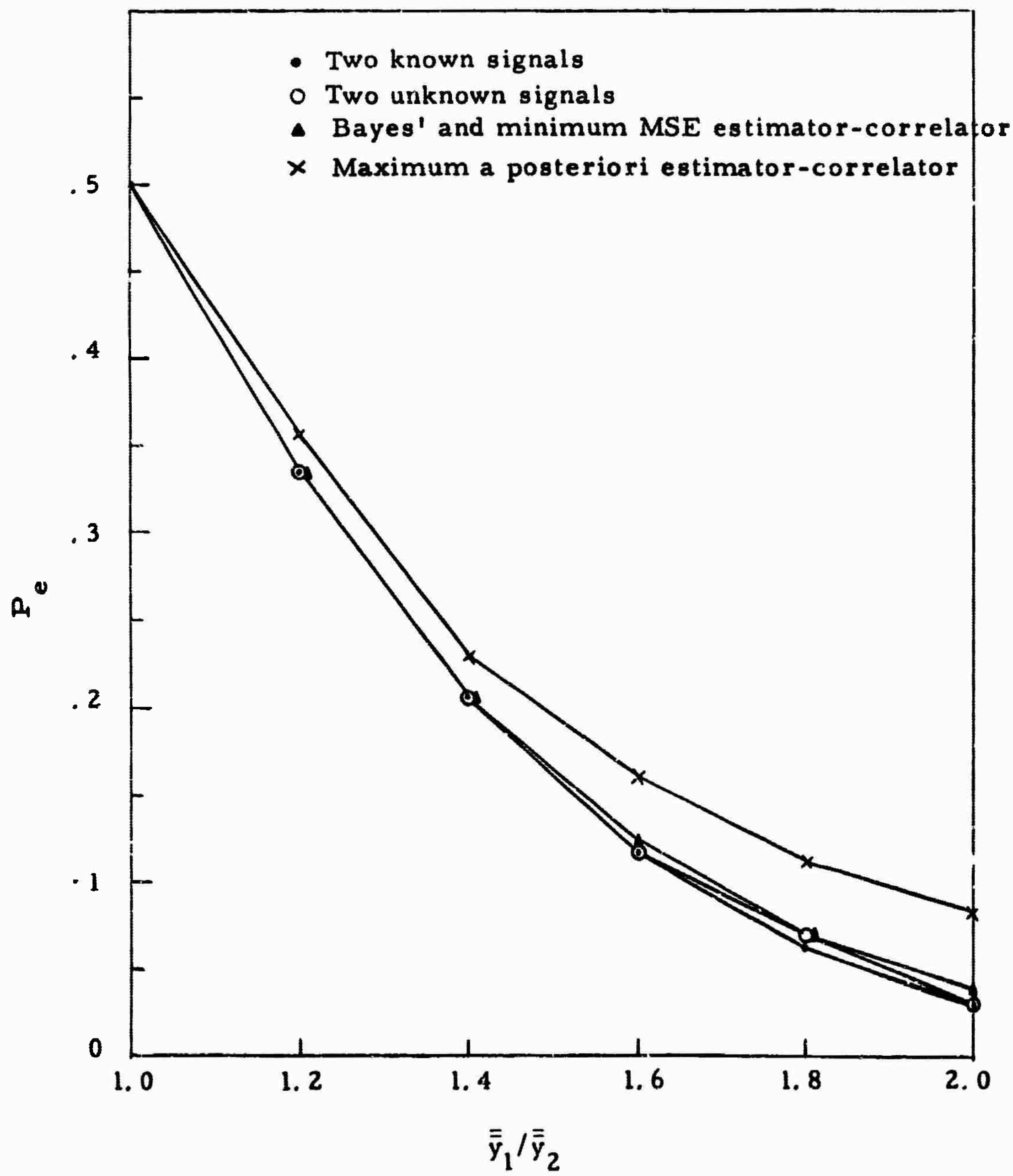

Figure 24. Calculated error probabilities versus ratio of $\bar{\bar{y}}_1$, and $\bar{\bar{y}}_2$ for various detectors, $u = 10$, $\bar{\bar{y}}_2 = 10$, and $\eta\tau = 2$.

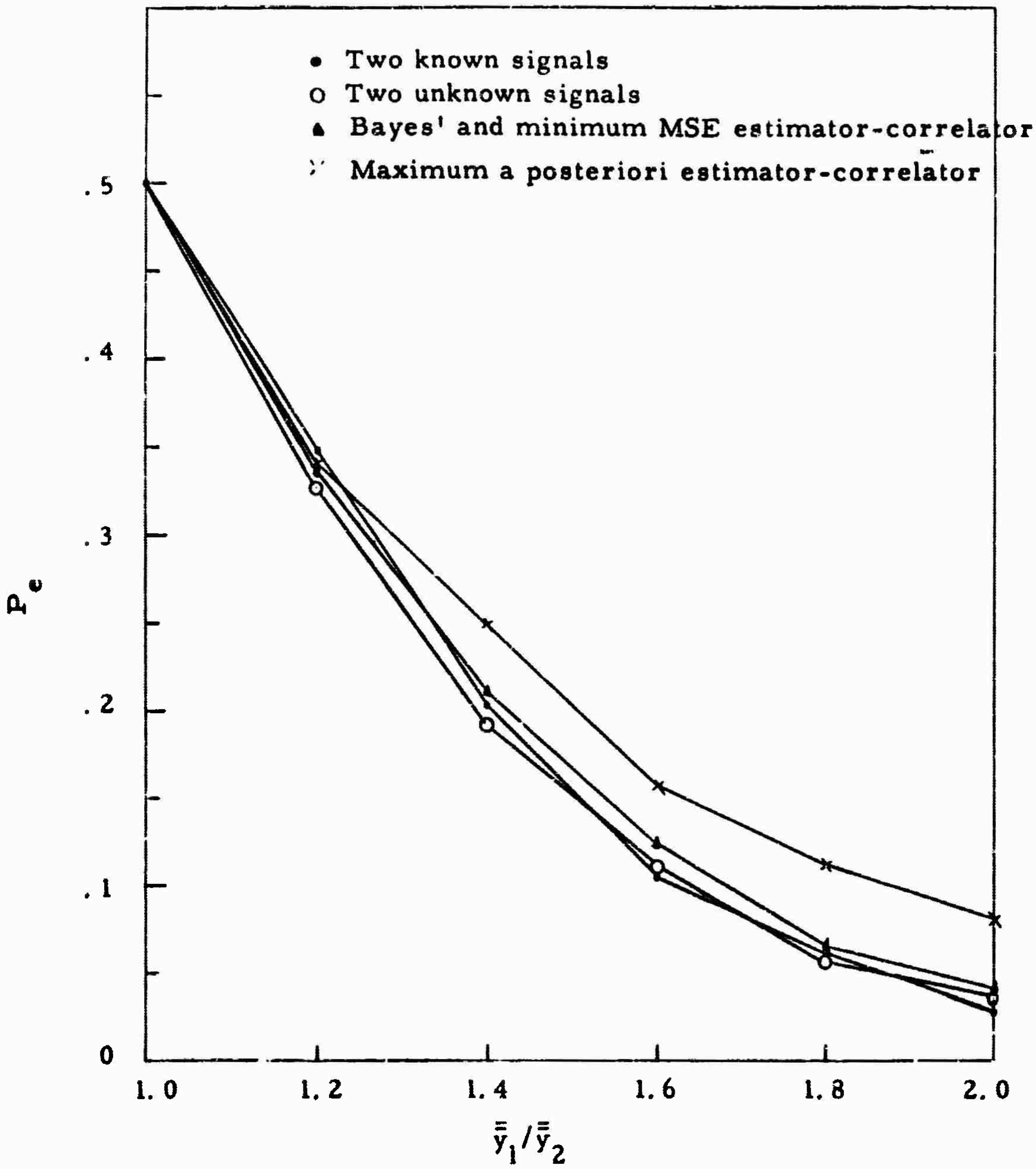

Figure 25. Calculated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 10, $\bar{\bar{y}}_2$ = 10, and $\eta\tau$ = 3.

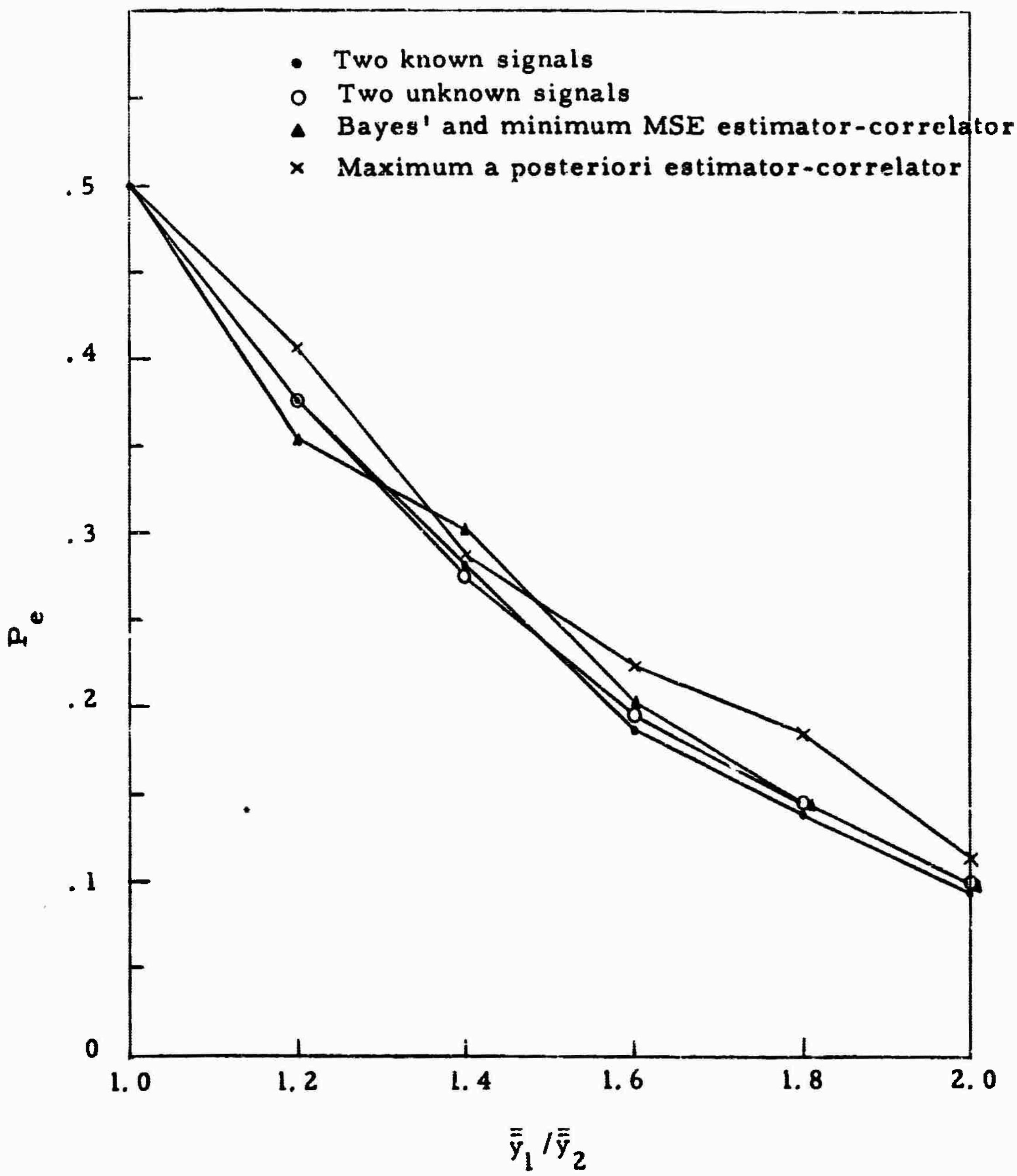

Figure 26. Computer-simulated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, $n = 10$, $\bar{\bar{y}}_2 = 10$, and $\eta\tau = 1$.

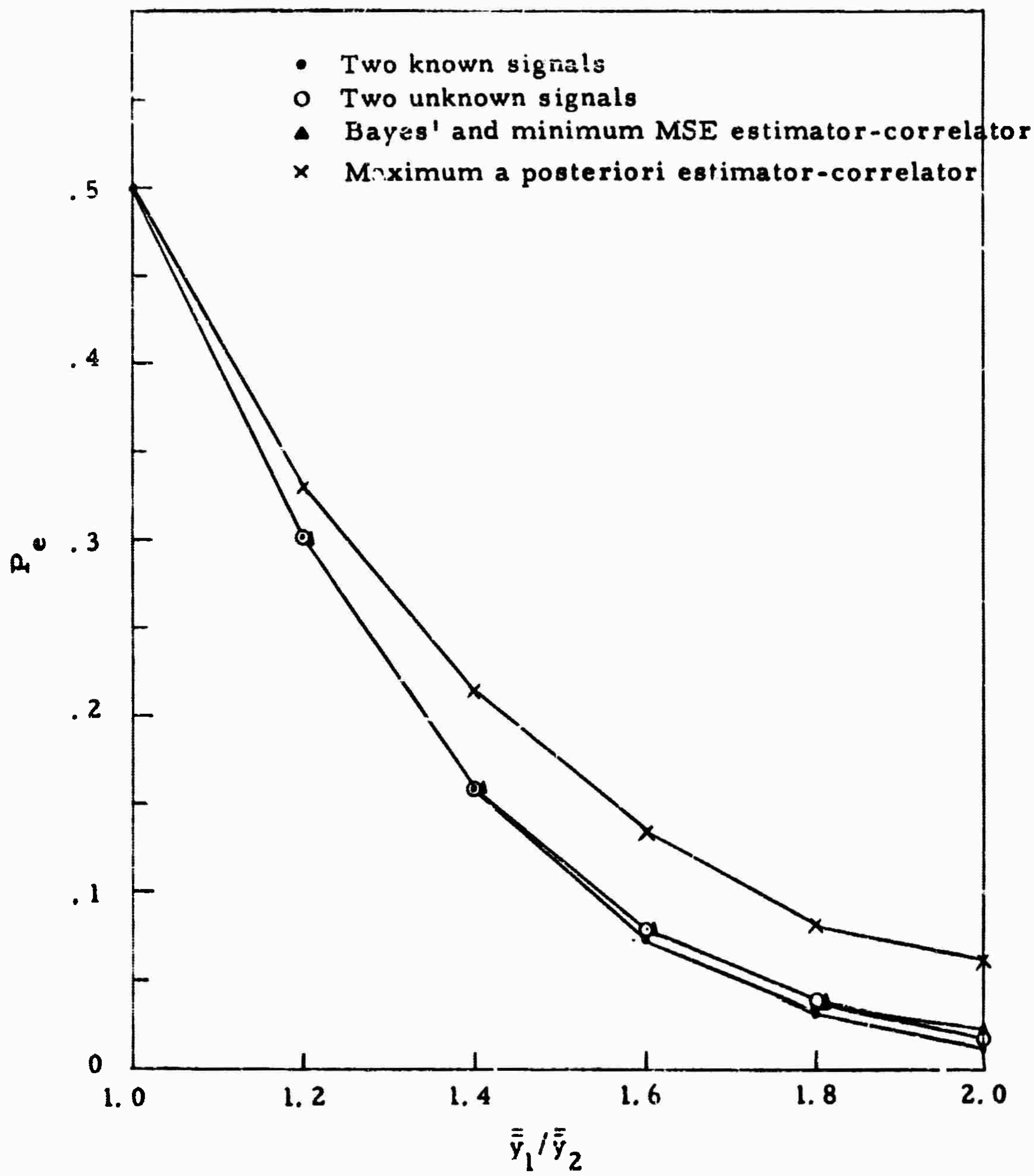

Figure 27. Computer-simulated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 10, $\bar{\bar{y}}_2$ = 10, and $\eta\tau$ = 2.

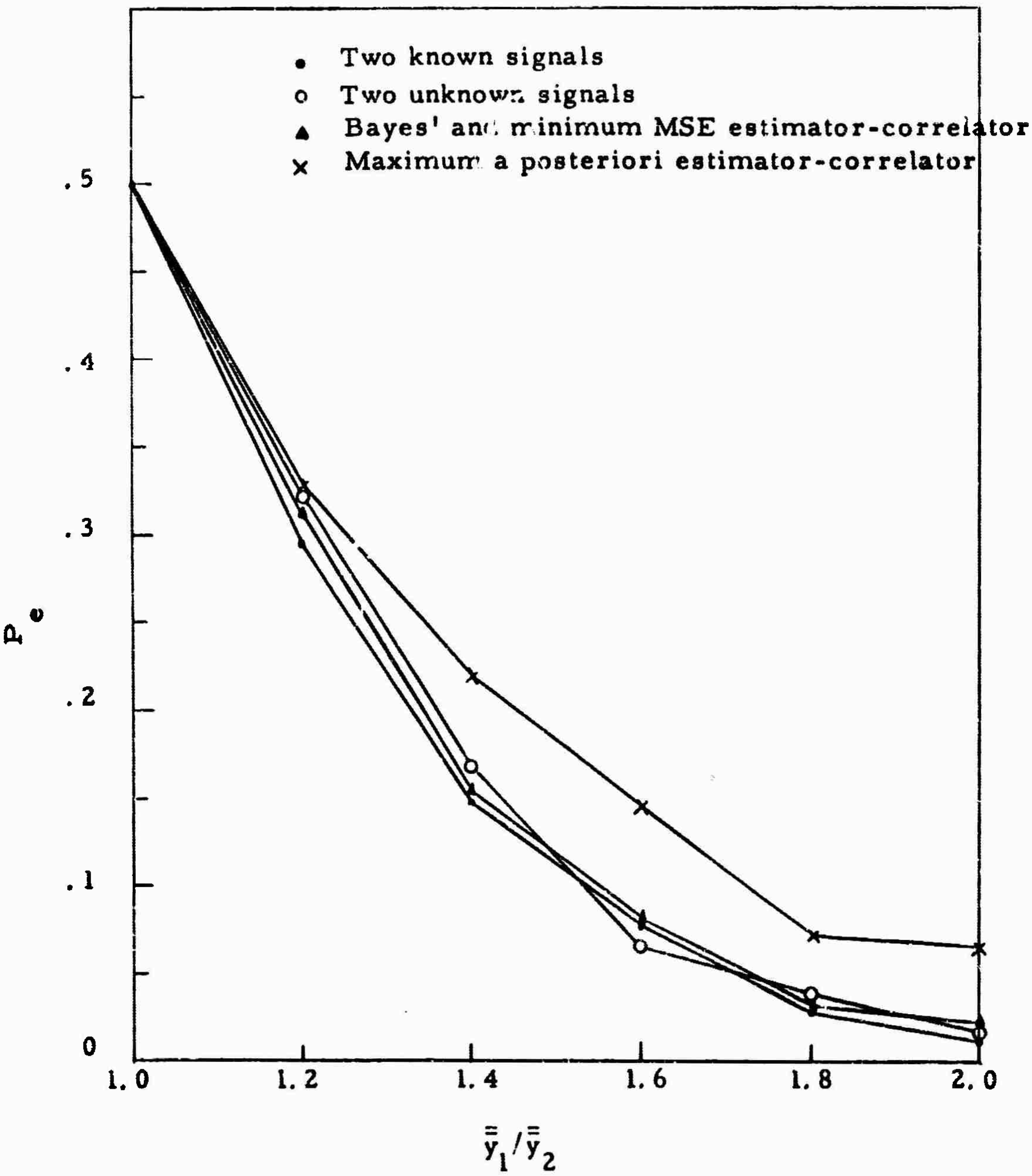

Figure 28. Computer-simulated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{\bar{y}}_2$ for various detectors, u = 10, $\bar{\bar{y}}_2$ = 10, and $\eta\tau$ = 3.

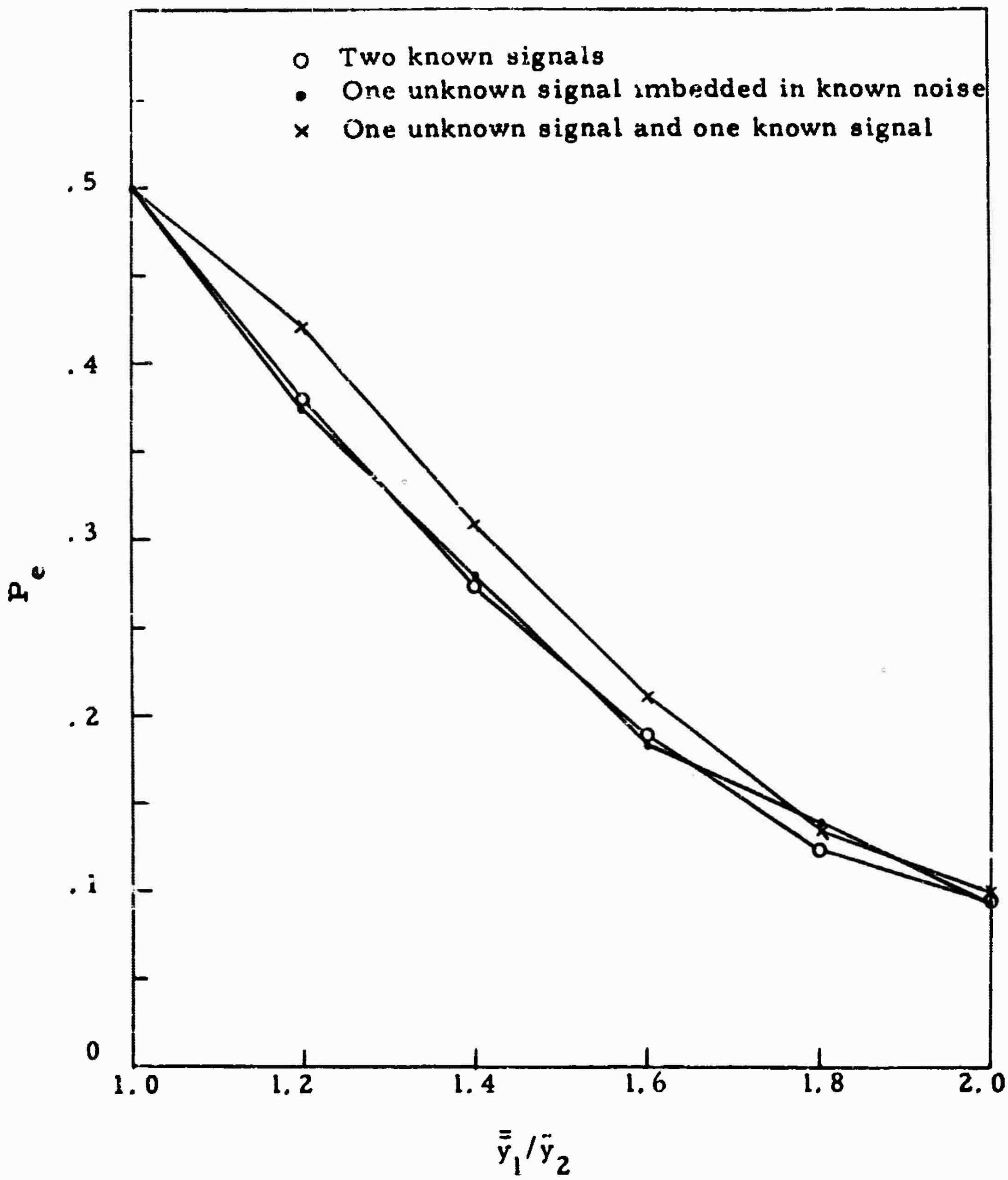

Figure 29. Computer-simulated error probabilities versus ratio of $\bar{\bar{y}}_1$ and $\bar{y}_2$ for both cases of one unknown signal, $u = 10$, $\bar{y}_2 = 10$, and $\eta\tau = 1$.

of the two-known-signals case. One thousand trials were made for each of the simulated results.

In general, if the error probabilities do not become less than about 0.1, it can be seen that knowing the signal parameters (mean rates) is not much better than only knowing their probability distributions, particularly if the Bayes' decision rule or the Bayes estimator-correlator is used. In many problems of signal detection, error probabilities << .1 are important. Errors of this size were more difficult to calculate on the computer because of overflow problems and hence fewer results were obtained for these small errors. For error probabilities of this size, the difference in error probabilities of the various fixed-sample detectors becomes significant (see Table 1).

SEQUENTIAL DETECTION

Introduction

In many problems of optical data processing, detection speed is important. For these situations a sequential detection procedure should considered as an alternative to the conventional fixed-sample detection procedure. The sequential detection test, which was developed by Wald (1947), minimizes the average test length for given P(FD) and P(FA) and hence has a shorter average test length than the fixed-sample test. In the sequential test we introduce two thresholds at the output of the detector such that we declare the "signal present" if one of the thresholds is exceeded and "signal absent" if we fall below the other threshold. The number of observations or test length is not fixed in advance. The number of observations required by the sequential test depends on the outcome of the experiment and is, therefore, not predetermined but a random variable.

Let $p(z/\omega_1)$ be the probability distribution of the observed random variable z when ω_1 is the true state of nature and $p(z/\omega_2)$ be the probability distribution when ω_2 is the true state of nature. If we make j successive observations of z, the probability density for the sample $(z^1, z^2 \ldots, z^j)$ is given by $p_j(z/\omega_1) = p(z^1, \ldots, z^j/\omega_1)$ when ω_1 is the true state of nature and $p_j(z/\omega_2) = p(z^1, \ldots, z^j/\omega_2)$ when ω_2 is the true state of nature. The quantity z^j is a vector whose elements [illegible] output of each of the m cells of the image plane (i.e., [illegible] $z^j_1, z^j_2, \ldots, z$ [illegible]

We need to determine two positive constants A and B, both depending on the error probabilities $P(FD) = \beta$ and $P(FA) = \alpha$ which are specified beforehand. P(FD) is the probability of false-dismissal and P(FA) is the probability of false-alarm. The constants A and B have the approximate realationships (Wald, 1947):

$$A \approx \frac{1 - \beta}{\alpha} \tag{223}$$

and

$$B \approx \frac{\beta}{1 - \alpha} \quad . \tag{224}$$

Bussgang and Middleton (1955) consider the sequential probability ratio test (SPRT) for testing ω_1 against ω_2. The definition of the SPRT for the m cell case is as follows: at each stage of the experiment, compute the probability ratio

$$\ell(j) \triangleq \frac{p_j(z/\omega_1)}{p_j(z/\omega_2)} = \frac{p(z^1, z^2, \ldots, z^j/\omega_1)}{p(z^1, z^2, \ldots, z^j/\omega_2)}, \tag{225}$$

choose ω_1 if $\ell(k) \geq A$,

choose ω_2 if $\ell(k) \leq B$,

continue the experiment by making another observation if $B < \ell(j) < A$.

We are also assuming that the number of electrons being emitted from the different regions or cells of the image plane are independent. Hence, given ω_q, the probability of z_1 electrons being emitted from region 1, ..., z_m electrons from region m all during the j observation intervals is

$$p_j(z/\omega_q) = \prod_{\ell=1}^{m} p(z_\ell/\omega_q) \quad (q = 1, 2) \tag{226}$$

where m is the number of cells in the image plane. The likelihood or probability ratio can then be written as

$$\ell(j) = \frac{p_j(z/\omega_1)}{p_j(z/\omega_2)} = \prod_{\ell=1}^{m} \frac{p_j(z_\ell/\omega_1)}{p_j(z_\ell/\omega_2)}. \tag{227}$$

If the j successive observations are statistically independent, the probability ratio becomes

$$\ell(j) = \frac{p(z^1, \ldots, z^j/\omega_1)}{p(z^1, \ldots, z^j/\omega_2)} = \prod_{i=1}^{j} \prod_{\ell=1}^{m} \frac{p(z_\ell^{\,i}/\omega_1)}{p(z_\ell^{\,i}/\omega_2)}. \tag{228}$$

The natural logarithm of this ratio may be taken for computational convenience. The probability ratio can then be expressed as

$$R_j \triangleq \ln \ell(j) = \sum_{i=1}^{j} \sum_{\ell=1}^{m} \ln \frac{p(z_\ell^{\,i}/\omega_1)}{p(z_\ell^{\,i}/\omega_2)} = \sum_{i=1}^{j} r_i \tag{229}$$

where

$$r_i = \sum_{\ell=1}^{m} \ln \frac{p(z_\ell^{\,i}/\omega_1)}{p(z_\ell^{\,i}/\omega_2)}. \tag{230}$$

The test procedure then becomes:

choose ω_1 if $R_k \geq \ln A$,

choose ω_2 if $R_k \leq \ln B$,

make another observation if $\ln B < R_j < \ln A$.

In our evaluation of this test procedure we are primarily interested in the average number of samples required to terminate the test and the Operating Characteristic Function (OCF). The OCF (L(y)) is defined as the probability of choosing ω_2 when the actual signal is $\bar{y}$. Bussgang and Middleton (1955) show that

$$L(\bar{y}) = \frac{A^h - 1}{A^h - B^h} \tag{231}$$

where h is chosen such that

$$\sum_{z^j=0}^{\infty} \cdots \sum_{z^1=0}^{\infty} \left[\frac{p_j(z/\omega_1)}{p_j(z/\omega_2)} \right]^h p_j(z/\bar{y}) = 1. \tag{232}$$

The quantity $L(\tilde{y})$ is needed in the calculation of the Average Sample Number (ASN). The average value of R_k (see (229)) is equal to the value of the bounds (thresholds $\ell n\ A$ and $\ell n B$) weighted by the probability that they will be reached. That is,

$$E(R_k) = L(\tilde{y})\ \ell n\ B + [1 - L(\tilde{y})]\ \ell n\ A. \tag{233}$$

Also,

$$E(R_k) = E\left[\sum_{\ell=1}^{m} \ell n \frac{p_k(z_\ell/\omega_1)}{p_k(z_\ell/\omega_2)}\right]. \tag{234}$$

For statistically independent observations we have

$$E(R_k) = E\left(\sum_{i=1}^{k} r_i\right) = \bar{k}E(r) \tag{235}$$

where r_i is given in (230).

The ASN for statistically independent observations becomes:

$$\bar{k} = \frac{\beta\ \ell n\ B + (1-\beta)\ \ell n\ A}{E(r)} \tag{236}$$

for ω_1 as the true state of nature,

$$\bar{k} = \frac{\alpha \ \ln A + (1-\alpha) \ \ln B}{E(r)} \tag{237}$$

for ω_2 as the true state of nature, and

$$\bar{k} = \frac{L(\bar{y}) \ \ln A + [1-L(\bar{y})] \ \ln B}{E(r)} \tag{238}$$

for general signal values $\bar{y}$. When the observations are not statistically independent, the ASN is found by solving for $\bar{k}$ from the following:

$$\beta \ \ln B + (1-\beta) \ \ln A = E\left[\sum_{\ell=1}^{m} \ln \frac{p_k(z_\ell/\omega_1)}{p_k(z_\ell/\omega_2)}\right] \tag{239}$$

for ω_1 as the true state of nature.

$$\alpha \ \ln A + (1-\alpha) \ \ln B = E\left[\sum_{\ell=1}^{m} \ln \frac{p_k(z_\ell/\omega_1)}{p_k(z_\ell/\omega_2)}\right] \tag{240}$$

for ω_2 as the true state of nature, and

$$L(\bar{y})\ \ln A + [1-L(\bar{y})]\ \ln B = E\left[\sum_{\ell=1}^{m} \ln \frac{p_k(z_\ell/\omega_1)}{p_k(z_\ell/\omega_2)}\right] \tag{241}$$

for general signal values $\bar{y}$.

Two Known Signals

We will first consider two known signals $\bar{y}_1$ and $\bar{y}_2$ as the possible input signals to the detector where $\bar{y}_1$ and $\bar{y}_2$ are the average number of photons that are incident upon the image plane or detector. The signals $\bar{y}_1$ and $\bar{y}_2$ contain any constant background noise that may be present. For this case we have the two states of nature:

$$\omega_1: \ \bar{y} = \bar{y}_1 \text{ (known)},$$

$$\omega_2: \ \bar{y} = \bar{y}_2 \text{ (known)},$$

where $\bar{y}_1 > \bar{y}_2$. The objective of our sequential test procedure is to decide which of the two states of nature, ω_1 or ω_2, is the true state of nature.

The probability of z_ℓ photoelectrons being emitted from the ℓ^{th} cell of the image plane in time τ during the i^{th} observation, given that

ω_q is the true state of nature, is

$$p(z_\ell^{\,i}/\omega_q) = \frac{(\eta\tau\bar{y}_{q\ell})^{z_\ell^i}\, e^{-\eta\tau\bar{y}_{q\ell}}}{z_\ell^i!} \qquad (q = 1, 2) \tag{242}$$

where $\bar{y}_{q\ell}$ is the ℓ^{th} element of signal $\bar{y}_q$ and $z_\ell^{\,i}$ is the actual number of photons emitted from the ℓ^{th} cell during the i^{th} observation.

Using (228) and (242), the likelihood ratio can be written as

$$\ell(j) = \prod_{\ell=1}^{m} \left(\frac{\bar{y}_{1\ell}}{\bar{y}_{2\ell}}\right)^{\sum_{i=1}^{j} z_\ell^i} e^{-\eta\tau j(\bar{y}_{1\ell}-\bar{y}_{2\ell})}. \tag{243}$$

The logarithm of $\ell(j)$ is

$$R_j = \sum_{\ell=1}^{m} \left[\left(\sum_{i=1}^{j} z_\ell^i\right) \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - \eta\tau j\,(\bar{y}_{1\ell}-\bar{y}_{2\ell})\right]. \tag{244}$$

Equation (244) can also be written as

$$R_j = R_{j-1} + \sum_{\ell=1}^{m} \left[z_\ell^j \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - \eta\tau(\bar{y}_{1\ell}-\bar{y}_{2\ell})\right] \tag{245}$$

where $R_o = 0$.

The SPRT procedure is:

choose ω_1 if $R_k \geq \ln A$,

choose ω_2 if $R_k \leq \ln B$,

continue testing if $\ln B < R_j < \ln A$.

In order to find the ASN for this test we must solve the following equation:

$$E(R_k) = L(\bar{y}) \ln B + [1-L(\bar{y})] \ln A \qquad (246)$$

Using (244) we have

$$E(R_k) = \bar{k}\eta\tau \sum_{\ell=1}^{m} [\bar{y}_\ell \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - (\bar{y}_{1\ell} - \bar{y}_{2\ell})]. \qquad (247)$$

Hence, for the ASN we have

$$\bar{k}(\bar{y}) = \frac{L(\bar{y}) \ln B + [1-L(\bar{y})] \ln A}{\eta\tau \sum_{\ell=1}^{m} [\bar{y}_\ell \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - (\bar{y}_{1\ell} - \bar{y}_{2\ell})]} \qquad (248)$$

for signal $\bar{y}$ in general.

$$\bar{k}(\bar{y}_1) = \frac{\beta \ln B + (1-\beta) \ln A}{\eta\tau \sum_{\ell=1}^{m} [\bar{y}_{1\ell} \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - \bar{y}_{1\ell} + \bar{y}_{2\ell}]} \tag{249}$$

for $\bar{y}_1$ present, and

$$\bar{k}(\bar{y}_2) = \frac{\alpha \ln A + (1-\alpha) \ln B}{\eta\tau \sum_{\ell=1}^{m} [\bar{y}_{2\ell} \ln(\bar{y}_{1\ell}/\bar{y}_{2\ell}) - \bar{y}_{1\ell} + \bar{y}_{2\ell}]} \tag{250}$$

for y_2 present.

Consider the single cell case (i.e., m = 1) where $\bar{y}_1$, $\bar{y}_2$, $\bar{y}$, z, and k are scalars. For this case

$$R_j = \left(\sum_{i=1}^{j} z^i\right) \ln(\bar{y}_1/\bar{y}_2) - \eta\tau j(\bar{y}_1 - \bar{y}_2), \tag{251}$$

$$I_j = R_j - R_{j-1} = z^j \ln(\bar{y}_1/\bar{y}_2) - \eta\tau(\bar{y}_1 - \bar{y}_2), \tag{252}$$

and

$$\bar{k}(\bar{y}) = \frac{L(\bar{y})\,\ln B + [1-L(\bar{y})]\,\ln A}{\eta\tau[\bar{y}\,\ln(\bar{y}_1/\bar{y}_2)-\bar{y}_1+\bar{y}_2]}\,. \tag{253}$$

The quantity $L(\bar{y})$ is found from

$$L(\bar{y}) = \frac{A^h - 1}{A^h - B^h} \tag{254}$$

where h is chosen such that

$$\sum_{z=0}^{\infty}\left[\frac{p(z/\omega_1)}{p(z/\omega_2)}\right]^h p(z/\bar{y}) = \sum_{z=0}^{\infty}\left[(\bar{y}_1/\bar{y}_2)^z e^{-\eta\tau(\bar{y}_1-\bar{y}_2)}\right]^h \frac{(\eta\tau\bar{y})^z e^{-\eta\tau\bar{y}}}{z!} = 1. \tag{255}$$

This equation can be solved to yield

$$\bar{y} = \frac{(\bar{y}_1-\bar{y}_2)h}{[(\bar{y}_1/\bar{y}_2)^h - 1]}\,. \tag{256}$$

By choosing values of h we can plot $L(\bar{y})$ versus $\bar{y}$. Using the corresponding values of $L(\bar{y})$ and $\bar{y}$ we can determine and plot $\bar{k}(\bar{y})$ versus $\bar{y}$. For the special case of $h = 0$

$$L(\bar{y}) = \ln A / \ln (A/B) \tag{257}$$

and

$$\bar{k}(\bar{y}) = \frac{-\ln A \ \ln B}{E\left[[z \ \ln(\bar{y}_1/\bar{y}_2) - \eta\tau(\bar{y}_1 - \bar{y}_2)]^2 \right]}$$

$$= \frac{-\ln A \ \ln B}{\left[[\mathrm{var}(z)+E^2(z)]\ln^2(\bar{y}_1/\bar{y}_2) - E(z) 2 \ln(\bar{y}_1/\bar{y}_2)\eta\tau(\bar{y}_1 - \bar{y}_2) + (\eta\tau)^2(\bar{y}_1 - \bar{y}_2)^2 \right]} \tag{258}$$

An example is shown in Figure 30 of the calculated ASN versus actual signal values $\bar{y}$ for a sequential detector designed for two particular signal mean rates whose preassigned values of α and β are equal. Figure 31 shows the OCF for the same example and sequential detector (OCF is defined as the probability of choosing state ω_2 to be present when the actual signal is $\bar{y}$). Figures 32 and 33 are simulated results corresponding to Figures 30 and 31 respectively. As the ASN becomes larger the calculated and simulated results compare more favorably. This is due to the approximation of the thresholds A and B, which becomes more accurate as the ASN becomes large. Figures 34 and 35 show calculated ASN and OCF versus actual signal values $\bar{y}$ for a sequential detector designed for two particular signal mean rates whose preassigned values of α and β are not equal. For this example the detector is inclined to make decisions in favor of signal $\bar{y}_2$ being present since for $\beta > \alpha$ the detector guards less against false dismissal (saying state ω_2 is present when ω_1 is the true

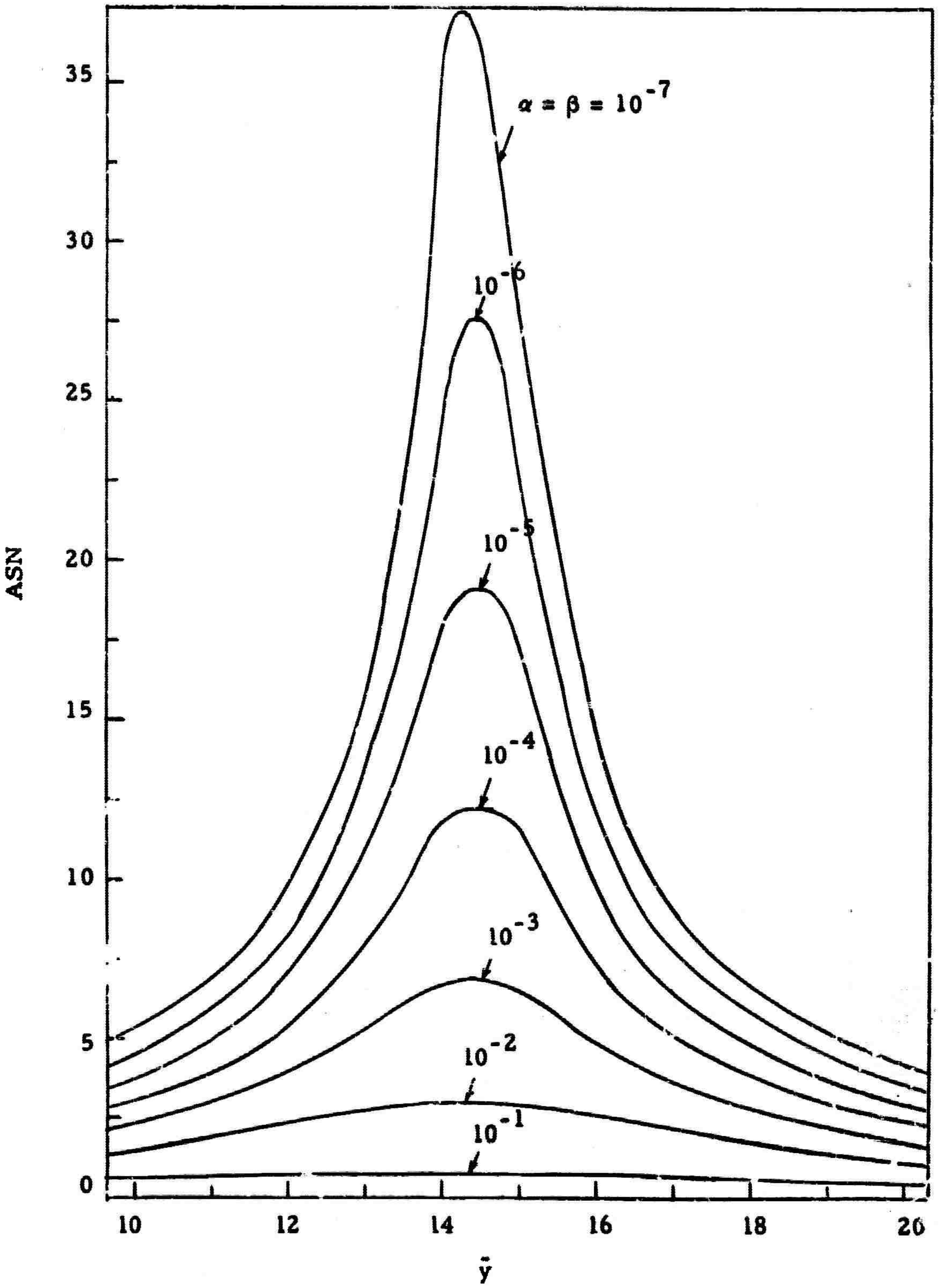

Figure 30. Calculated ASN versus actual signal values $\bar{y}$ for a sequential detector designed for two known signals, various values of $\alpha = \beta$, $\bar{y}_1 = 20$, and $\bar{y}_2 = 10$.

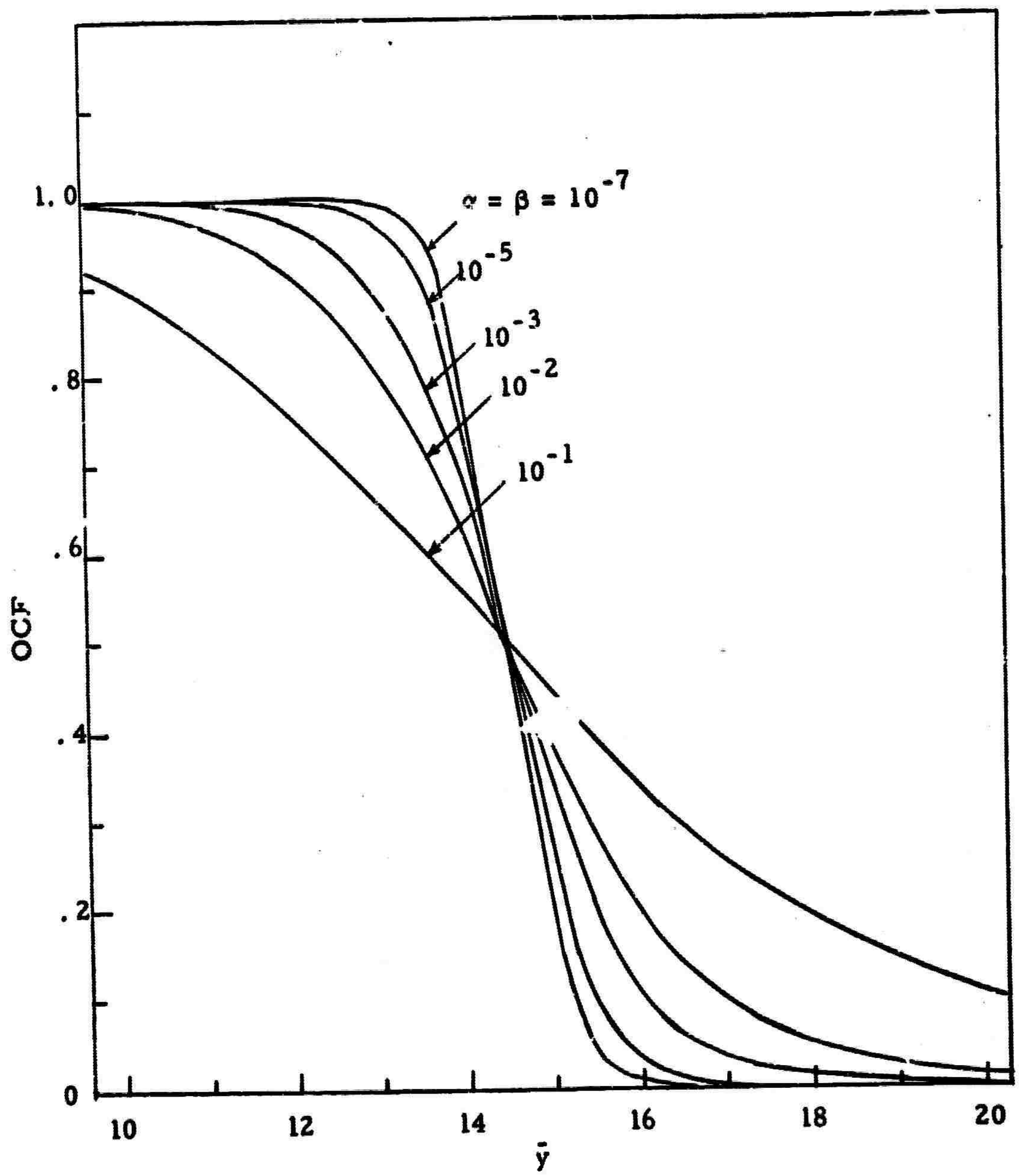

Figure 31. Calculated OCF versus actual signal values $\bar{y}$ for a sequential detector designed for two known signals, various values of $\alpha = \beta$, $\bar{y}_1 = 20$, and $\bar{y}_2 = 10$.

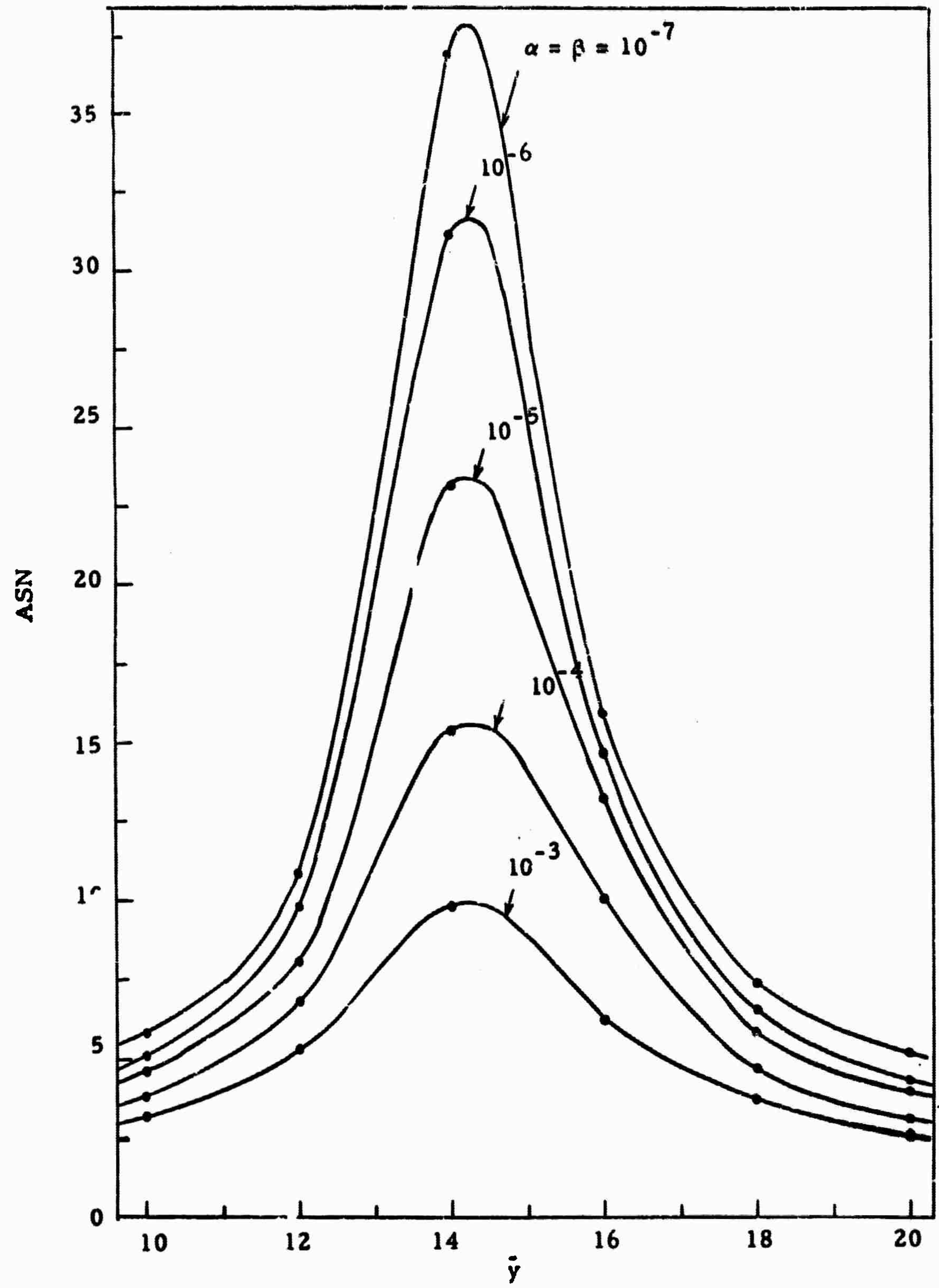

Figure 32. Computer-simulated ASN versus actual signal values $\bar{y}$ for a sequential detector designed for two known signals, various values of $\alpha = \beta$, $\bar{y}_1 = 20$, and $\bar{y}_2 = 10$.

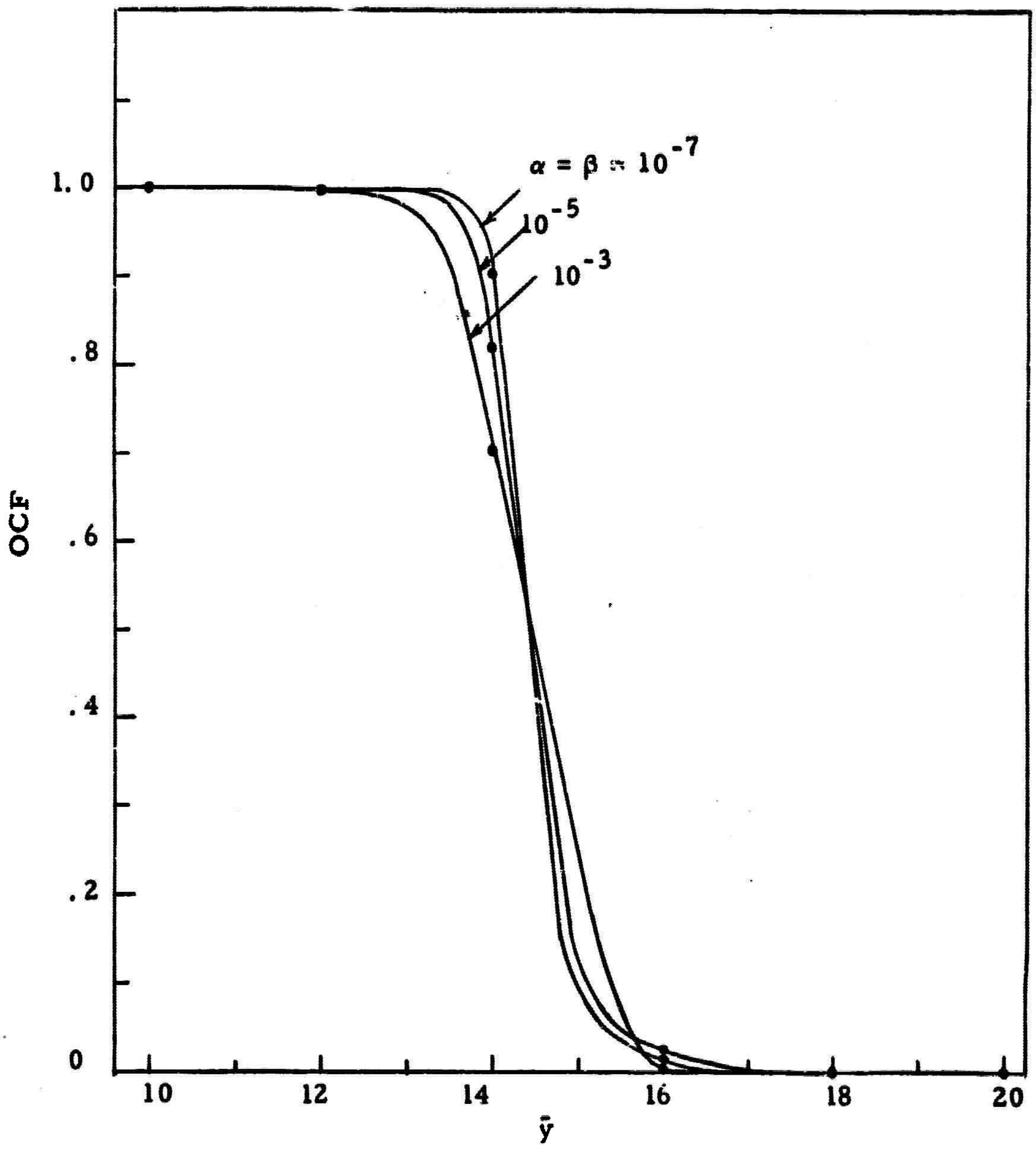

Figure 33. Computer-simulated OCF versus actual signal values $\bar{y}$ for a sequential detector designed for two known signals, various values at $\alpha = \beta$, $\bar{y}_1 = 20$, and $\bar{y}_2 = 10$.

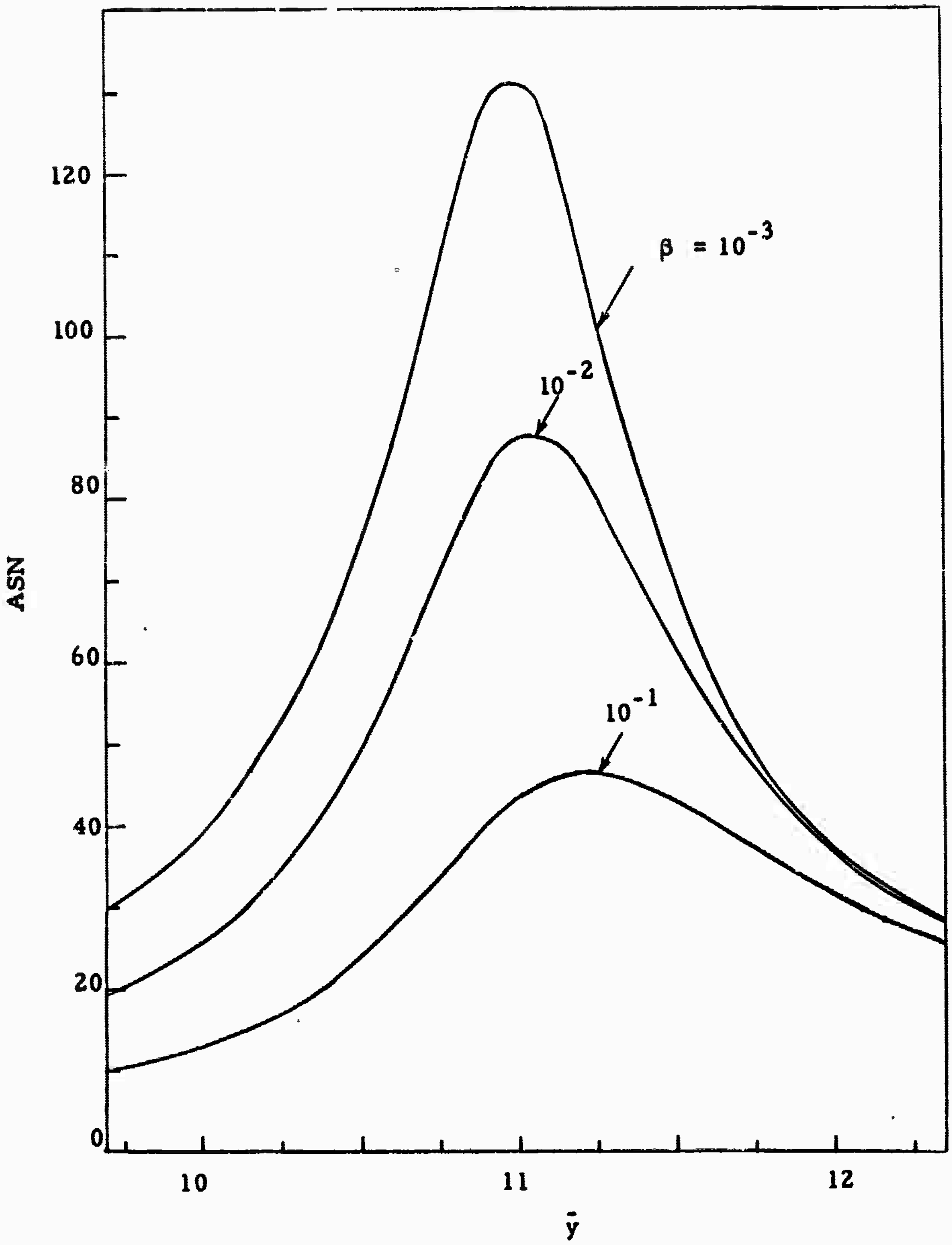

Figure 34. Calculated ASN versus actual signal values $\bar{y}$ for a sequential designed for two known signals, various values of β, $\alpha = 10^{-3}$, $\bar{y}_1 = 12$, and $\bar{y}_2 = 10$.

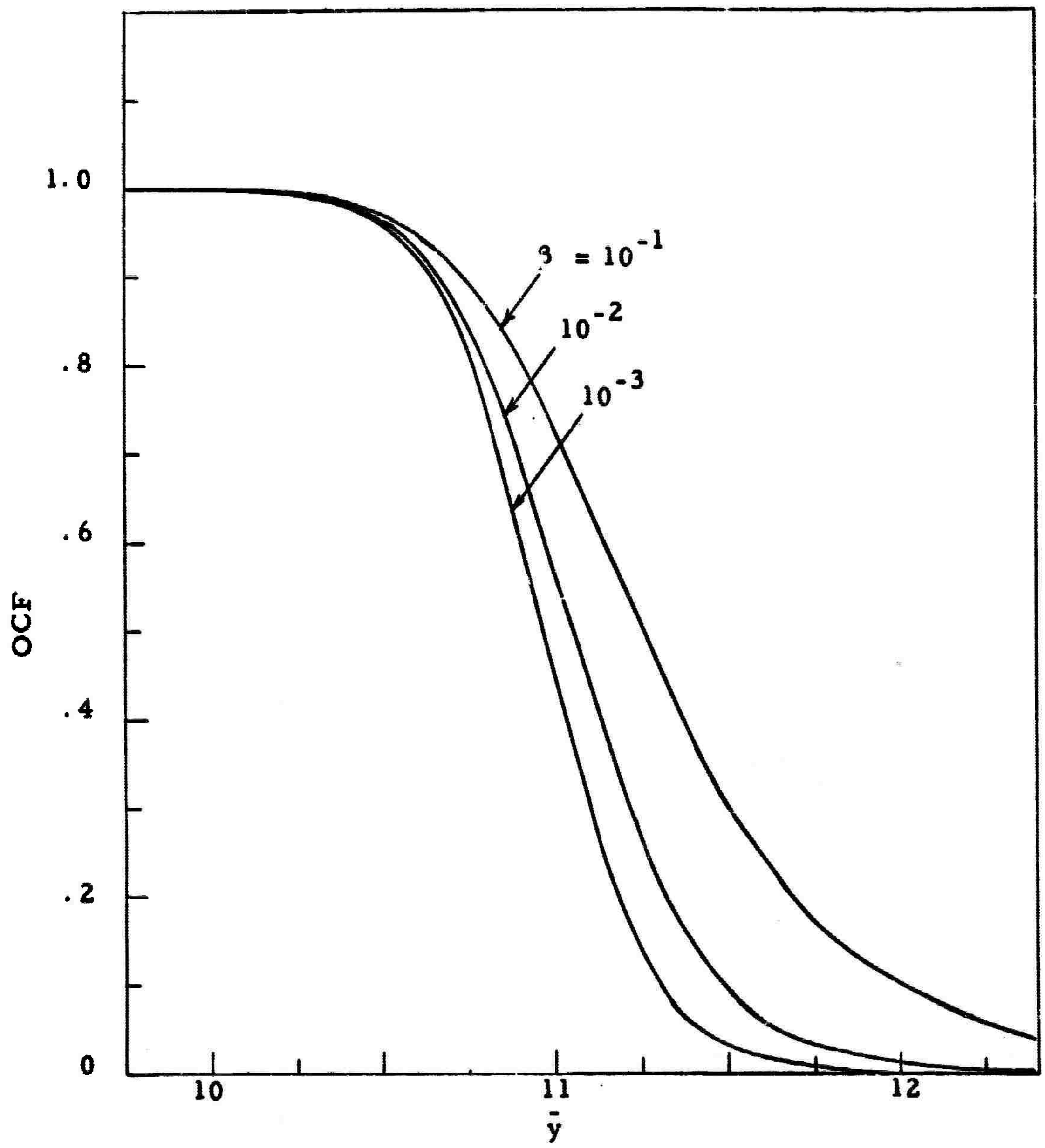

Figure 35. Calculated OCF versus actual signal values $\bar{y}$ for a sequential detector designed for two known signal, various values of β, $\alpha = 10^{-3}$, $\bar{y}_1 = 12$, and $\bar{y}_2 = 10$.

state of nature) than it does against false alarm (saying state ω_1 is present when ω_2 is the true state of nature). One hundred trials were m ade for each of the simulated results of Figures 32 and 33.

The case just discussed may, as a special case, be considered as having two known signals $\bar{s}_1$ and $\bar{s}_2$ imbedded in known noise $\bar{n}$. The two states of nature for this case are:

$$\omega_1: \bar{y} = \bar{s}_1 + \bar{n} \text{ (known signal plus noise),}$$

$$\omega_2: \bar{y} = \bar{s}_2 + \bar{n} \text{ (known signal plus noise),}$$

where $\bar{s}_1 > \bar{s}_2$. The values of $\bar{s}_2$ may or may not be zero.

Using these conditions for the single cell case we have:

$$R_j = \left(\sum_{i=1}^{j} z^i\right) \ln\left(\frac{\bar{s}_1+\bar{n}}{\bar{s}_2+\bar{n}}\right) - \pi j(\bar{s}_1-\bar{s}_2), \tag{259}$$

$$I_j = z^j \ln\left(\frac{\bar{s}_1+\bar{n}}{\bar{s}_2+\bar{n}}\right) - \pi (\bar{s}_1-\bar{s}_2), \tag{260}$$

$$\bar{k} = \frac{L(\bar{s}) \ln B + [1-L(\bar{s})] \ln A}{\pi \left[\bar{s} \ln\left(\frac{\bar{s}_1+\bar{n}}{\bar{s}_2+\bar{n}}\right) - (\bar{s}_1-\bar{s}_2)\right]}, \tag{261}$$

$$L(\bar{s}) = \frac{A^h - 1}{A^h - B^h} . \quad (262)$$

and

$$\bar{s} = \frac{(\bar{s}_1 - \bar{s}_2)\, h}{\left[\left(\frac{\bar{s}_1 + \bar{n}}{\bar{s}_2 + \bar{n}}\right)^h - 1\right]} . \quad (263)$$

Two Unknown Signals

Assume that we are to discriminate between two signals $\bar{y}_1$ and $\bar{y}_2$ both of which contain any background noise that may be present. These signals are fixed but unknown, and each of their elements are taken from statistically independent gamma distributions with known parameters:

$$f(\bar{y}_q) = \prod_{\ell=1}^{m} f(\bar{y}_{q\ell}) = \prod_{\ell=1}^{m} \frac{\alpha_{q\ell}\, (\alpha_{q\ell}\bar{y}_{q\ell})^{u_{q\ell}-1}\, e^{-\alpha_{q\ell}\bar{y}_{q\ell}}}{\Gamma(u_{q\ell})} \quad (q=1,2),\ \bar{y}_{q\ell} \geq 0 \quad (264)$$

$$= 0, \qquad \text{otherwise.}$$

Our two states of nature are:

$$\omega_1 : \; \bar{y} = \bar{y}_1 \quad \text{(unknown)},$$

$$\omega_2 : \; \bar{y} = \bar{y}_2 \quad \text{(unknown)},$$

where $\bar{\bar{y}}_1 > \bar{\bar{y}}_2$. We want to apply the SPRT procedure to decide which of these two states is present.

The conditional probability distribution for the observed random variable z is

$$P_j(z/\omega_q, \bar{y}_q) = \prod_{\ell=1}^{m} \prod_{i=1}^{j} \frac{(\eta\tau\bar{y}_{q\ell})^{z_\ell^i}\, e^{-\eta\tau\bar{y}_{q\ell}}}{z_\ell^i\,!} \quad (q=1,\ 2). \tag{265}$$

The likelihood ratio becomes

$$\ell(j) = \frac{\int_0^\infty P_j(z/\omega_1,\ \bar{y}_1)\ f(\bar{y}_1)\ d\bar{y}_1}{\int_0^\infty P_j(z/\omega_2,\ \bar{y}_2)\ f(\bar{y}_2)\ d\bar{y}_2}$$

$$= \prod_{\ell=1}^{m} \frac{\alpha_{1\ell}^{u_{1\ell}}\,\Gamma(u_{2\ell})\,\Gamma\left(\sum_{i=1}^{j} z_\ell^i + u_{1\ell}\right)\ (\alpha_{2\ell}+j\eta\tau)^{\sum_{i=1}^{j} z_\ell^i + u_{2\ell}}}{\alpha_{2\ell}^{u_{2}}\,\Gamma(u_{1\ell})\,\Gamma\left(\sum_{i=1}^{j} z_\ell^i + u_{2\ell}\right)\ (\alpha_{1\ell}+j\eta\tau)^{\sum_{i=1}^{j} z_\ell^i + u_{1\ell}}} \;. \tag{266}$$

Taking the logarithm of this equation yields

$$R_j = \ln \ell(j) = \sum_{\ell=1}^{m} \ln \left[\frac{\alpha_{1\ell}^{u_{1\ell}} \Gamma(u_{2\ell}) \Gamma\left(\sum_{i=1}^{j} z_\ell^i + u_{1\ell}\right) (\alpha_{2\ell} + j\eta\tau)^{\sum_{i=1}^{j} z_\ell^i + u_{2\ell}}}{\alpha_{2\ell}^{u_2} \Gamma(u_{1\ell}) \Gamma\left(\sum_{i=1}^{j} z_\ell^i + u_{2\ell}\right) (\alpha_{1\ell} + j\eta\tau)^{\sum_{i=1}^{j} z_\ell^i + u_{1\ell}}} \right]. \quad (267)$$

The SPRT procedure is then:

choose ω_1 if $R_k \geq \ln A$,

choose ω_2 if $R_k \leq \ln A$,

continue testing if $\ln B < R_j < \ln A$.

Consider the single cell case (i. e., $m = 1$) where $\bar{y}_1$, $\bar{y}_2$, $\bar{y}$, z, and k are all scalars. Assume that $u_1 = u_2 = u$. For these conditions we have

$$R_j = \left(\sum_{i=1}^{j} z^i\right) \ln \left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) + u \left[\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) - \ln\left(\frac{\alpha_2}{\alpha_1}\right) \right]. \quad (268)$$

Also,

$$I_j = R_j - R_{j-1} = \left(\sum_{i=1}^{j-1} z^i + u\right) \left[\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) - \ln\left(\frac{\alpha_2 + j\eta\tau - \eta\tau}{\alpha_1 + j\eta\tau - \eta\tau}\right) \right] + z^j \left(\frac{\alpha_2 + j\eta\tau}{u_1 + j\eta\tau}\right). \quad (269)$$

We can now apply the SPRT procedure to discriminate between the two states of nature, ω_1 and ω_2.

We now attempt to find the ASN, $\bar{k}$, from

$$E(R_k) = L(\bar{y}) \ln B + [1-L(\bar{y})] \ln A. \tag{270}$$

If we assume that $E(R_k) \approx E(R_{\bar{k}})$ we can write

$$E(R_k) \approx \ln\left(\frac{\alpha_2+\bar{k}\eta\tau}{\alpha_1+\bar{k}\eta\tau}\right)[\bar{k}\eta\tau E(\bar{y})+u] - u \ln(\alpha_2/\alpha_1). \tag{271}$$

Thus we have

$$[\bar{k}\eta\tau E(\bar{y})+u] \ln\left(\frac{\alpha_2+\bar{k}\eta\tau}{\alpha_1+\bar{k}\eta\tau}\right) \approx L(\bar{y}) \ln B + [1-L(\bar{y})] \ln A + u \ln(\alpha_2/\alpha_1) \tag{272}$$

for general signal values $\bar{y}$,

$$[\bar{k}\eta\tau u/\alpha_1+u] \ln\left(\frac{\alpha_2+\bar{k}\eta\tau}{\alpha_1+\bar{k}\eta\tau}\right) \approx \beta \ln B + (1-\beta) \ln A + u \ln(\alpha_2/\alpha_1) \tag{273}$$

for $\bar{y}_1$ present, and

$$[\bar{k}\eta\tau u/\alpha_2+u]\ \ln\left(\frac{\alpha_2+\bar{k}\eta\tau}{\alpha_1+\bar{k}\eta\tau}\right) = \alpha \ln A + (1-\alpha)\ln B + u\ln(\alpha_2/\alpha_1) \tag{274}$$

for $\bar{y}_2$ present.

These last two equations must be solved by a "cut and try" procedure (or graphically) to obtain $\bar{k}$ (ASN). For general signal values, $\bar{y}$, $L(\bar{y})$ is very difficult, if not impossible to obtain analytically.

Figures 36, 37, and 38 show graphical solutions of ASN for the case of $\bar{y}_1$ present for both the two-known-signals case and the two-unknown-signals case. It can be seen from these figures that for the two-unknown-signals case and small enough values of α and β, given some value of $\bar{R}_j$ ($R_j \triangleq E(\bar{R}_j/\bar{y}, j)$), neither threshold will be crossed regardless of how large ASN becomes.

For general signal values $\bar{y}$ the ASN and $L(\bar{y})$ can be found by computer simulation. Figure 39 shows the simulated ASN versus actual signal values $\bar{y}$ for the two-unknown-signals case that we have just considered. Simulated and calculated ASN for the sequential detector designed for two particular signal mean rates are also shown for comparison. Figure 40 shows the OCF for these cases. One thousand trials were made for each of the simulated results of Figures 39 and 40.

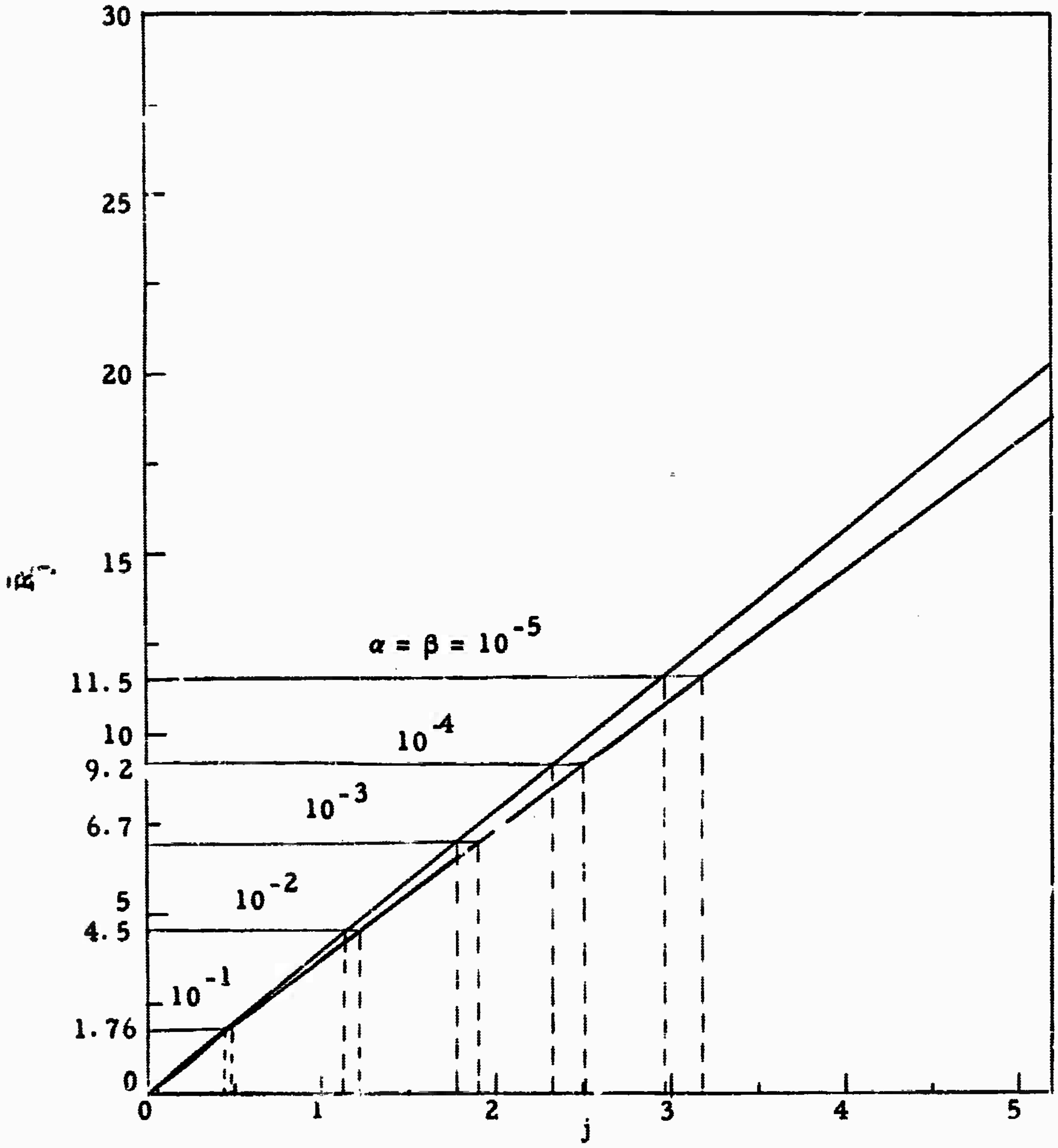

Figure 36. Calculated mean cumulative information versus sample number for sequential detectors designed for the two-known-signals case and the two-unknown signals case for u = 1000, $\bar{y}_1 = 20$, $\bar{\bar{y}}_2 = 10$, and $\bar{y} = 20$.

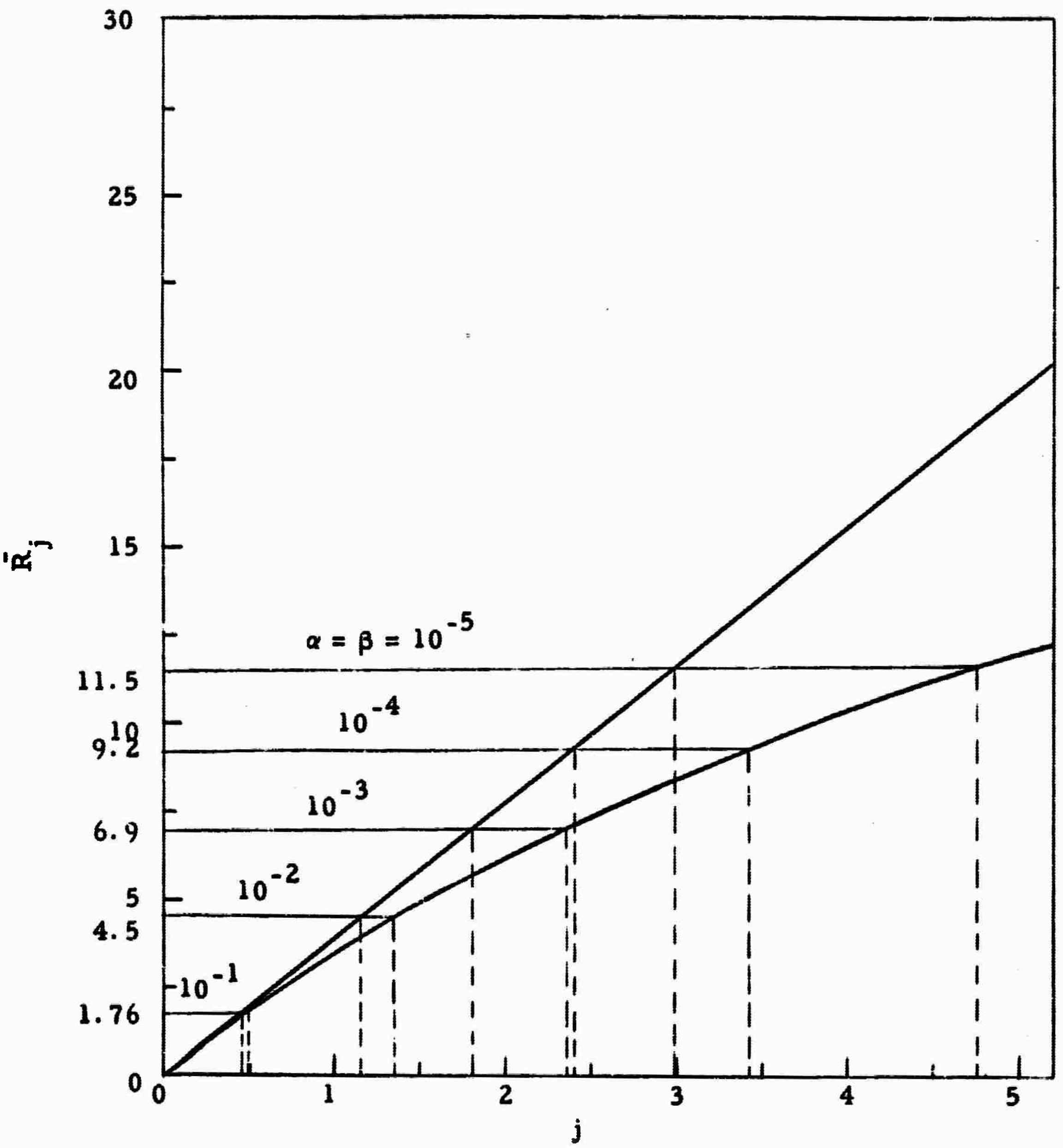

Figure 37. Calculated mean cumulative information versus sample number for sequential detectors designed for the two-known-signals case and the two-unknown-signals case for u = 100, $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, and $\bar{y} = 20$.

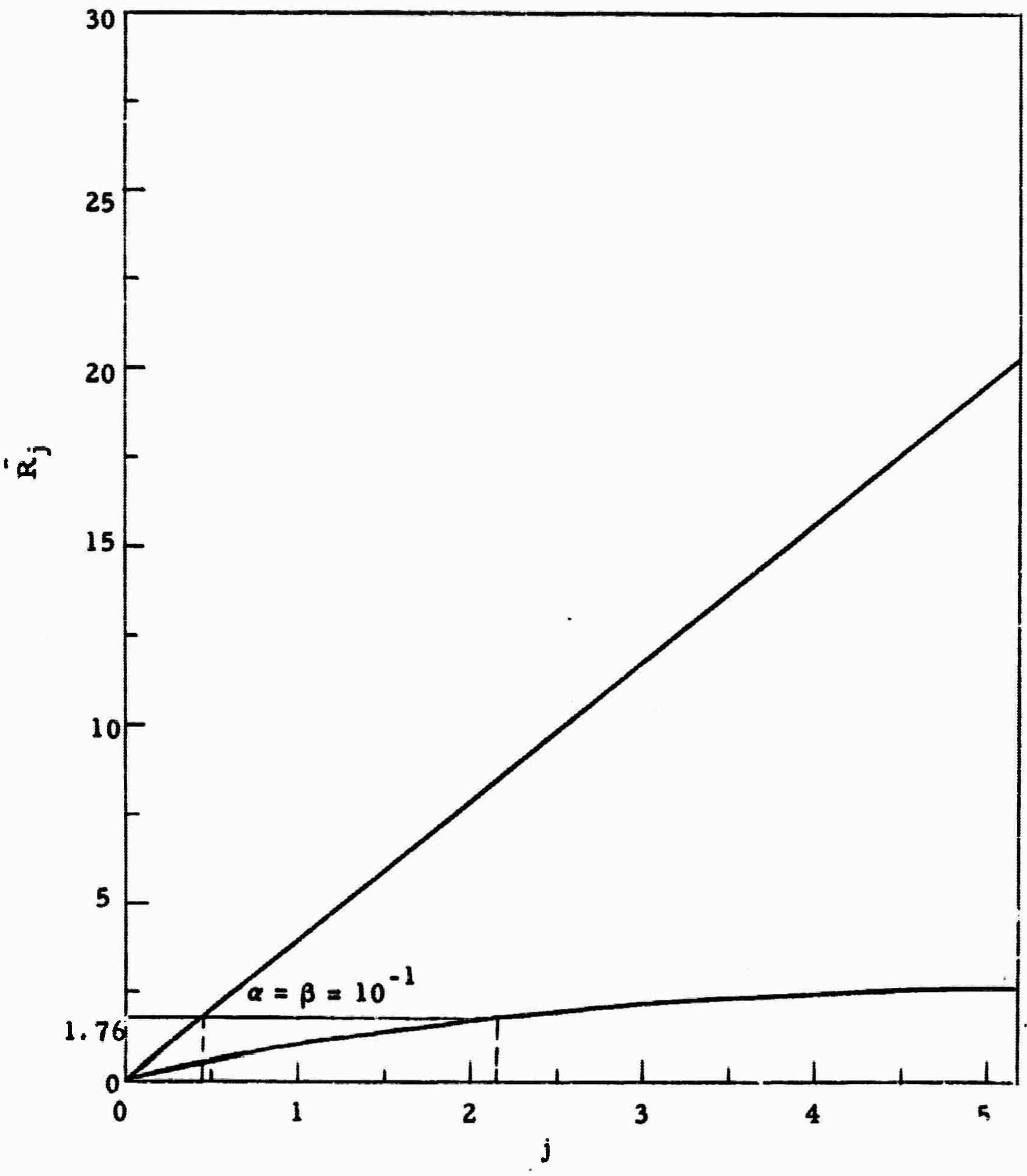

Figure 38. Calculated mean cumulative information versus sample number for sequential detectors designed for the two-known-signals case and the two-unknown-signals case for $u = 10$, $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, and $\bar{y} = 20$.

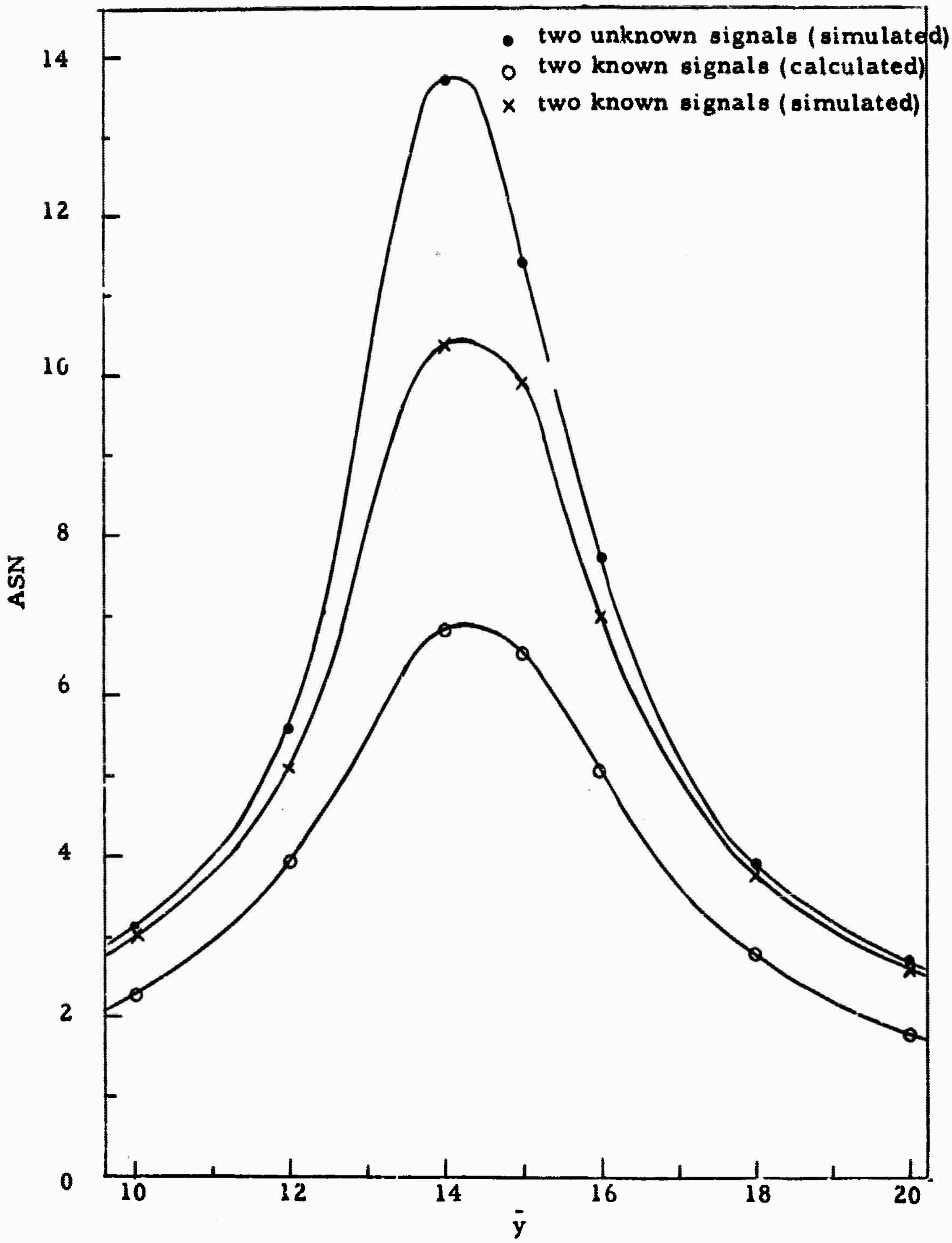

Figure 39. Computer-simulated ASN versus actual signal values $\bar{v}$ for a sequential detector designed for two unknown signals and compared with calculated and simulated ASN of the two-known-signals case for $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, $\alpha = \beta = 10^{-3}$, and u = 1000.

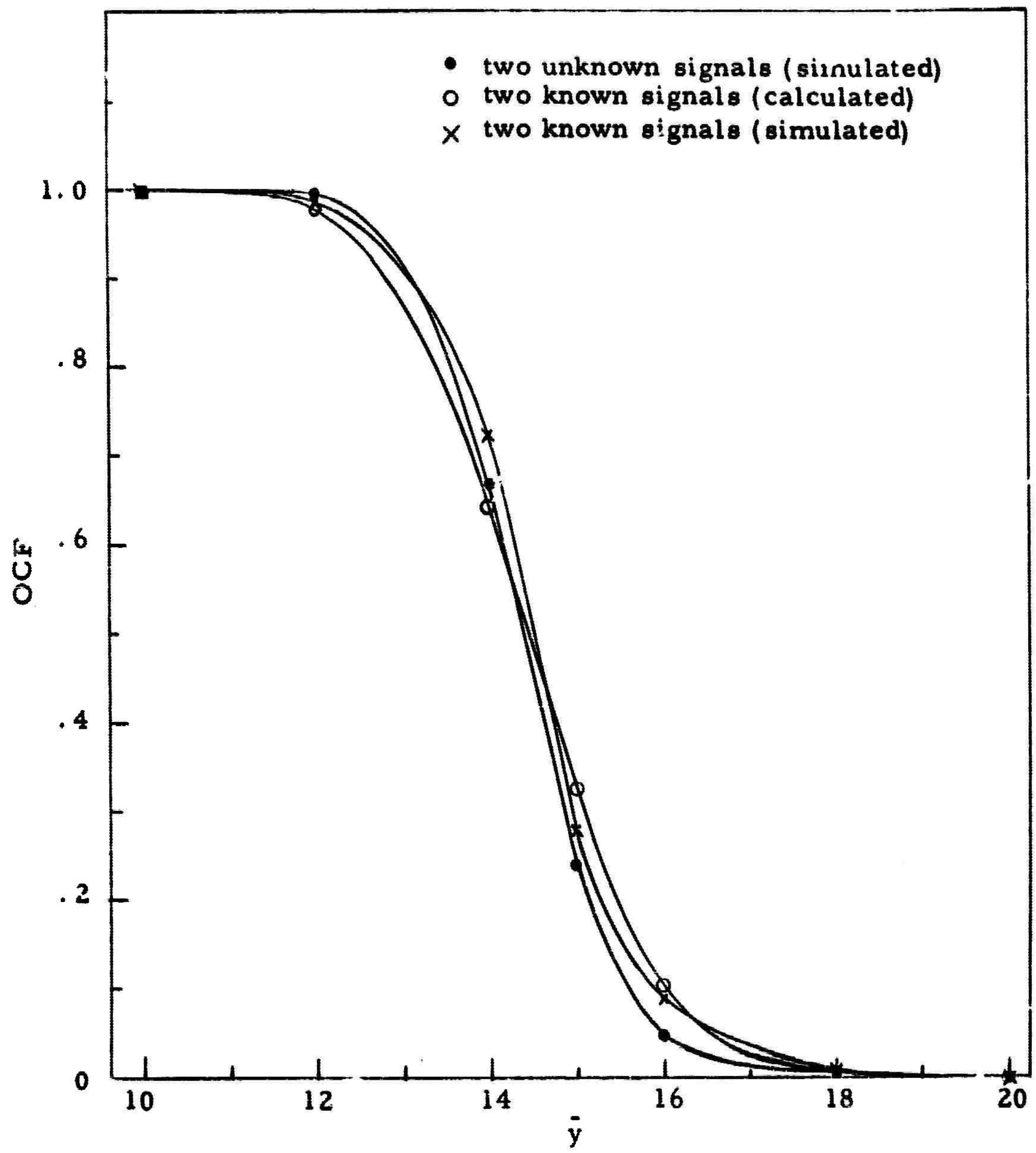

Figure 40. Computer-simulated OCF versus actual signal values $\bar{y}$ for a sequential detector designed for two unknown signals and compared with calculated and simulated OCF of the two-known-signals case for $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, $\alpha = \beta = 10^{-3}$, and $u = 1000$.

One Unknown Signal and One Known Signal

Assume that at the input of the detector we have either the known signal $\bar{y}_2$ or the unknown signal $\bar{y}_1$. We have as the two states of nature:

$$\omega_1 : \bar{y} = \bar{y}_1 \quad \text{(unknown)},$$

$$\omega_2 : \bar{y} = \bar{y}_2 \quad \text{(known)},$$

where $\bar{\bar{y}}_1 > \bar{y}_2$. We again consider this case because of its advantages in analyzing the detection error from which we can gain additional insight into the detection process. The unknown signal $\bar{y}_1$ is assumed fixed but unknown and having each of its elements taken from statistically independent gamma distributions with known parameters:

$$f(\bar{y}_1) = \prod_{\ell=1}^{m} f(\bar{y}_{1\ell}) = \prod_{\ell=1}^{m} \frac{\alpha_{1\ell}(\alpha_{1\ell}\bar{y}_{1\ell})^{u_{1\ell}-1} e^{-\alpha_{1\ell}\bar{y}_{1\ell}}}{\Gamma(u_{1\ell})}, \quad \bar{y}_{1\ell} \geq 0 \qquad (275)$$

$$= 0, \qquad \text{otherwise.}$$

The conditional probability distribution for the observed random variable z, given ω_1 and $\bar{y}_1$, is

$$P_j(z/\omega_1,\bar{y}_1) = \prod_{\ell=1}^{m} \prod_{i=1}^{j} \frac{(n\tau\bar{y}_{1\ell})^{z_\ell^i} e^{-n\tau\bar{y}_{1\ell}}}{z_\ell^i\,!}. \qquad (276)$$

Also,

$$P_j(z/\omega_2) = \prod_{\ell=1}^{m} \prod_{i=1}^{j} \frac{(n\pi\bar{y}_{2\ell})^{z_\ell^i} e^{-n\pi\bar{y}_{2\ell}}}{z_\ell^i\,!} \,. \tag{277}$$

The likelihood ratio becomes

$$\ell(j) = \frac{P_j(z/\omega_1)}{P_j(z/\omega_2)} = \int_0^\infty \frac{P_j(z/\omega_1,\bar{y}_1)\; f\;(\bar{y}_1)\; d\bar{y}_1}{P_j(z/\omega_2)}$$

$$= \prod_{i=1}^{m} \left(\frac{\alpha_{1\ell}}{\alpha_{1\ell}+j\pi\pi}\right)^{u_{1\ell}} \frac{\Gamma\left(\sum_{i=1}^{j} z_\ell^i+u_{1\ell}\right) e^{j\pi\pi\bar{y}_{2\ell}}}{[\bar{y}_{2\ell}(\alpha_{1\ell}+j\pi\pi)]^{\sum_{i=1}^{j} z_\ell^i}\;\Gamma(u_{1\ell})} \,. \tag{278}$$

Taking the logarithm of this equation yields

$$R_j = \sum_{\ell=1}^{m} \Bigg[\ell n\Gamma\left(\sum_{i=1}^{j} z_\ell^i+u_{1\ell}\right) + u_{1\ell}\ell n\left(\frac{\alpha_{1\ell}}{\alpha_{1\ell}+j\pi\pi}\right) + j\pi\pi\bar{y}_{2\ell}$$

$$-\ell n\Gamma(u_{1\ell}) - \left(\sum_{i=1}^{j} z_\ell^i\right)\ell n\;[\bar{y}_{2\ell}(\alpha_{1\ell}+j\pi\pi)]\Bigg] . \tag{279}$$

We can now use this expression in the SPRT procedure to discriminate between the states of nature ω_1 and ω_2.

Now consider the single cell case where $\bar{y}_2$, $\bar{y}_1$, $\bar{y}$, z, and k are scalars. For the single cell case we have

$$R_j = \ln \Gamma \left(\sum_{i=1}^{j} z^i + u_1\right) + j\pi \bar{y}_2 + u_1 \ln\left(\frac{\alpha_1}{\alpha_1 + j\pi}\right) - \ln \Gamma(u_1) - \left(\sum_{i=1}^{j} z^i\right) \ln [\bar{y}_2(\alpha_1 + j\pi)]. \quad (280)$$

Using the approximation

$$\ln\Gamma\ (x+1) \approx \tfrac{1}{2}\ln 2\pi - x + (x+\tfrac{1}{2})\ \ln x \quad (281)$$

we can write

$$R_j \approx \left(\sum_{i=1}^{j} z^i + u_1 - \tfrac{1}{2}\right) \ln \left(\sum_{i=1}^{j} z^i + u_1 - 1\right) - (u_1 - \tfrac{1}{2}) \ln (u_1 - 1) + u_1 \ln\left(\frac{\alpha_1}{\alpha_1 + j\pi}\right) + j\pi\bar{y}_2 - \left(\sum_{i=1}^{j} z^i\right)\left[1 + \ln[\bar{y}_2(\alpha_1 + j\pi)]\right] \quad (282)$$

and

$$I_j = R_j - R_{j-1} = \left(\sum_{i=1}^{j-1} z^i + u_1 - \tfrac{1}{2}\right) \ln\left(\frac{\sum_{i=1}^{j} z^i + u_1 - 1}{\sum_{i=1}^{j-1} z^i + u_1 - 1}\right) + z^j \ln\left(\sum_{i=1}^{j} z^i + u_1 - 1\right)$$

$$-z^j - \left(\sum_{i=1}^{j-1} z^i\right) \ln\left(\frac{\alpha_1 + jn\tau}{\alpha_1 + jn\tau - n\tau}\right) - z^j \ln\left[\bar{y}_2(\alpha_1 + jn\tau)\right] \tag{283}$$

$$-u_1 \ln\left(\frac{\alpha_1 + jn\tau}{\alpha_1 + jn\tau - n\tau}\right) + n\tau\bar{y}_2 .$$

To solve for ASN and OCF analytically appears extremely difficult in this case. We will consider only computer-simulated solutions. Figures 41 and 42 show respectively the simulated ASN and OCF versus actual signal values $\bar{y}$ for a detector designed for one signal with known mean rate and one signal with unknown mean rate. Figure 43 shows, for comparison, the ASN for the two-known-signals case, the two-unknown-signals case, and the case just considered. It can be seen from this figure how the ASN increases as less is known about the signals. Two hundred trials were made for each of the simulated results of Figures 41 and 42.

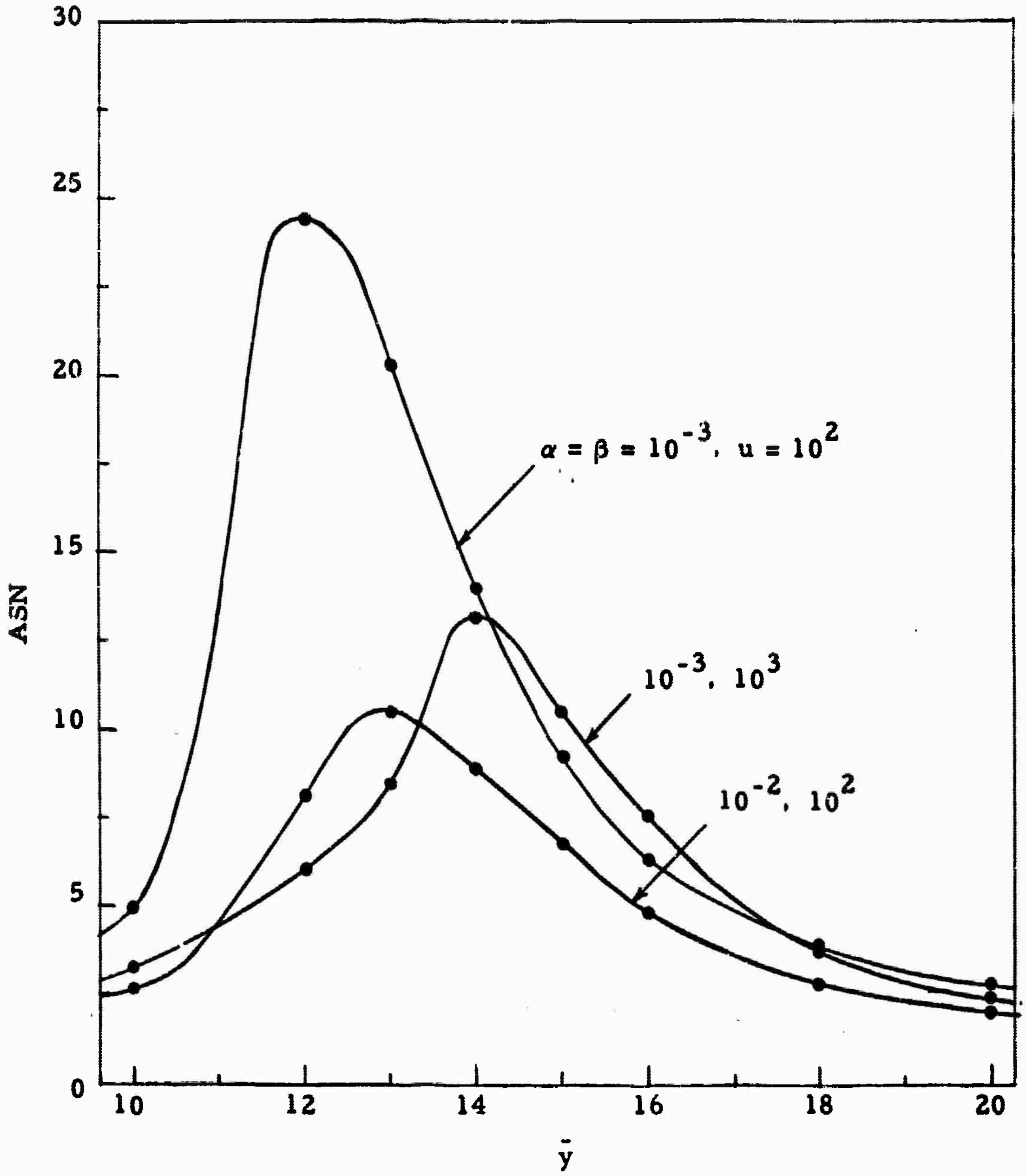

Figure 41. Computer-simulated ASN versus actual signal values $\bar{y}$ for a sequential detector designed for one known signal $\bar{y}_2$ and one unknown signal $\bar{y}_1$ where $\bar{\bar{y}}_1 = 20$ and $\bar{y}_2 = 10$.

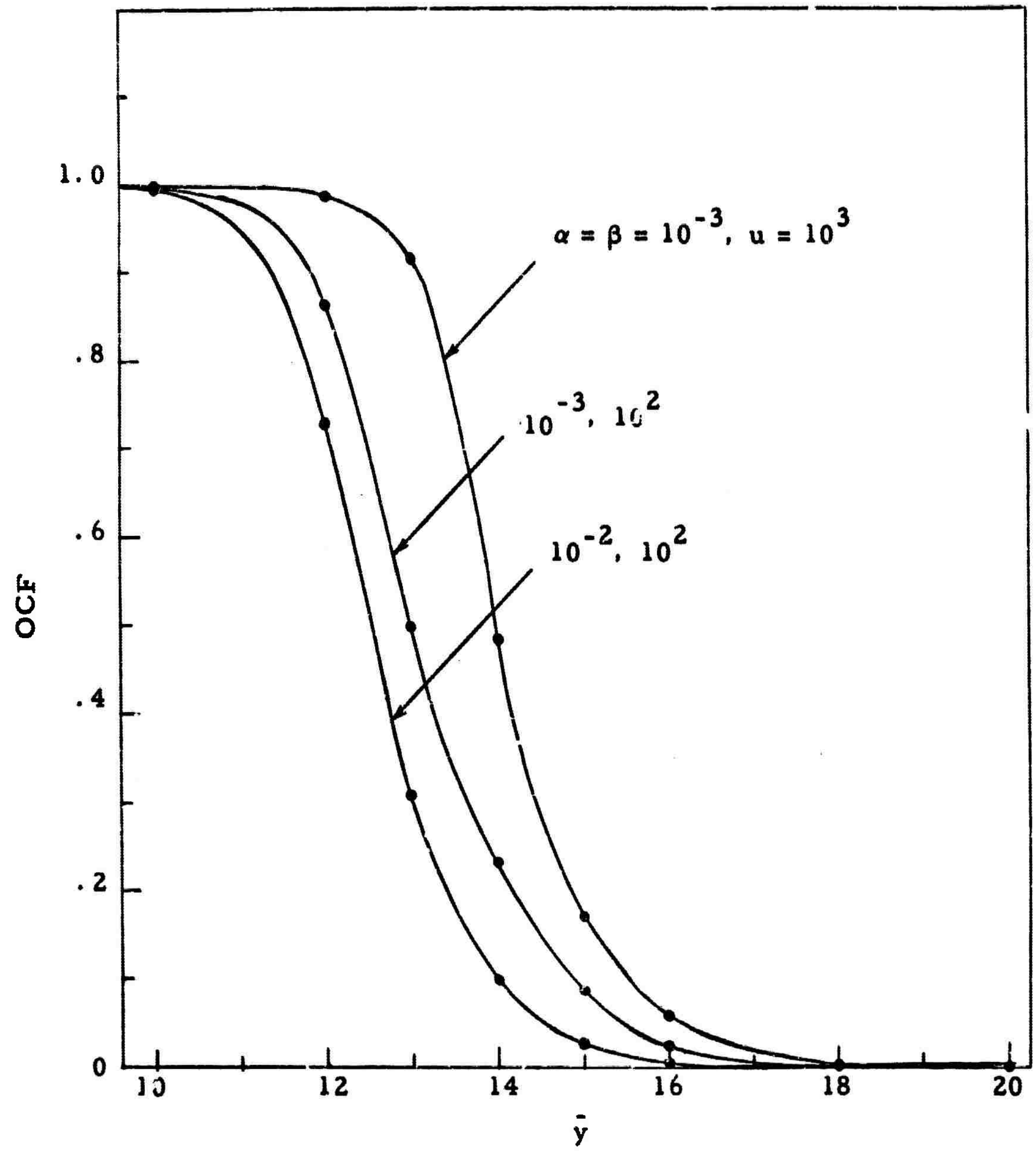

Figure 42. Computer-simulated OCF versus actual signal values $\bar{y}$ for a sequentail detector designed for one known signal $\bar{y}_2$ and one unknown signal $\bar{y}_1$ where $\bar{\bar{y}}_1 = 20$ and $\bar{y}_2 = 10$.

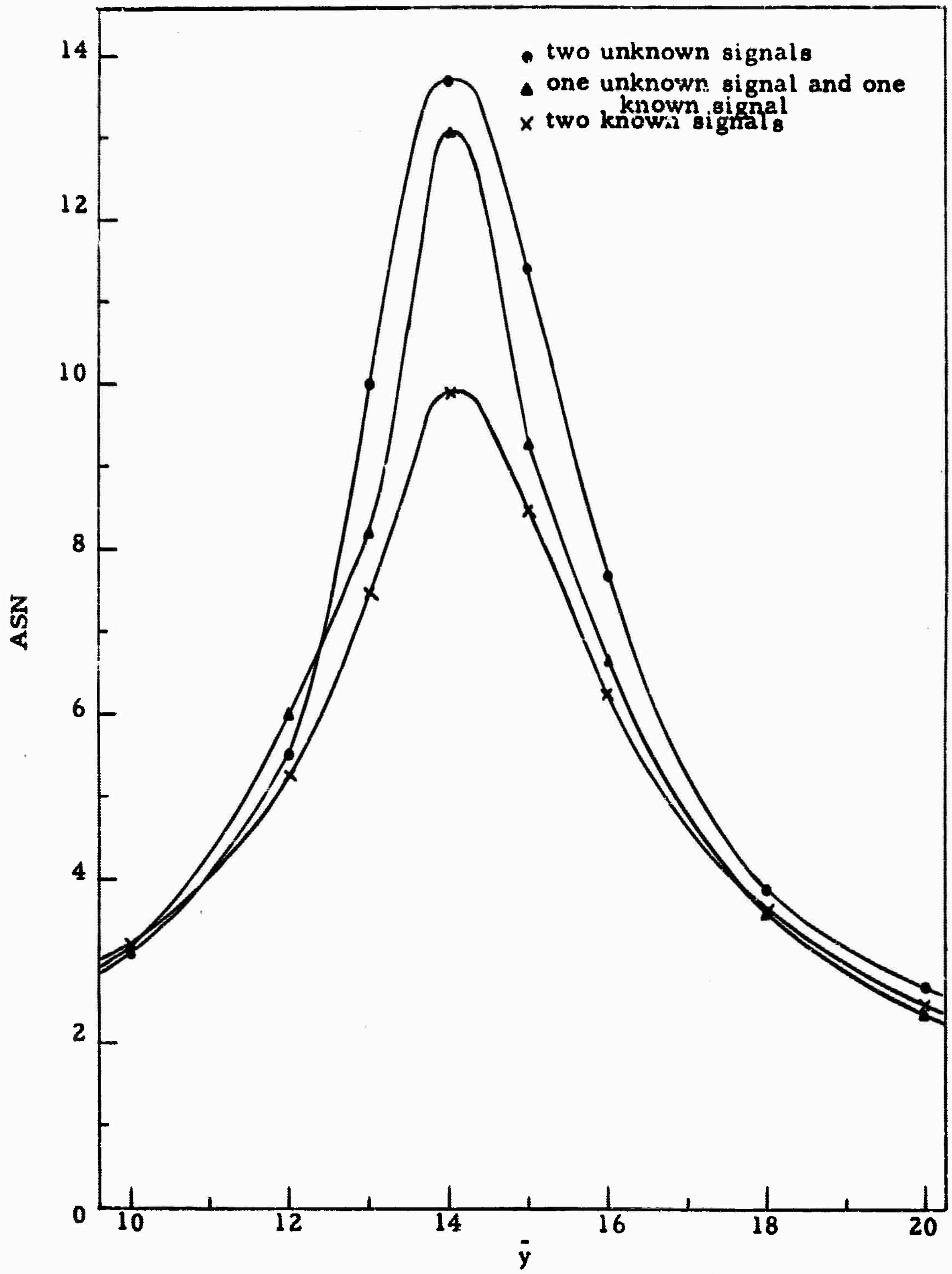

Figure 43. Comparison of simulated ASN versus actual signal $\bar{y}$ for the sequential detectors designed for the two-known-signals case, the two-unknown-signals case and the case of one known signal and one unknown signal where $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, u = 1000, and $\alpha = \beta = 10^{-3}$.

Unknown Signal Imbedded in Known Noise

Consider the case where we have one unknown signal $\bar{s}$ which, if present at the input of the detector, is imbedded in known noise $\bar{n}$. The two states of nature are:

$$\omega_1: \ \bar{y} = \bar{s} + \bar{n} \ \text{(unknown signal plus noise)},$$

$$\omega_2: \ \bar{y} = \bar{n} \ \text{(noise alone)}.$$

The signal $\bar{s}$ is assumed fixed but unknown and having each of its elements taken from statistically independent gamma distributions with known parameters:

$$f(\bar{s}) = \prod_{\ell=1}^{m} f(\bar{s}_1) = \prod_{\ell=1}^{m} \frac{\beta_\ell(\beta_\ell \bar{s}_\ell)^{u_\ell - 1} e^{-\beta_\ell \bar{s}_\ell}}{\Gamma(u_\ell)}, \quad \bar{s}_\ell \geq 0 \tag{284}$$

$$= 0, \qquad \text{otherwise.}$$

The conditional likelihood ratio is

$$\ell\,(j/\bar{s}) = \frac{P_j(z/\omega_1,\ \bar{s})}{P_j(z/\omega_2)} = \prod_{\ell=1}^{m} \left(\frac{\bar{n}_\ell + \bar{s}_\ell}{\bar{n}_\ell}\right)^{\sum_{i=1}^{j} z_\ell^i} e^{-j\eta\tau\bar{s}_\ell} \ . \tag{285}$$

The likelihood ratio becomes

$$\ell(j) = \int_0^\infty \ell(j/\bar{s}) f(\bar{s}) d\bar{s}$$

$$= \prod_{\ell=1}^{m} \left(\frac{\beta_\ell}{\beta_\ell + j\eta\tau}\right)^{u_\ell} \frac{\sum_{k=0}^{\sum_{i=1}^{j} z_\ell^i} \binom{\sum_{i=1}^{j} z_\ell^i}{k} [\bar{n}_\ell(\beta_\ell + j\eta\tau)]^k \Gamma\left(\sum_{i=1}^{j} z_\ell^i + u_\ell - k\right)}{\Gamma(u_\ell)[\bar{n}_\ell(\beta_\ell + j\eta\tau)]^{\sum_{i=1}^{j} z_\ell^i}} . \quad (286)$$

Taking the logarithm of this equation and assuming the single cell case we have

$$R_j = u \ln\left(\frac{\beta}{\beta + j\eta\tau}\right) - \ln \Gamma(u) + \ln\left[\left(\sum_{i=1}^{j} z^i\right)!\right] - \left(\sum_{i=1}^{j} z^i\right) \ln[\bar{n}(\beta + j\eta\tau)]$$

$$+ \ln\left[\sum_{k=0}^{\sum_{i=1}^{j} z^i} \frac{[\bar{n}(\beta + j\eta\tau)]^k \Gamma\left(\sum_{i=1}^{j} z^i + u - k\right)}{\left(\sum_{i=1}^{j} z^i - k\right)! \, k!}\right] . \quad (287)$$

This expression can then be used in the SPR procedure to discriminate between the states of nature, ω_1 and ω_2.

An exact analytical solution of ASN and OCF for this case appears extremely difficult, if not impossible. It appears that the only reasonable method of analysis of this case is by computer simulation.

Figure 44 shows the simulated ASN versus actual signal values $\bar{y}$ for a detector designed for a signal with unknown mean rate imbedded in Poisson noise with known mean rate. Figure 45 shows the corresponding OCF. Two hundred trials were made for each of the simulated results of Figures 44 and 45.

Information Content of Samples

An important property of the sequential detector is the information content of the samples. We will consider the mean information of a sample to be

$$\bar{I}_j \triangleq E(I_j/\bar{y}, j) = E(R_j/\bar{y}, j) - E(R_{j-1}/\bar{y}, j). \qquad (288)$$

The information is favorable to the hypothesis that ω_1 is the true state of nature when the above expression is positive. When the information is negative, the sequential detector tends to choose ω_2 as the state of nature. We will determine the amount of information provided by the samples of the various sequential detectors that we have previously considered.

For the detector designed for two known signals, the mean information per sample is constant regardless of the actual signal $\bar{y}$ which may be present and is given by

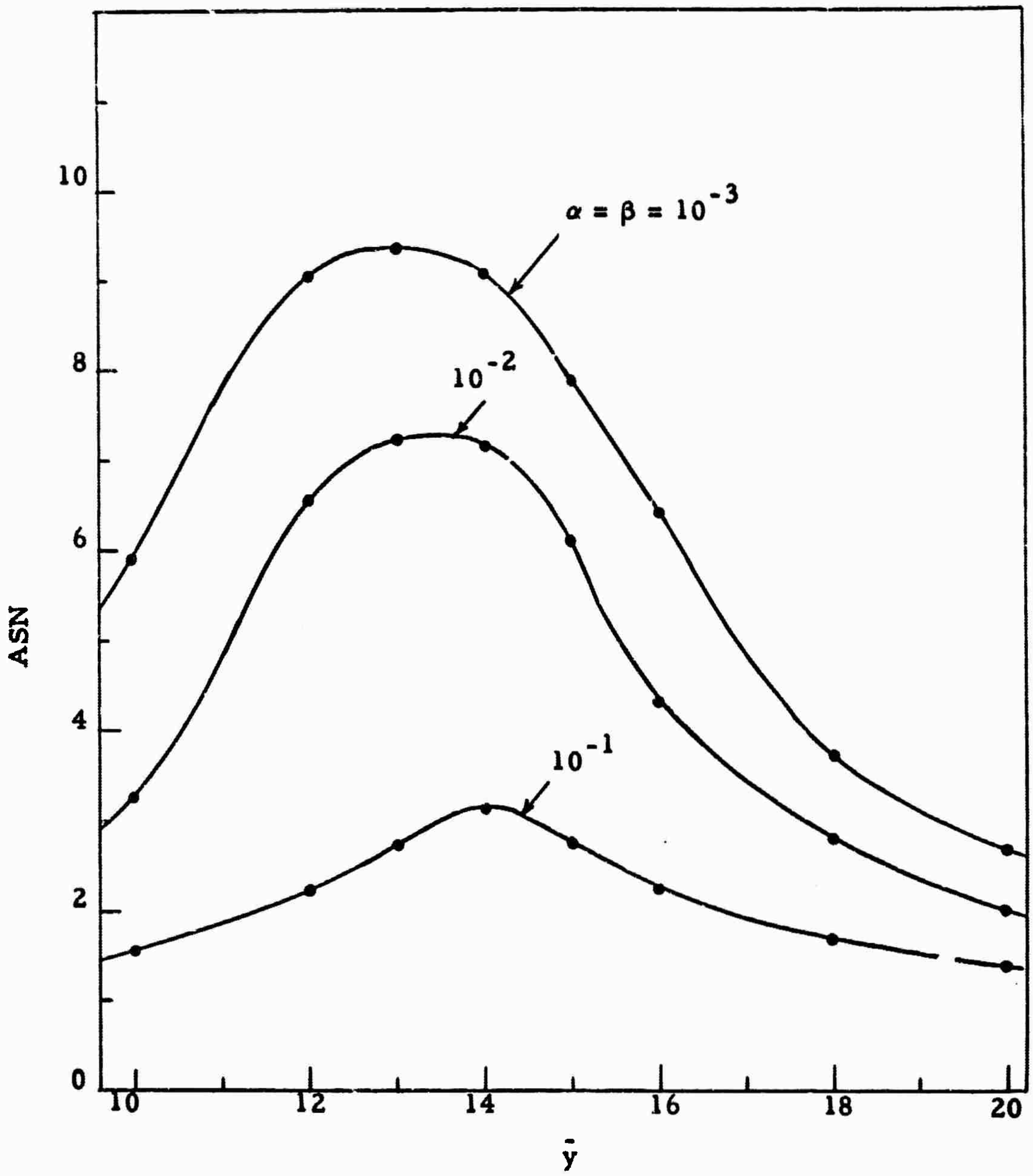

Figure 44. Computer-simulated ASN versus actual signal values $\bar{y}$ for a sequential detector designed for an unknown signal $\bar{s}$ which if present is imbedded in known noise $\bar{n}$ where $\bar{\bar{s}} = 10$, $\bar{n} = \bar{y}_2 = 10$, $\bar{\bar{y}}_1 = 20$, and $u = 10$.

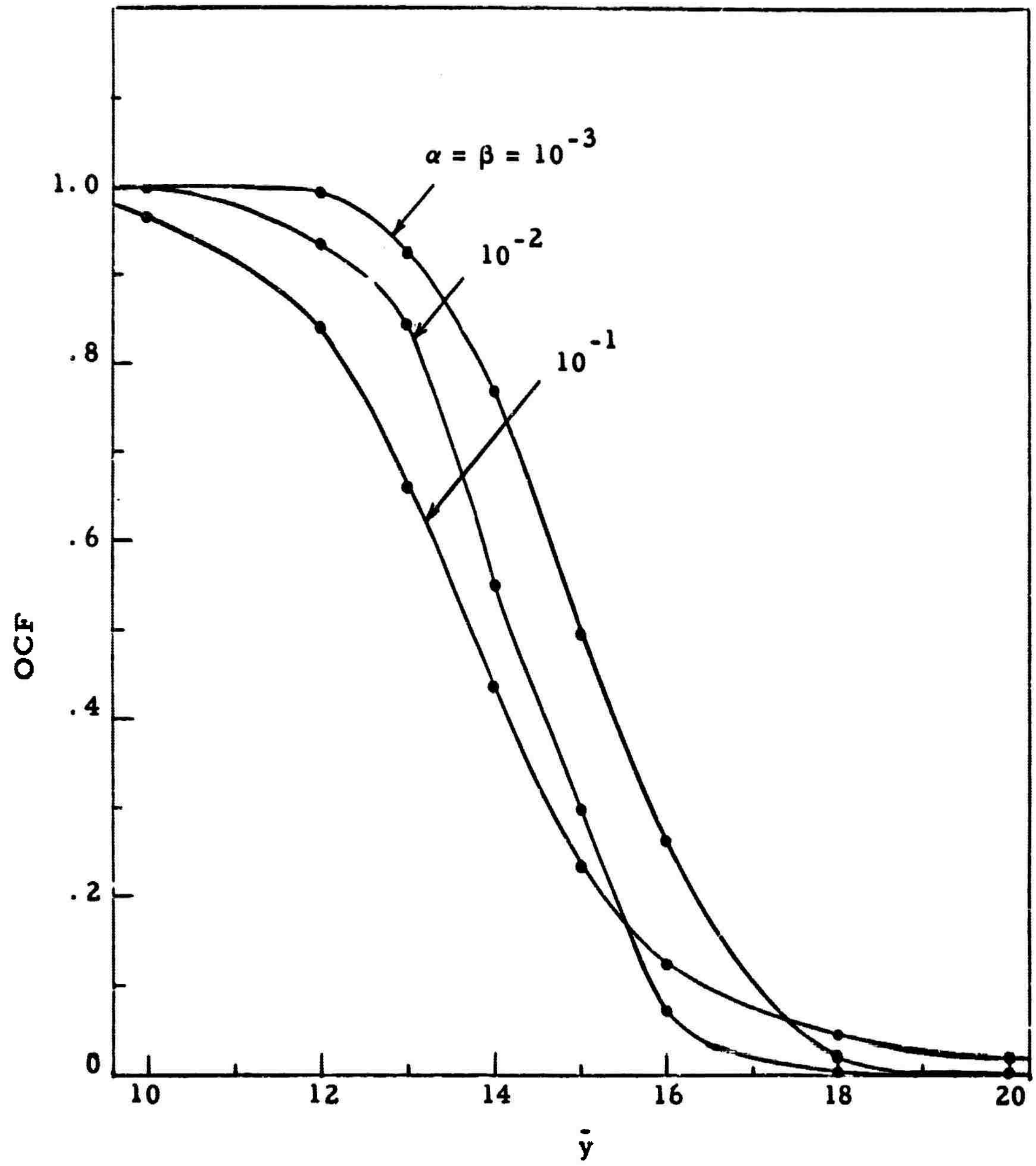

Figure 45. Computer-simulated OCF versus actual signal values $\bar{y}$ for a sequential detector designed for an unknown signal $\bar{s}$ which if present is imbedded in known noise $\bar{n}$ where $\bar{\bar{s}} = 10$, $\bar{n} = \bar{y}_2 = 10$, $\bar{\bar{y}}_1 = 20$, and $u = 10$.

$$\bar{I}_j = \eta\tau \, [\bar{y} \, \ln (\bar{y}_1/\bar{y}_2) - \bar{y}_1 + \bar{y}_2]. \tag{289}$$

The expected cumulative information for the j^{th} sample is

$$\bar{R}_j \triangleq E(R_j/\bar{y}, j) = \eta\tau j [\bar{y} \, \ln (\bar{y}_1/\bar{y}_2) - \bar{y}_1 + \bar{y}_2] \tag{290}$$

where the true mean rate is $\bar{y}$. Figure 46 shows the mean cumulative information versus sample number for this case where the detector is designed for $\bar{y}_2 = 10$ and for various values of $\bar{y}_1$. The positive values result when ω_1 is the state of nature ($\bar{y} = \bar{y}_1$), and the negative values result when ω_2 is the state of nature ($\bar{y} = \bar{y}_2$).

Whenever either or both of the two signals have unknown mean rates, the information per sample is no longer constant. For the detector designed for one unknown signal $\bar{y}_1$ and one known signal $\bar{y}_2$ the mean cumulative information for the j^{th} sample is

$$\bar{R}_j = E[\ln \Gamma(\sum_{i=1}^{j} z^i + u_1)] - j\eta\tau\bar{y} \ln[\bar{y}_2(\alpha_1 + j\eta\tau)] + j\eta\tau\bar{y}_2 \tag{291}$$

$$+ u_1 \ln \left(\frac{\alpha_1}{\alpha_1 + j\eta\tau}\right) - \ln \Gamma(u_1).$$

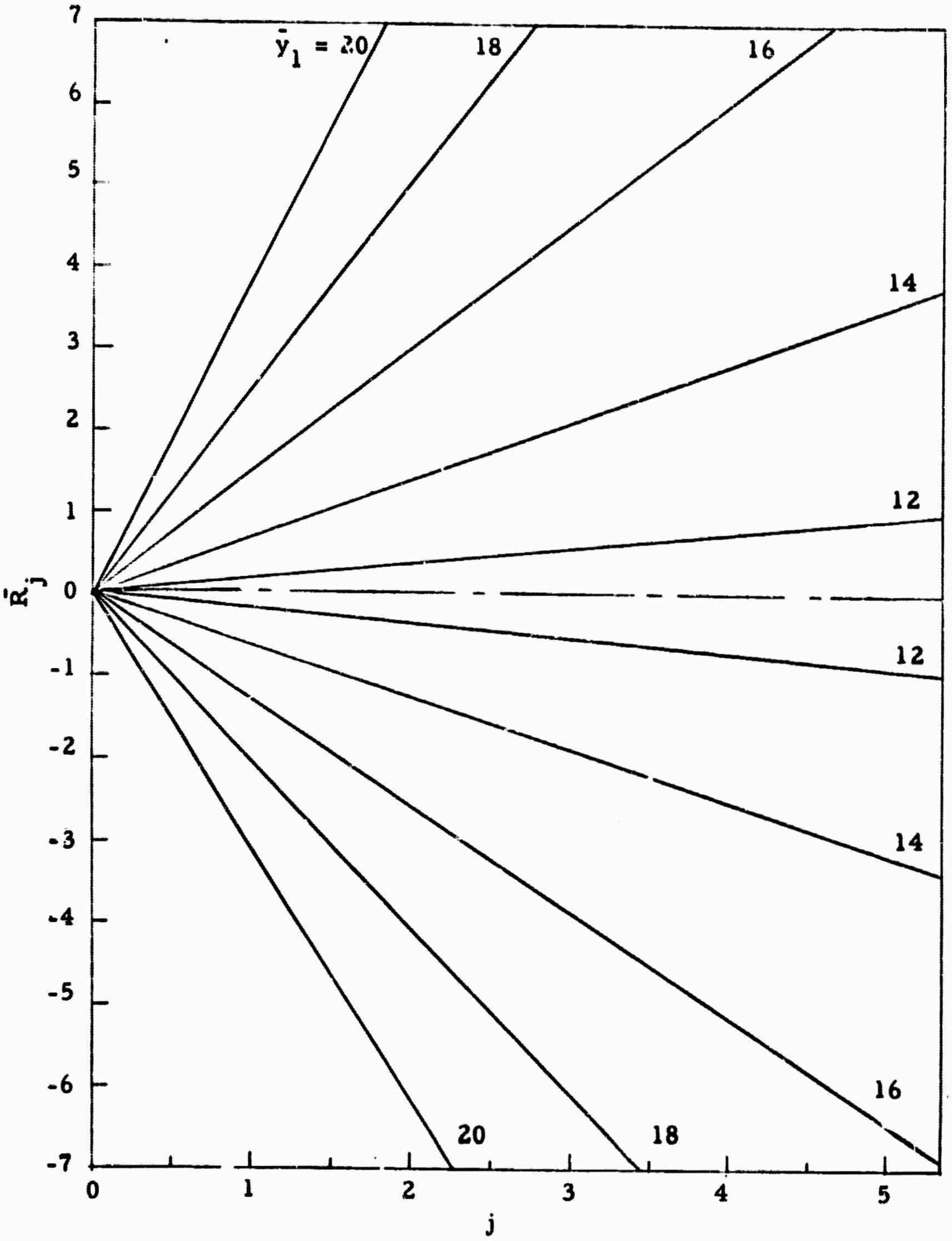

Figure 46. Calculated mean cumulative information versus sample number for a sequential detector designed for two known signals where $\bar{y}_2 = 10$ and $\bar{y} = \bar{y}_1$.

Using the approximations that

$$E(R_j(z)/\bar{y}, j) \approx R_j(\bar{z}) \tag{292}$$

and

$$\ln\Gamma(x+1) \approx \tfrac{1}{2}\ln 2\pi - x + (x+\tfrac{1}{2})\ln x \tag{293}$$

we can write

$$\bar{R}_j \approx (j\eta\tau\bar{y} + u_1 - \tfrac{1}{2})\ln(j\eta\tau\bar{y} + u_1 - 1) - (u_1 - \tfrac{1}{2})\ln(u_1 - 1) + u_1 \ln\left(\frac{\alpha_1}{\alpha_1 + j\eta\tau}\right) + j\eta\tau\bar{y}_2 - j\eta\tau\bar{y} - j\eta\tau\bar{y}\ln[\bar{y}_2(\alpha_1 + j\eta\tau)] \tag{294}$$

and

$$\bar{I}_j \approx (j\eta\tau\bar{y} - \eta\tau\bar{y} + u_1 - \tfrac{1}{2})\ln\left(\frac{j\eta\tau\bar{y} + u_1 - 1}{j\eta\tau\bar{y} - \eta\tau\bar{y} + u_1 - 1}\right) + j\eta\tau\bar{y}\ln(j\eta\tau + u_1 - 1) - j\eta\tau\bar{y} - (j\eta\tau\bar{y} - \eta\tau\bar{y})\ln\left(\frac{\alpha_1 + j\eta\tau}{\alpha_1 + j\eta\tau - \eta\tau}\right) - \eta\tau\bar{y}\ln[\bar{y}_2(\alpha_1 + j\eta\tau)] - u_1 \ln\left(\frac{\alpha_1 + j\eta\tau}{\alpha_1 + j\eta\tau - \eta\tau}\right) + \eta\tau\bar{y}_2. \tag{295}$$

The limit of $\bar{I}_j$ as j approaches infinity gives us an indication of what the detector is doing as the sample number becomes large. The mean information per sample provides us with a measure of the adaptive capability and the performance of the detector. It provides some insight into the speed by which the detector will make a decision. The limit of $\bar{R}_j$ as j approaches infinity gives us some upper bound to the information that we can attain and may point out some limitations of the detector. For the case of the detector designed for one unknown signal $\bar{y}_1$ and one known signal the limit of the mean cumulative information as j approaches infinity is

$$\lim_{j\to\infty} \bar{R}_j \approx u_1 \ln \alpha_1 \bar{y} - \alpha_1 \bar{y} - (u_1 - \tfrac{1}{2}) \ln(u_1 - 1) + (u_1 - 1) + \lim_{j\to\infty} \Big[j\eta\tau[\bar{y} \ln(\bar{y}/\bar{y}_2) - \bar{y} + \bar{y}_2] - \tfrac{1}{2} \ln(j\eta\tau\bar{y} + u_1 - 1) \Big] . \tag{296}$$

The limit of the mean information per sample as j approaches infinity is

$$\lim_{j\to\infty} \bar{I}_j = \eta\tau[\bar{y} \ln(\bar{y}/\bar{y}_2) - \bar{y} + \bar{y}_2]. \tag{297}$$

For ω_1 as the state of nature the detector initially does not know the true mean rate of the signal and the mean information per sample provided about the state ω_1 is small. As more samples are taken, the detector adapts to $\bar{y}_1$ and thus becomes more effective. As the sample number becomes large, the detector learns $\bar{y}_1$, and then the mean information per sample becomes the same as for the case when the signal was known beforehand. If ω_2 is the state of nature the detector is most effective to begin with, and as more samples are taken the detector performance begins to deteriorate since the mean information per sample decreases and in the limit the mean information per sample approaches zero. The mean cumulative information approaches infinity as j approaches infinity when either ω_1 or ω_2 is the state of nature but the mean cumulative information for state ω_1 approaches infinity faster than for state ω_2. The mean cumulative information versus sample number for this case is shown in Figure 47. The detector for this case is designed for a known signal with a mean rate of $\bar{y}_2 = 10$ and for an unknown signal with various expected mean rates $\bar{\bar{y}}_1$. The positive values result when ω_1 is the state of nature ($\bar{y} = \bar{\bar{y}}_1$) and negative values result when ω_2 is the state of nature ($\bar{y} = \bar{y}_2$).

For the case of one unknown signal $\bar{s}$ imbedded in known Poisson noise $\bar{n}$, a similar result holds. For this case when ω_1 is the true state of nature the detector approaches the two-known-signals case faster than does the detector of the case just discussed where one signal is known and one signal is unknown. Also, when ω_2 is the true state of nature, the detector does not deteriorate as rapidly. The mean cumulative information versus sample number for this case is shown in Figure 48. The detector is designed for known noise $\bar{n} = \bar{y}_2 = 10$, and several values

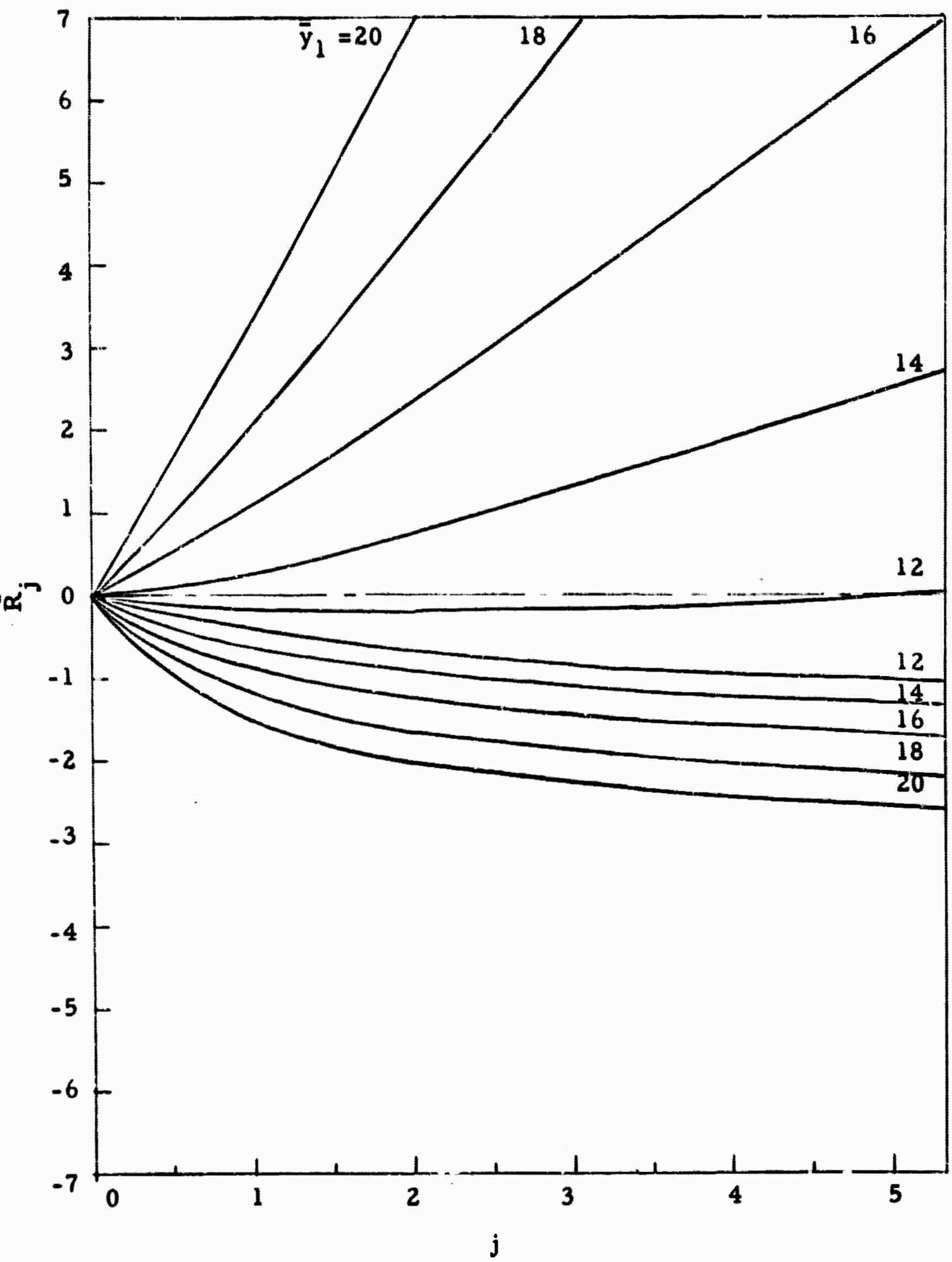

Figure 47. Calculated mean cumulative information versus sample number for a sequential detector designed for one known signal $\bar{y}_2$ and several mean values of the unknown signal $\bar{y}_1$ where $\bar{y}_2$ = 10, $\bar{y} = \bar{\bar{y}}$, and u = 10.

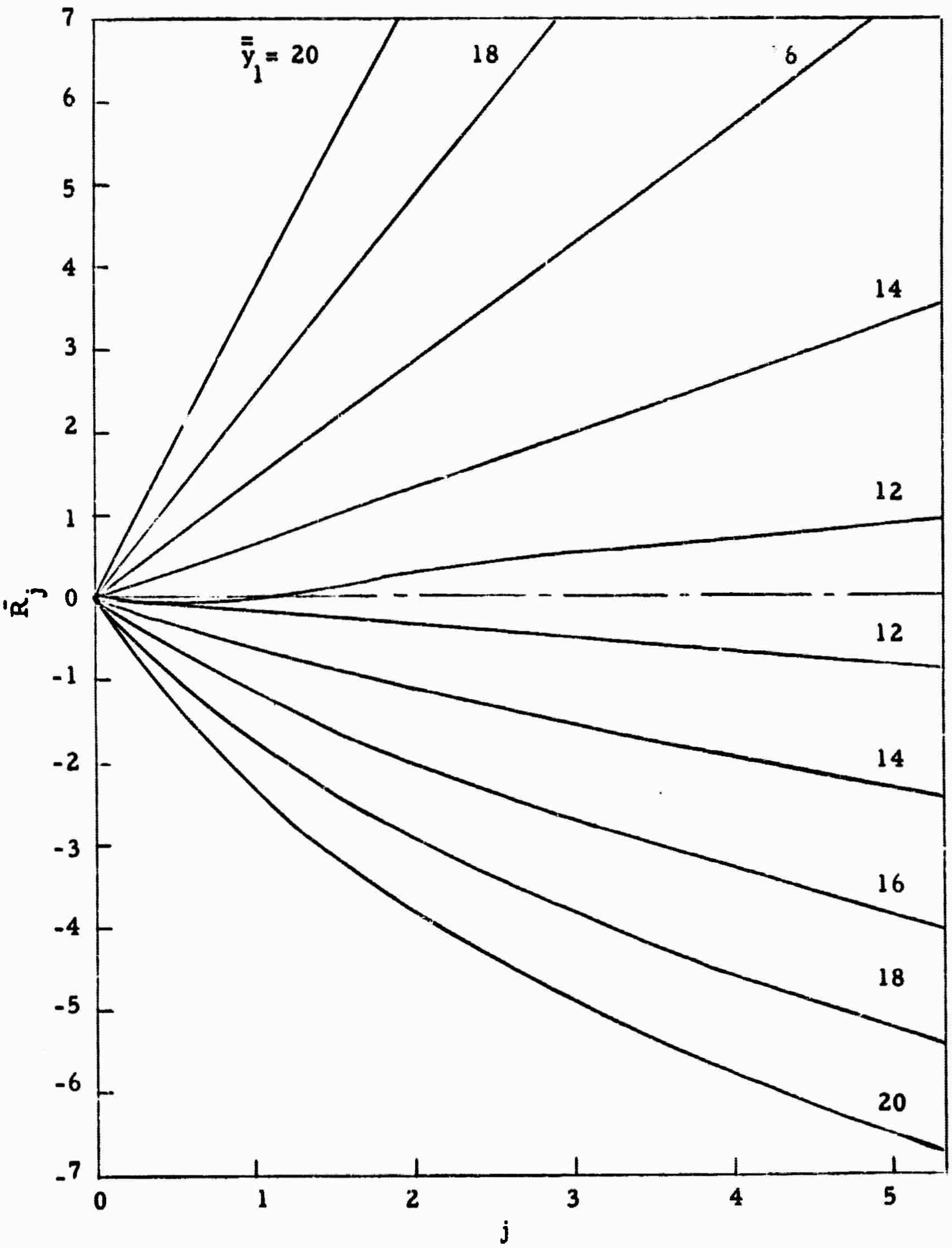

Figure 48. Calculated mean cumulative information versus sample number for a sequential detector designed for the unknown signal $\bar{s}$ which if present is imbedded in known noise $\bar{n}$ where $\bar{y}_2 = \bar{n} = 10$, $\bar{y} = \bar{\bar{y}}_1 = \bar{\bar{s}} + \bar{n}$, and u = 10.

of unknown signal $\bar{s}$ or $\bar{y}_1 = \bar{s} + \bar{n}$. Positive values result when ω_1 is the state of nature ($\bar{y} = \bar{\bar{y}}_1$), and negative values result when ω_2 is the state of nature ($\bar{y} = \bar{y}_2$).

When we have the two-unknown-signals case, the mean information per sample for the j^{th} sample is

$$\bar{I}_j = (j\eta\tau\bar{y} - \eta\tau\bar{y} + u)\left[\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) + \ln\left(\frac{\alpha_2 + j\eta\tau - \eta\tau}{\alpha_1 + j\eta\tau - \eta\tau}\right)\right] + \eta\tau\bar{y}\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) \tag{298}$$

and the mean cumulative information for the j^{th} sample is

$$\bar{R}_j = j\eta\tau\bar{y}\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) + u\left[\ln\left(\frac{\alpha_2 + j\eta\tau}{\alpha_1 + j\eta\tau}\right) - \ln(\alpha_2/\alpha_1)\right]. \tag{299}$$

For this case the detector is most effective to begin with, regardless of the true state of nature, and as more samples are taken, the detector performance begins to deteriorate because of the decrease in the mean information per sample and in the limit the mean information per sample approaches zero. The limit of the mean cumulative information approaches some constant depending on the actual signal present and is given by

$$\lim_{j\to\infty} \bar{R}_j = \bar{y}(\alpha_2 - \alpha_1) - u\ln(\alpha_2/\alpha_1). \tag{300}$$

Figure 49 shows the mean cumulative information versus sample number for this case when the detector is designed for an unknown signal with an expected mean rate of $\bar{\bar{y}}_2 = 10$ and for an unknown signal with various expected mean rates $\bar{\bar{y}}_1$. The positive values result when ω_1 is the state of nature ($\bar{y} = \bar{\bar{y}}_1$) and negative values result when ω_2 is the state of nature ($\bar{y} = \bar{\bar{y}}_2$).

Figure 50 shows the mean information per sample versus sample number for the two-known-signals case and the two-unknown-signals case. The positive information is obtained when ω_1 is the state of nature ($\bar{y} = \bar{\bar{y}}_1$), and negative information is obtained when ω_2 is the state of nature ($\bar{y} = \bar{\bar{y}}_2$). Figure 51 shows the calculated mean cumulative information versus sample number and also a computer-simulated sample of the cumulative information for a detector designed for two known signals when ω_1 is the state of nature ($\bar{y} = \bar{y}_1$). Also shown in Figure 51 is the computer-simulated mean of the cumulative information versus sample number. This simulated sample mean compares very favorably with the calculated values. Figure 52 shows the results for the same detector when ω_2 is the state of nature ($\bar{y} = \bar{y}_2$). Figures 53 and 54 show results that correspond to Figures 51 and 52 respectively, for a detector designed for two unknown signals. Figure 55 shows the computer-simulated mean cumulative information versus sample number for a detector designed for two known signals, $\bar{y}_1 = 20$ and $\bar{y}_2 = 10$, and for various values of actual signals present. Figure 56 shows the computer-simulated mean

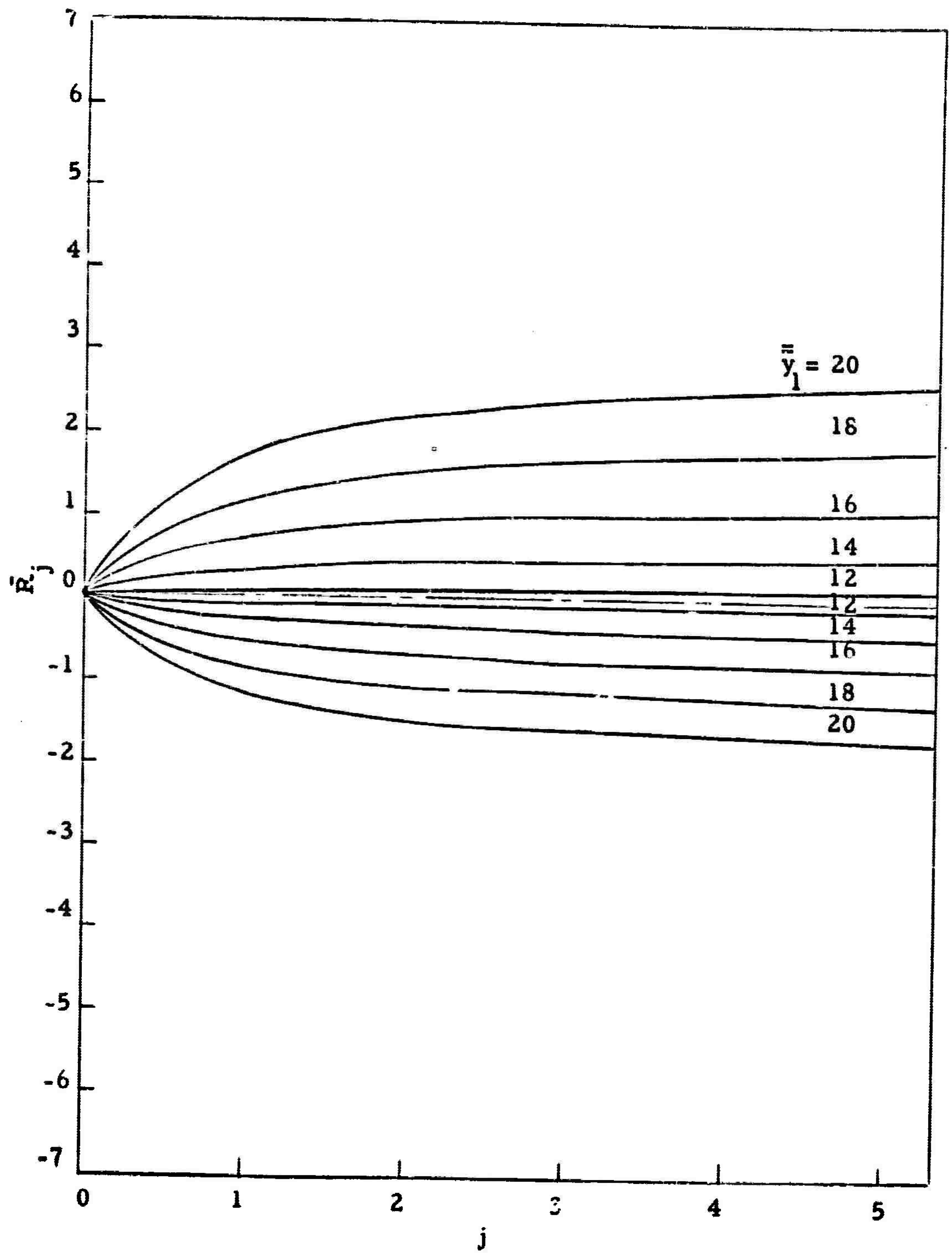

Figure 49. Calculated mean cumulative information versus sample number for a sequential detector designed for two unknown signals where $\bar{\bar{y}}_2 = 10$, $\bar{y} = \bar{\bar{y}}_1$, and $u = 10$.

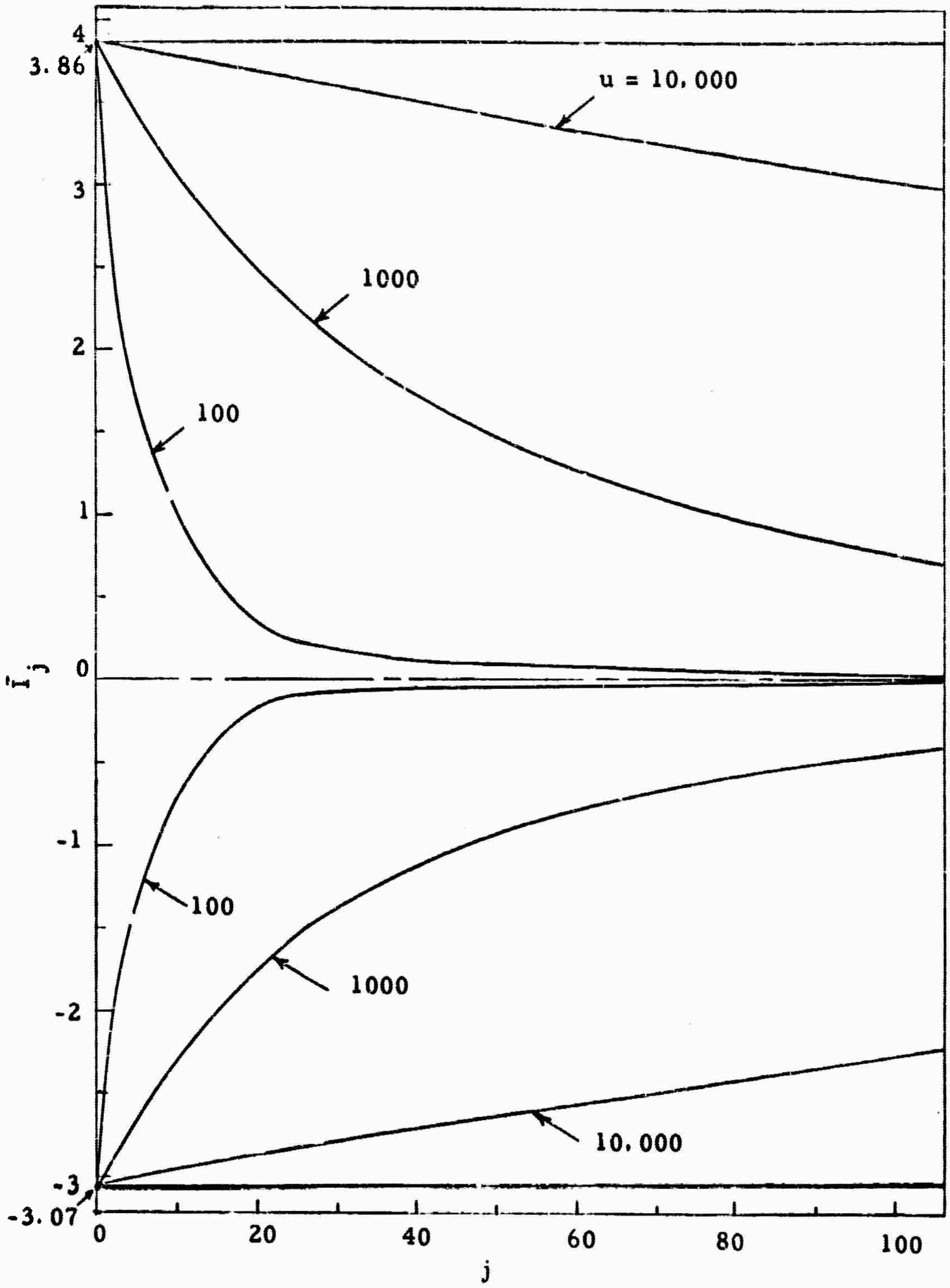

Figure 50. Calculated mean information per sample versus sample number for a sequential detector designed for the two-known-signals case and the two-unknown-signals case where $\bar{\bar{y}}_1 = 20$, and $\bar{\bar{y}}_2 = 10$.

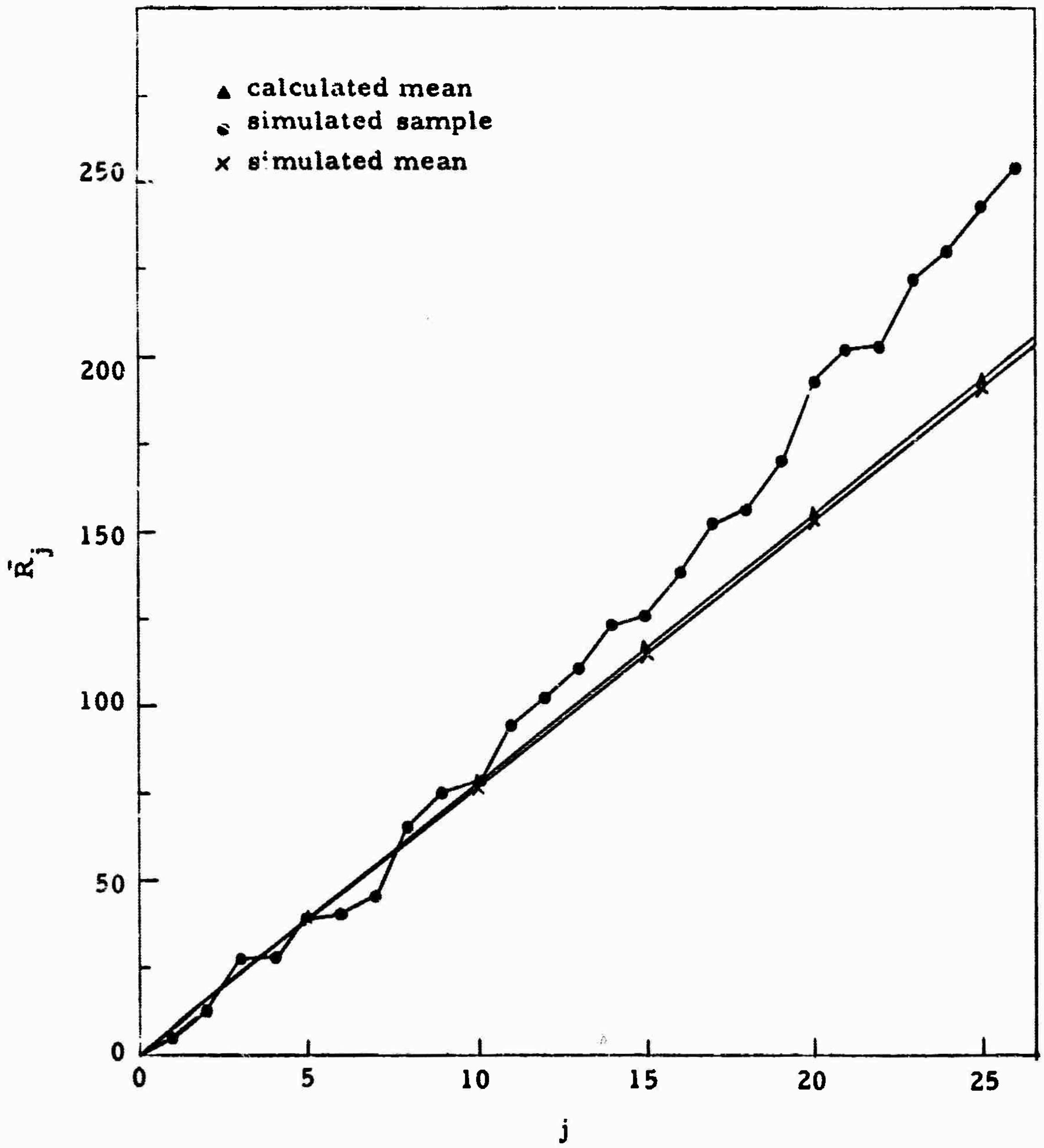

Figure 51. Calculated and simulated mean cumulative information versus sample number and a simulated sample of the cumulative information versus sample number for a sequential detector designed for two known signals when ω_1 is the state of nature and $\bar{y}_1 = 20$, $\bar{y} = 10$, and $\bar{y} = 20$.

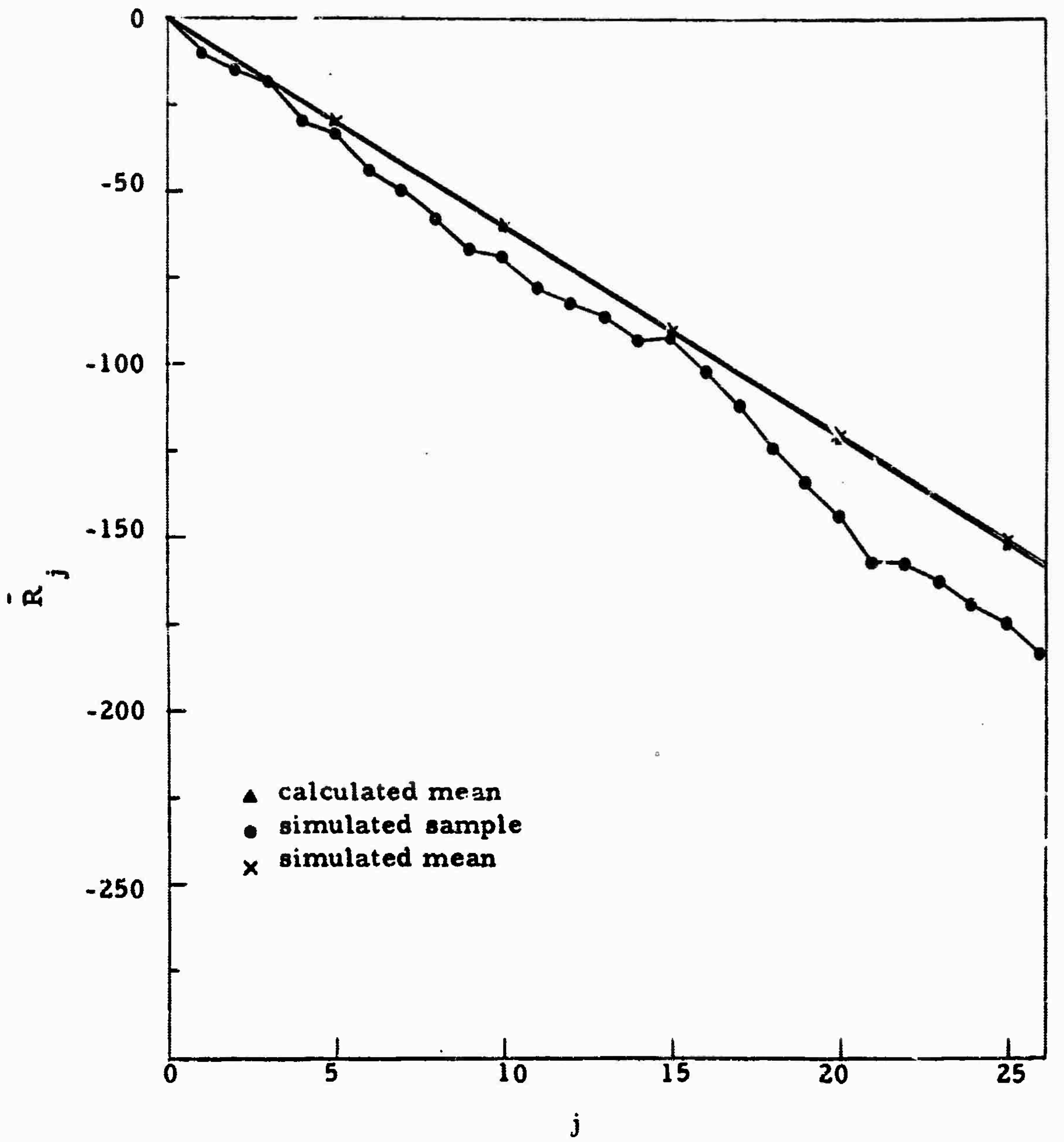

Figure 52. Calculated and simulated mean cumulative information versus sample number and a simulated sample of the cumulative information versus sample number for a sequential detector designed for two known signals when ω_2 is the state of nature and $\bar{y}_1 = 20$, $\bar{y}_2 = 10$, and $\bar{y} = 10$.

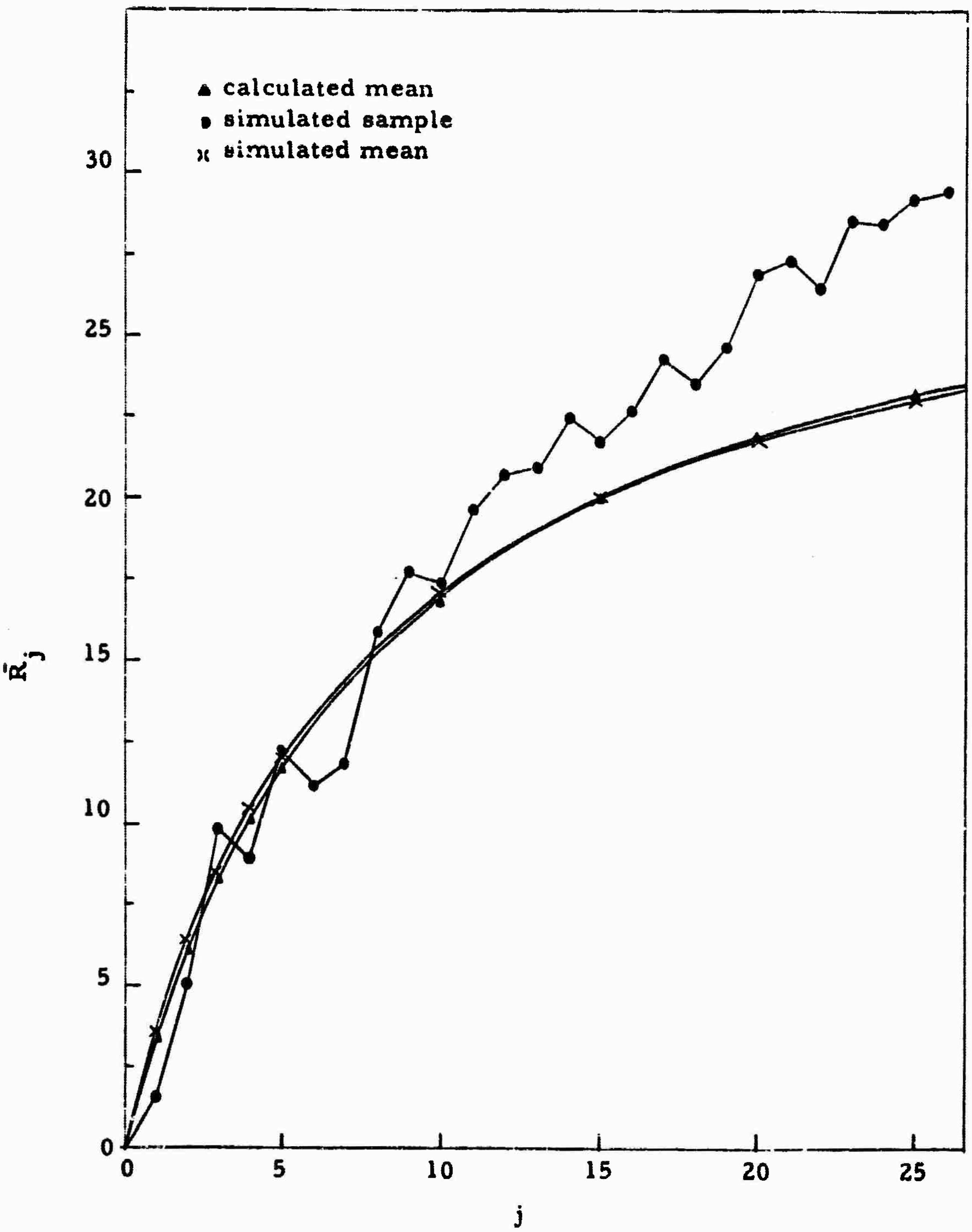

Figure 53. Calculated and simulated mean cumulative information versus sample number and a simulated sample of the cumulative information versus sample number for a sequential detector designed for two unknown signals when ω_1 is the state of nature and $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, $\bar{y} = 20$, and $u = 100$.

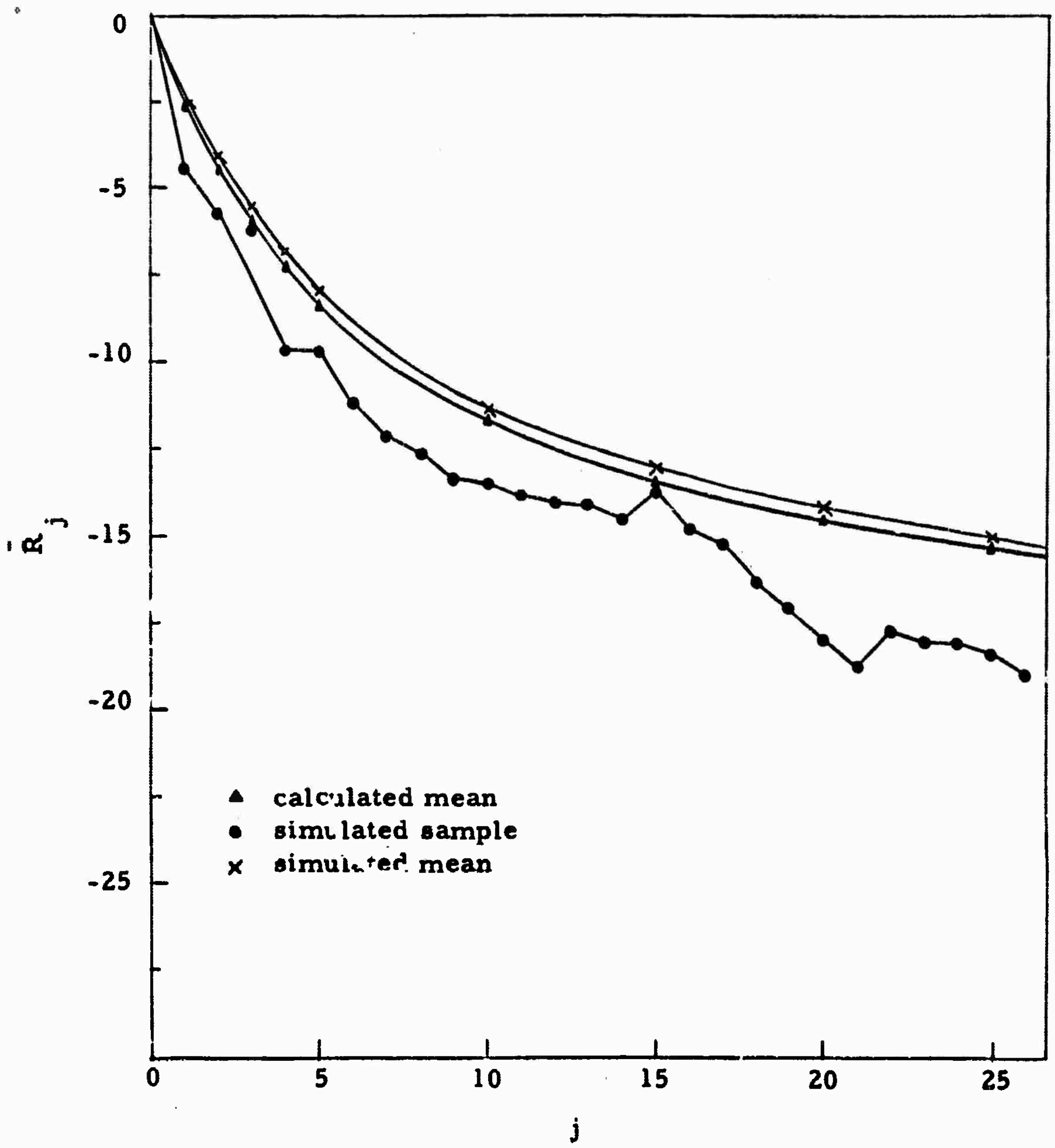

Figure 54. Calculated and simulated mean cumulative information versus sample number and a simulated sample of the cumulative information versus sample number for a sequential detector designed for two unknown signals when ω_2 is the state of nature and $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, $\tilde{y} = 10$, and $u = 100$.

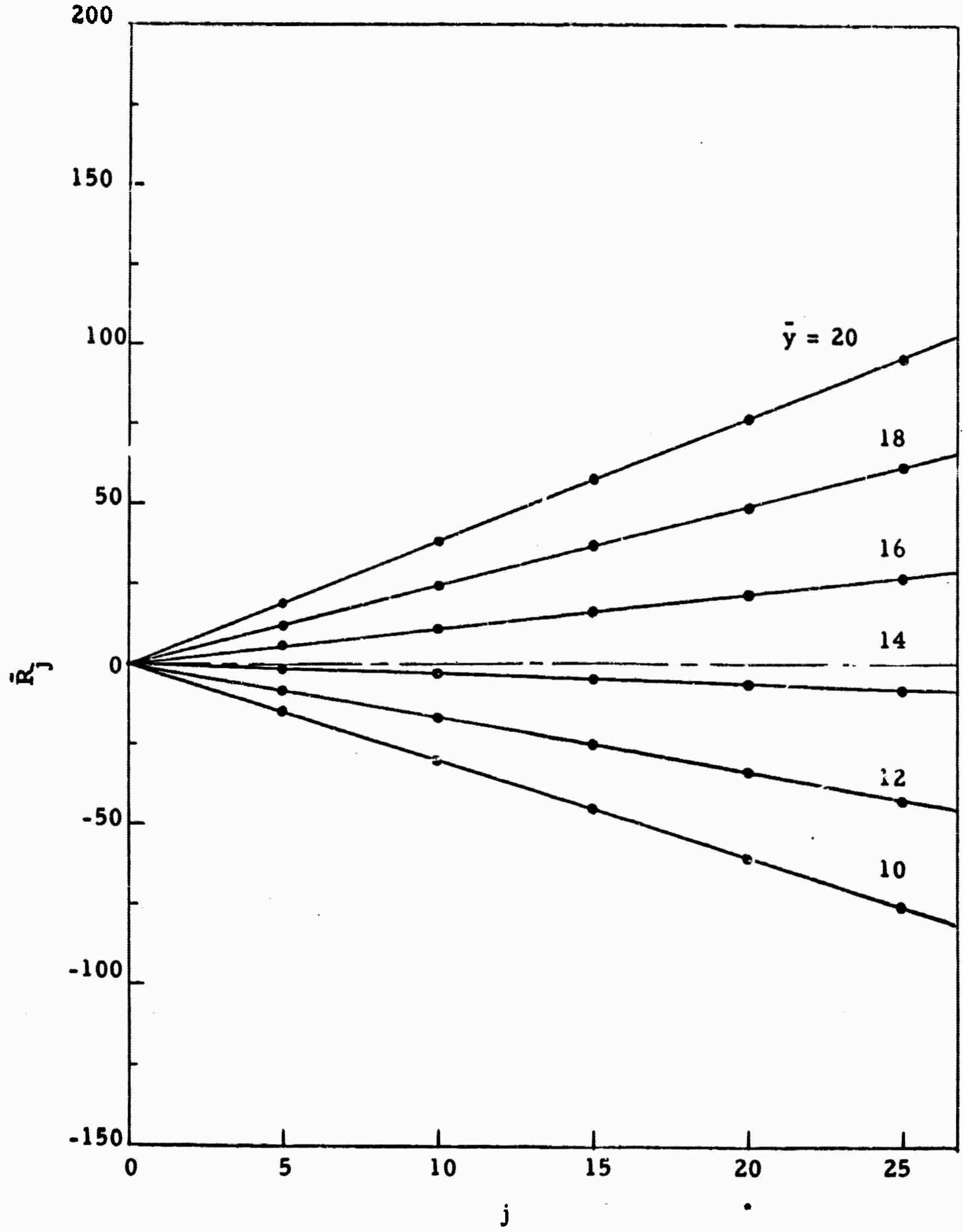

Figure 55. Computer-simulated mean cumulative information versus sample number for a sequential detector designed for two known signals, $\bar{y}_1 = 20$ and $\bar{y}_2 = 10$, and for various values $\bar{y}$ of actual signal present.

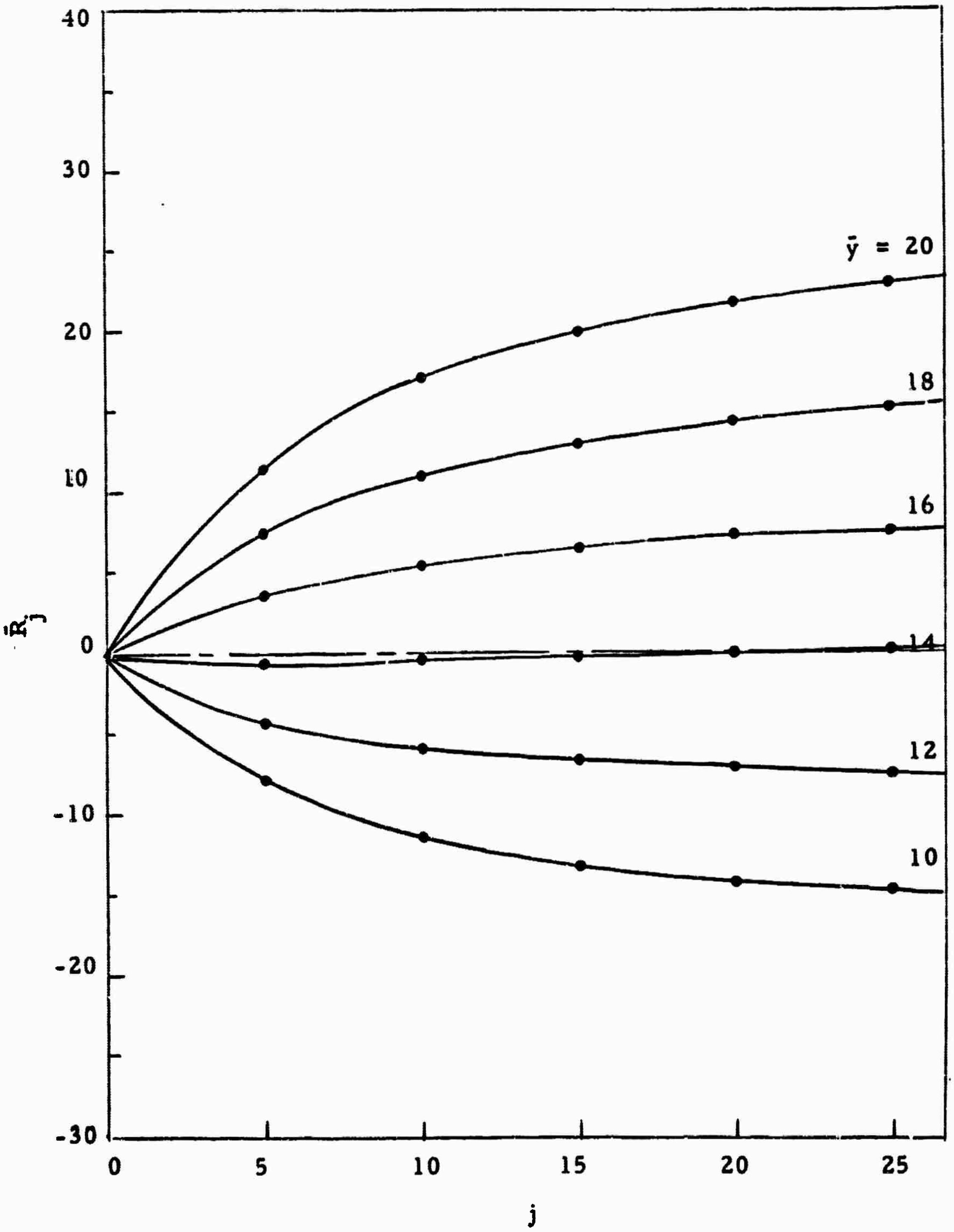

Figure 56. Computer-simulated mean cumulative information versus sample number for a sequential detector designed for two unknown signals and for various values $\bar{y}$ of actual signal present where $\bar{\bar{y}}_1 = 20$, $\bar{\bar{y}}_2 = 10$, and u = 100.

cumulative information versus sample number for a detector designed for two unknown signals, with mean rates $\bar{y}_1 = 20$ and $\bar{y}_2 = 10$, and for various values of actual signals present. Figures 36, 37, and 38 show the calculated mean cumulative information versus sample number for different values of u. These figures can be used to find graphical solutions of ASN for the case of $\bar{y}_1$ present (ω_1 as the state of nature). Two hundred trials were made for each of the simulated results of Figures 51, 52, 53, 54, 55, and 56.

To simulate the mean cumulative information, it was necessary to find the sample means $\bar{I}_1, \bar{I}_2, \ldots, \bar{I}_j$ (i.e., averages of 200 samples). These sample averages were then added to yield the simulated mean cumulative information.

Savings of ASN

One of the advantages of sequential detection over fixed-sample detection is in the savings of the average number of samples needed to achieve a specified reliability of decision. We are interested in making some comparison of these two detectors to obtain an idea of the difference of ASN.

Figure 57 shows the calculated ASN of the two-known-signals case versus error probability for the fixed-sample detector, where $p(\omega_1) = p(\omega_2) = 1/2$, and the sequential detector for both ω_1 and ω_2 as the states of nature. This figure shows that for the particular example considered,

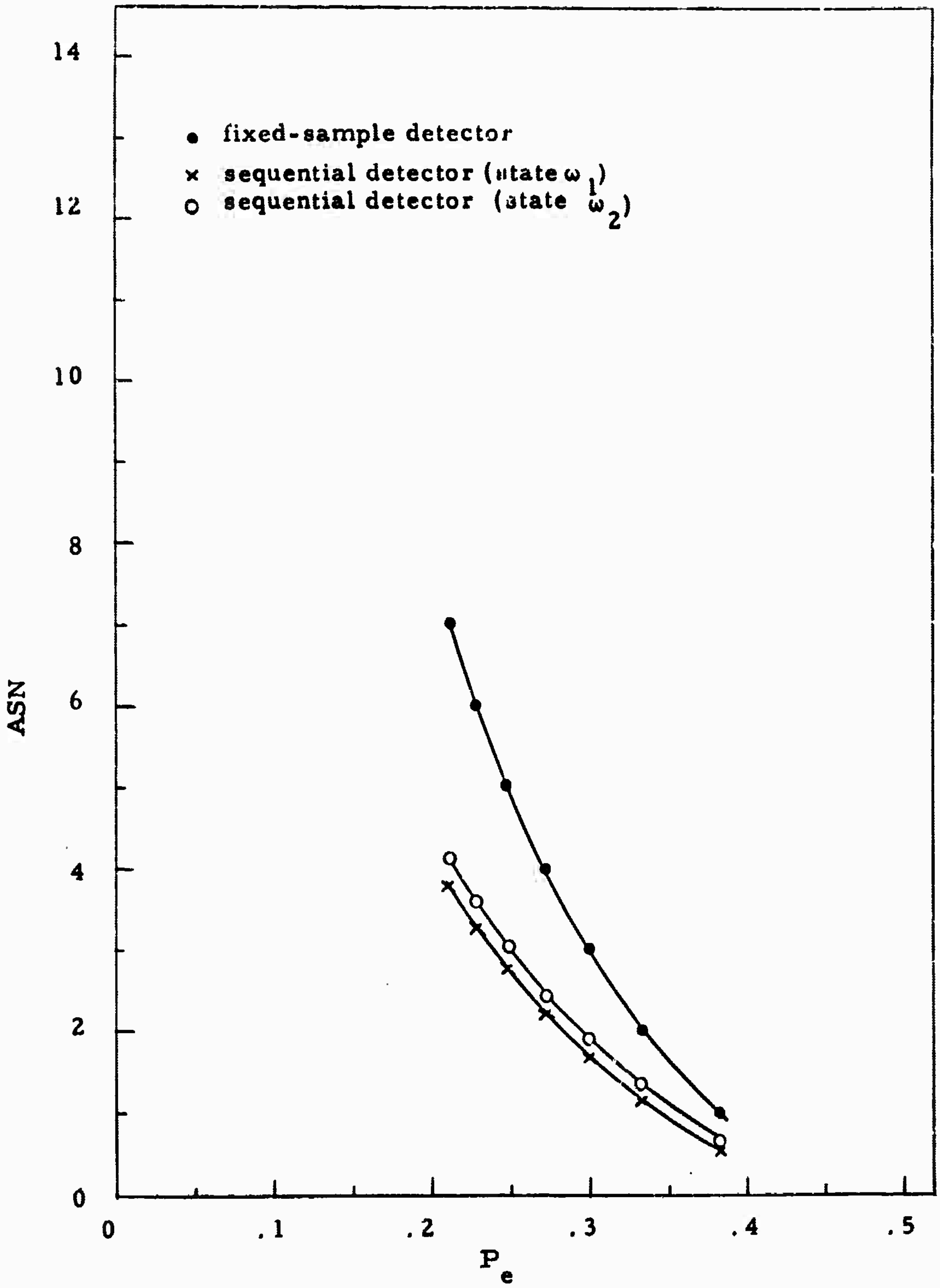

Figure 57. Calculated ASN for the two-known-signals case versus error probability for the fixed-sample detector ($p(\omega_1) = p(\omega_2) = 1/2$) and the sequential detector for both states of nature ω_1 and ω_2 where $\bar{y}_1 = 12$ and $\bar{y}_2 = 10$.

the sequential detector has about a 40% savings of ASN. Figure 58 shows the computer-simulated results corresponding to Figure 57. The simulated results show that for small sample numbers the fixed-sample detector actually does better than the sequential detector, but as the sample number becomes large the sequential detector is substantially better than the fixed-sample detector. Figure 59 shows the computer-simulated results corresponding to Figure 58 except that the sequential detector is designed for two unknown signals. Four hundred trials were made for each of the simulated results of Figures 58 and 59.

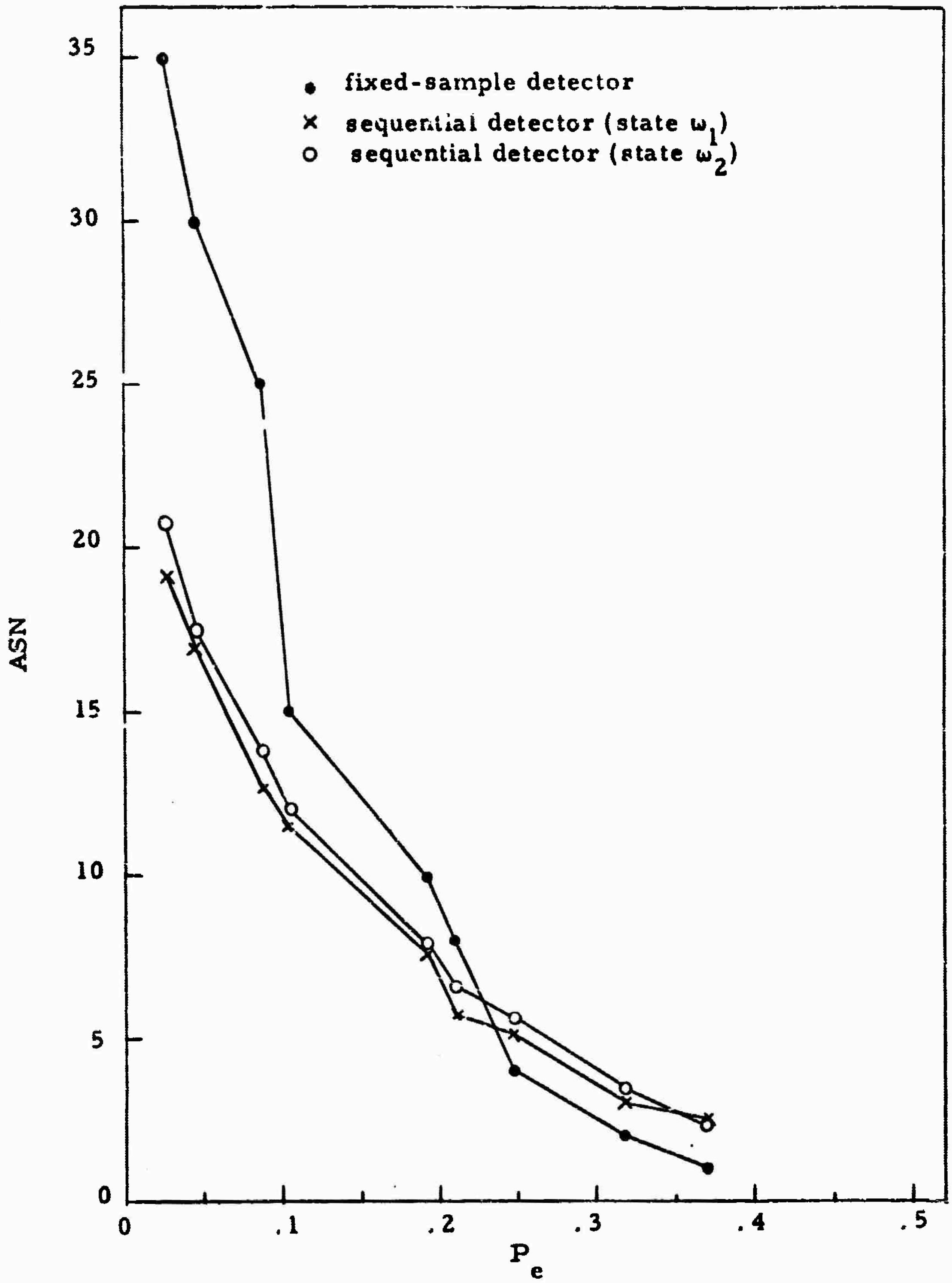

Figure 58. Computer-simulated ASN for the two-known-signals case versus error probability for the fixed-sample detector ($p(\omega_1) = p(\omega_2) = 1/2$) and the sequential detector for both states of nature ω_1 and ω_2 where $\bar{y}_1 = 12$ and $\bar{y}_2 = 10$.

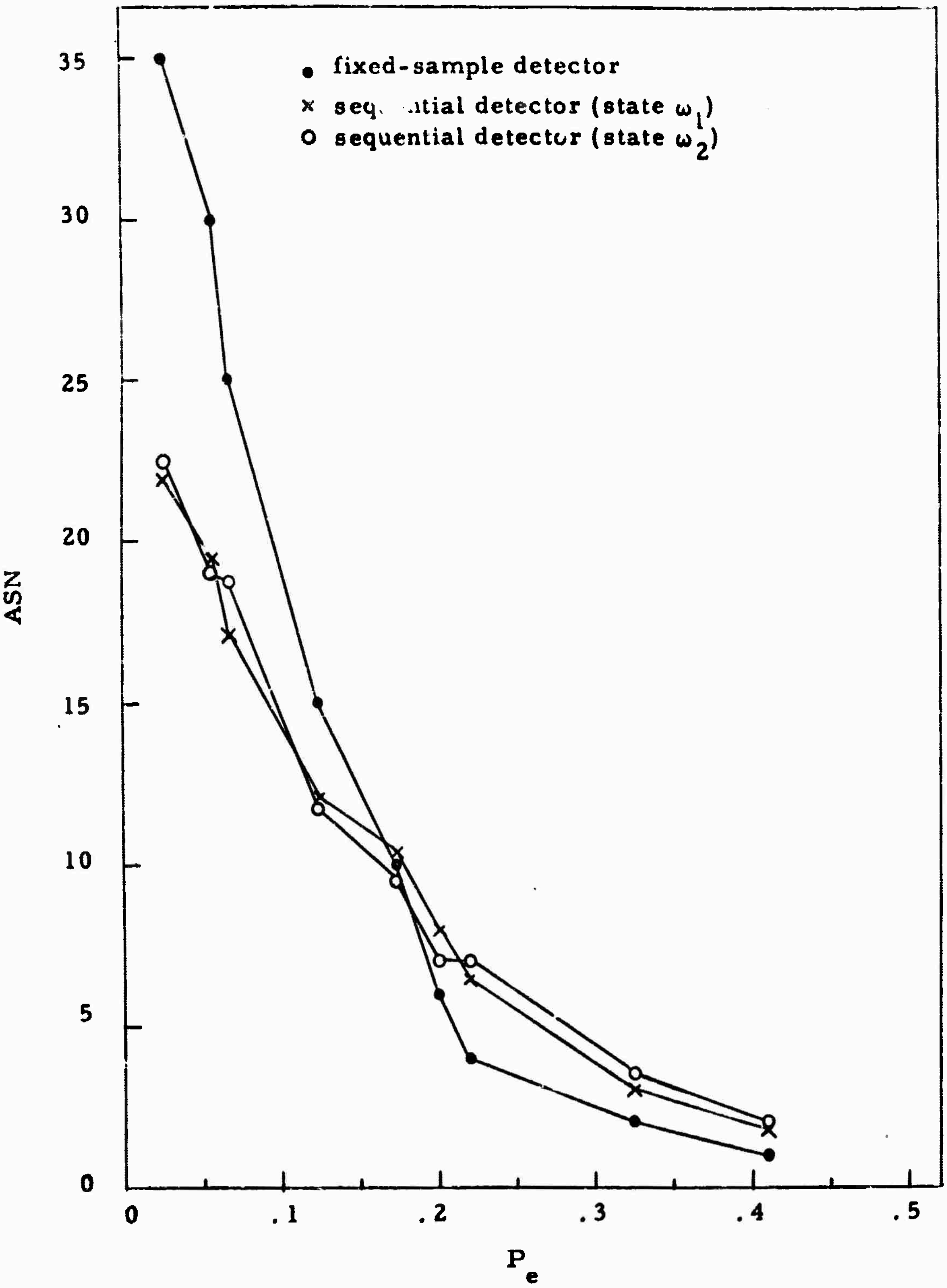

Figure 59. Computer-simulated ASN for the two-unknown-signals case versus error probability for the fixed-sample detector ($p(\omega_1) = p(\omega_2) = 1/2$) and the sequential detector for both states of nature ω_1 and ω_2 where $\bar{\bar{y}}_1 = 12$, $\bar{\bar{y}}_2 = 10$, and $u = 1000$.

SUMMARY AND CONCLUSIONS

In this paper the problems of estimating and discriminating between extended optical signals which have been distorted by diffraction, additive background noise, and multiplicative noise have been studied. Poisson statistics have been assumed throughout this paper since in many real situations they are more realistic than Gaussian statistics which are more commonly used.

For the estimation problem, an optimum linear estimate using the minimum mean-square-error criterion was considered. Detection noise (multiplicative noise) as well as additive noise was considered since for any measurement technique there will be some interaction between photons and matter which in turn will give rise to detection noise. The performance of the minimum mean-square-error estimation procedure was evaluated for several special cases. Some results pertinent to optimum sampling schemes were obtained for both white and colored noise.

Other estimates besides the minimum mean-square-error estimate were considered. These other estimates included the Bayes' estimate, the maximum likelihood estimate, and the maximum a posteriori estimate. These estimates were considered primarily to find estimates of the mean rate of signal-plus-noise photons $\bar{y}$ incident upon the image plane. These estimates were then used in an estimator-correlator in conjunction with fixed-sample detection. Given that $\bar{y}_i$ is taken from a gamma distribution with known parameters, the minimum mean-square-error estimate and the Bayes' estimate of $\bar{y}_i$ were found to be the same.

The Bayes' decision rule was used in the fixed-sample detection procedure to discriminate between extended optical signals. Only problems involving two possible states of nature, ω_1 and ω_2, were considered. Various amounts of a priori information were assumed about the two possible signals which may be present at the input of the detector. In the error analysis only a single cell of the image plane was considered. For cases of unknown signal parameters (mean rates), the parameters were assumed fixed but initially unknown having been taken from a known gamma distribution. Several estimator-correlators were considered. The Bayes' and minimum mean-square-error estimator-correlators were found to be superior to the maximum a posteriori estimator-correlator for the cases considered. It was found in general that if the error probabilities are greater than about 0.1, knowing the signal parameters (mean rates) is not much more of an advantage than knowing only the probability distributions from which they were taken, particularly if the Bayes' decision rule or the Bayes' estimator-correlator is used. For error probabilities << .1 the above statement does not hold and the difference in error probabilities becomes significant.

When speed is important in processing optical data, sequential detection should be considered. The sequential detection test developed by Wald (1947) minimizes the average test length and hence has a shorter average test length than the fixed-sample test for a given P(FA) and P(FD). A comparison was made of the test length of these two detectors to obtain some insight into the savings of time that the sequential detector has over the fixed-sample detector. Sequential detectors were derived for various cases of known and unknown optical signals and their performances compared. For the case of a sequential detector designed for two known signals, the information per sample remained constant for all samples. It was found that if only one of the signals for which the sequential detector was designed were unknown and if that particular state of nature (ω_1) were the state of nature present, the information per sample is smallest for the first sample and as more samples are taken the information per sample increase and approaches the constant information per sample of the two-known-signals case. The performance of the sequential detector for these cases thus improves as more samples are taken. It was also found that if the other state of nature (ω_2) were present the information per sample is greatest for the first sample and as more samples are taken the information per sample decreases and thus the performance of the sequential detector deteriorates as more samples are taken. For the case of the sequential detector designed for two unknown signals, the information per sample is greatest for the first sample and decreases as more samples are taken regardless of which state of nature is present. The average sample number and Operating

Characteristic Function (probability of choosing state ω_2 when signal $\bar{y}$ is present) versus actual signal $\bar{y}$ were determined for the cases mentioned above. For most of the cases considered, analytical solutions appeared extremely difficult to obtain; hence, the analysis was to a large extent carried out by computer simulation.

For future work in this area, it is recommended that the discrete sequential detector that has been discussed in this paper be considered in terms of a continuous sequential detector where only one observation is made and the duration time of the observation is varied. Also, multiple detection should be considered for cases with more than two states of nature.

APPENDIX

Two Dimensional Optical Transform Theory

Imaging Configuration

We will use Huyghen's principle in this development. According to Huyghen's principle, each point (α, β) acts as the source of a wavelet that propagates with the Green's function e^{ikr}/r for radiation from a point source where r is the distance from the source and k is equal to $2\pi\nu/c$ (c is the speed of light and ν is the frequency of the radiation). The electric field at the aperture is then found by adding the vector amplitudes of all the wavelets originating at the object plane. The strength of each wavelet depends on the complex amplitude of the electric field in the object plane.

We again employ Huyghen's principle just to the right of the aperture. Each point (ρ, γ) of the aperture acts as the source of a secondary wavelet that again propagates with the Green's function e^{ikr}/r. The total electric field at the image plane is found by adding the vector amplitudes of all the wavelets originating at the aperture plane where the strength of these wavelets depends on the incident electric field at the points of the aperture.

Let $E(\alpha, \beta)$ be the complex amplitude of the monochromatic source point in the object at point (α, β). Using Huyghen's principle, the electric field just to the left of the aperture is

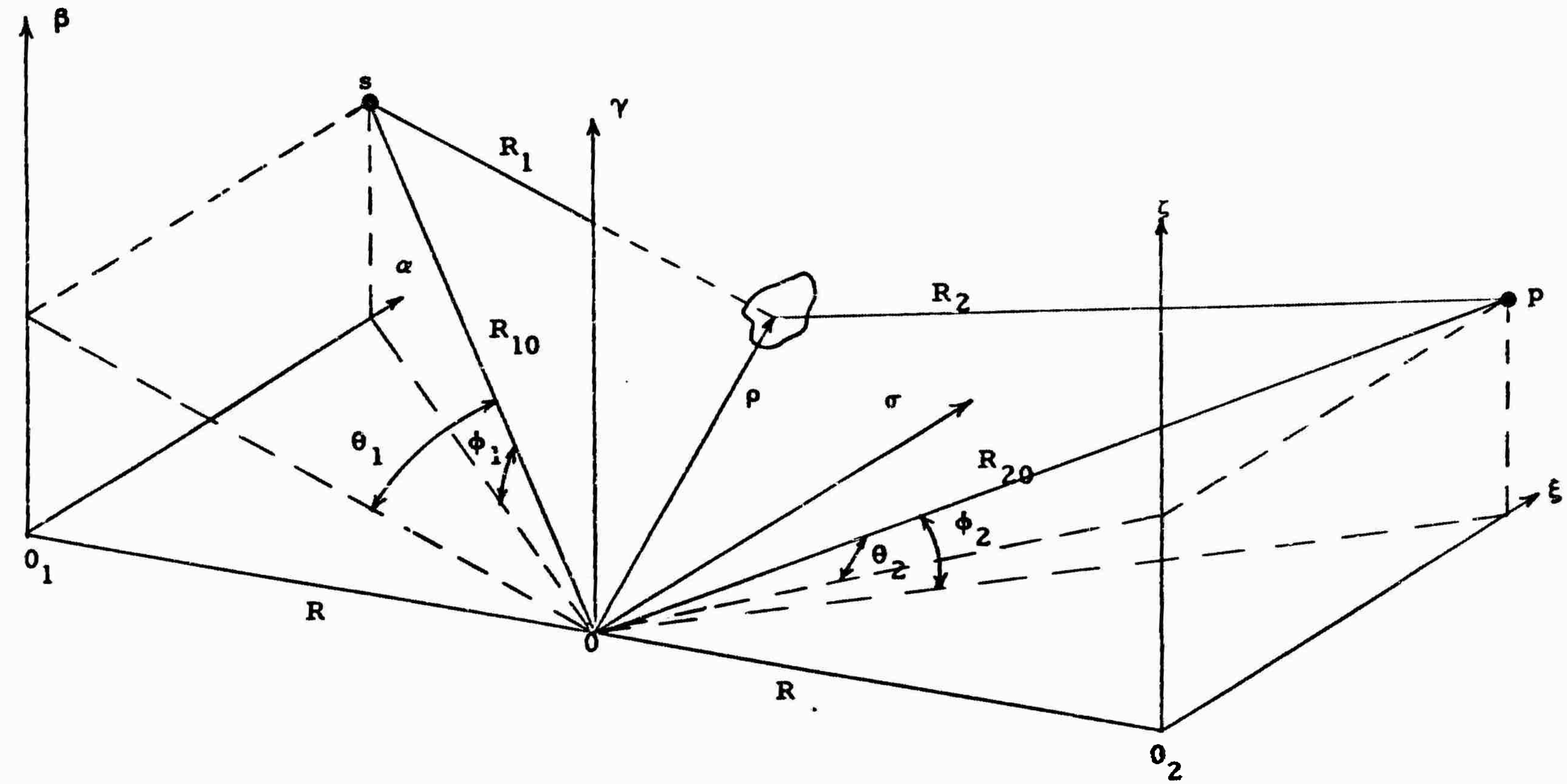

Figure 60. Optical imaging configuration.

$$E(\alpha, \beta) \frac{e^{ikR_1}}{R_1} . \tag{301}$$

We are assuming Fraunhofer diffraction; hence,

$$R_1 + R_{10} \approx 2R_{10} \approx 2R, \tag{302}$$

$$R_2 + R_{20} \approx 2R_{20} \approx 2R, \tag{303}$$

and

$$R_1 \approx R. \tag{304}$$

We can rewrite (301) as

$$E(\alpha, \beta)\, e^{ikR_1}. \tag{305}$$

$$R^2 = R_1^{\,2} - (\sigma - \alpha)^2 - (\gamma - \beta)^2 = R_{10}^{\,2} - \alpha^2 - \beta^2. \tag{306}$$

$$R^2 = R_2^{\,2} - (\sigma + \xi)^2 - (\gamma + \zeta)^2 = R_{20}^{\,2} - \xi^2 - \zeta^2. \tag{307}$$

$$R_1 - R_{10} = \frac{\rho^2}{R_1 + R_{10}} - \frac{2}{R_1 + R_{10}} (\sigma\alpha + \gamma\beta), \tag{308}$$

where $\rho^2 = \sigma^2 + \gamma^2$.

$$R_2 - R_{20} = \frac{\rho^2}{R_2 + R_{20}} + \frac{2}{R_2 + R_{20}} (\sigma\xi + \gamma\zeta). \tag{309}$$

Then

$$R_1 - R_{10} = \frac{\rho^2}{2R} - \frac{(\sigma\alpha + \gamma\beta)}{R}, \tag{310}$$

$$R_2 - R_{20} = \frac{\rho^2}{2R} + \frac{(\sigma\xi + \gamma\zeta)}{R}. \tag{311}$$

For small aperture, $\rho^2/2R$ can be neglected. Hence,

$$R_1 - R_{10} = \frac{-(\sigma\alpha + \gamma\beta)}{R} \tag{312}$$

and

$$R_2 - R_{20} = \frac{\sigma\xi + \gamma\zeta}{R}. \tag{313}$$

Let $f_\sigma = \sigma / \nu R$ and $f_\gamma = \Upsilon / \nu R$. We now can rewrite (305) as

$$E(\alpha,\beta)e^{ikR_{10}}e^{ik(R_1-R_{10})} = E(\alpha,\beta)e^{ikR_{10}}e^{-ik(\frac{\sigma\alpha+\gamma\beta}{R})} \tag{314}$$

$$= E(\alpha,\beta)e^{ikR_{10}}e^{-i2\pi(f_\sigma\alpha+f_\gamma\beta)}.$$

Coherent Light

Since R_{10} is approximately constant over the object pattern, we can neglect $e^{ikR_{10}}$ or absorb it into $E(\alpha,\beta)$ since it is only a constant phase term, and we cannot observe phase but only intensity.

The electric field at point (σ, Υ) just to the left of the aperture plane is the sum of all contributions at that point due to all the points in the object plane. This is true because amplitudes add for coherent light.

$$W(f_\sigma, f_\gamma) = \iint_{-\infty}^{\infty} E(\alpha,\beta)\, e^{-i2\pi(f_\sigma\alpha+f_\gamma\beta)}\, d\alpha d\beta. \tag{315}$$

Let $T(f_\sigma, f_\gamma)$ represent the aperture function. The electric field at point (σ, γ) just to the right of the aperture plane is then

$$U(f_\sigma, f_\gamma) = W(f_\sigma, f_\gamma)\, T(f_\sigma, f_\gamma). \tag{316}$$

Again using Huyghen's principle, the electric field at (ξ, ζ) in the image plane due to the point (σ, γ) in the aperture plane is proportional to

$$U(f_\sigma, f_\gamma)\, e^{ikR_{20}}\, e^{ik[\rho^2/2R + (\sigma\xi+\gamma\zeta)/R]}. \tag{317}$$

Neglecting the term $\rho^2/2R$ for small aperture and neglecting the constant phase term $e^{ikR_{20}}$ we can write

$$U(f_\sigma, f_\gamma)\, e^{ik(\sigma\xi+\gamma\zeta)/R} = U(f_\sigma, f_\gamma)\, e^{i2\pi(f_\sigma\xi+f_\gamma\zeta)}. \tag{318}$$

Now the total electric field at the point (ξ, ζ) in the image plane due to all the points in the aperture plane is

$$u(\xi, \zeta) = \iint_{-\infty}^{\infty} U(f_\sigma, f_\gamma)\, e^{i2\pi(f_\sigma\xi+f_\gamma\zeta)}\, df_\sigma df_\gamma. \tag{319}$$

This is the inverse Fourier transform of $U(f_\sigma, f_\gamma)$, i.e.,

$$F\{u(\xi, \zeta)\} = U(f_\sigma, f_\gamma) = W(f_\sigma, f_\gamma)\, T(f_\sigma, f_\gamma). \tag{320}$$

Using the convolution theorem we can write

$$u(\xi, \zeta) = E(\xi, \zeta) * t(\xi, \zeta) \tag{321}$$

where

$$t(\xi, \zeta) = \iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\, e^{i2\pi(f_\sigma \xi + f_\gamma \zeta)}\, df_\sigma\, df_\gamma \tag{322}$$

is the impulse response of the system.

In the image plane, intensity is the quantity that is actua observed and not the amplitude. That is, we observe

$$|u(\xi, \zeta)|^2 = v(\xi, \zeta). \tag{323}$$

<u>Incoherent Light</u>

For incoherent light, intensities rather than amplitudes add. We will now find the expression for the intensity at (ξ, ζ) due to a point source in the object plane (α, β) and then integrate over the object plane.

In this case we can take advantage of the previous discussion and write for the electric field at point (σ, γ) just to the right of the aperture the expression

$$U_{\alpha,\beta}(f_\sigma, f_\gamma) = T(f_\sigma, f_\gamma)E(\alpha,\beta)e^{ikR_{10}}e^{-i2\pi(f_\sigma \alpha + f_\gamma \beta)} \tag{324}$$

where the subscripts α and β infer that this expression is due to the point source at point (α, β). The electric field at (ξ, ζ) due to the electric field at (σ, γ) is

$$U_{\alpha,\beta}(f_\sigma, f_\gamma)e^{ikR_{20}}e^{ik(\sigma\xi+\gamma\zeta)/R}. \tag{325}$$

Let $f_\sigma = \sigma/\nu R$ and $f_\gamma = \gamma/\nu R$. Then we have

$$U_{\alpha,\beta}(f_\sigma, f_\gamma)e^{ikR}e^{i2\pi(f_\sigma\xi+f_\gamma\zeta)} = T(f_\sigma, f_\gamma)E(\alpha,\beta)e^{ik(R_{10}+R_{20})} \cdot e^{i2\pi(f_\sigma[\xi-\alpha] + f_\gamma[\zeta-\beta])}. \tag{326}$$

The total electric field at point (ξ, ζ) due to the point source at point (α, β) is

$$u_{\alpha,\beta}(\xi, \zeta) = e^{ik(R_{10}+R_{20})} \iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)E(\alpha,\beta)e^{i2\pi(f_\sigma[\xi-\alpha]+f_\gamma[\zeta-\beta])} df_\sigma df_\gamma. \tag{327}$$

The intensity (the observable) at (ξ, ζ) due to the point source at point (α, β) is

$$v_{\alpha,\beta}(\xi,\zeta) = |u_{\alpha,\beta}(\xi,\zeta)|^2 = u_{\alpha,\beta}(\xi,\zeta)\, u_{\alpha,\beta}{}^*(\xi,\zeta)$$

$$= |E(\alpha,\beta)|^2 \iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\, T^*(f_\sigma', f_\gamma')\, e^{-i2\pi(\alpha-\xi)(f_\sigma - f_\sigma')} \cdot e^{-i2\pi(\beta-\zeta)(f_\gamma - f_\gamma')}\, df_\sigma\, df_\gamma\, df_\sigma'\, df_\gamma'. \quad (328)$$

To find the total intensity at (ξ, ζ) due to all the point sources in the object plane we must integrate over the object plane. That is,

$$v(\xi,\zeta) = |u(\xi,\zeta)|^2 = \iiiint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\, T^*(f_\sigma', f_\gamma')\, W(f_\sigma - f_\sigma', f_\gamma - f_\gamma') \cdot e^{i2\pi\xi(f_\sigma - f_\sigma')}\, e^{i2\pi\zeta(f_\gamma - f_\gamma')}\, df_\sigma\, df_\gamma\, df_\sigma'\, df_\gamma' \quad (329)$$

where

$$W(f_\sigma - f_\sigma', f_\gamma - f_\gamma') = \iint_{-\infty}^{\infty} |E(\alpha,\beta)|^2\, e^{-i2\pi\alpha(f_\sigma - f_\sigma')}\, e^{-i2\pi\beta(f_\gamma - f_\gamma')}\, d\alpha\, d\beta. \quad (330)$$

Now let $|E(\alpha,\beta)|^2 = w(\alpha,\beta)$ and

$$z = f_\sigma - f'_\sigma,\ f'_\sigma = f_\sigma - z,\ df'_\sigma = -dz,$$

$$y = f_\gamma - f'_\gamma,\ f'_\gamma = f_\gamma - y,\ \text{and } df'_\gamma = -dy. \qquad (331)$$

Then

$$v(\xi,\zeta) = |u(\xi,\zeta)|^2 = \iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\ T^*(f_\sigma - z, f_\gamma - y)\ e^{i2\pi(\xi z+\zeta y)}$$

$$\cdot\, w(z,y)\ df_\sigma\, df_\gamma\, dz\ dy \qquad (332)$$

$$= \iint_{-\infty}^{\infty} w(z,y) e^{i2\pi(\xi z+\zeta y)} \left[\iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\ T^*(f_\sigma - z, f_\gamma - y) df_\sigma\, df_\gamma\right] dz\ dy.$$

Let

$$A(z,y) = \iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma) T^*(f_\sigma - z,\ f_\gamma - y)\ df_\sigma\, df_\gamma. \qquad (333)$$

$$A(z,y) = \quad T(f_\sigma, f_\gamma) * T'(f_\sigma, f_\gamma). \qquad (334)$$

$$a = F\{A\} = F\{T^*T\}. \qquad (335)$$

Thus

$$v(\xi, \zeta) = \iint_{-\infty}^{\infty} A(z, y)\ W\ (z, y)\ e^{i2\pi(\xi z + \zeta y)}\ dzdy. \tag{336}$$

and

$$V(f_\sigma, f_\gamma) = A(f_\sigma, f_\gamma)\ W\ (f_\sigma, f_\gamma). \tag{337}$$

The quantity $A(f_\sigma, f_\gamma)$ is the transfer function for the incoherent light case. Using the convolution theorem we can write

$$v(\xi, \zeta) = a(\xi, \zeta) * w(\xi, \zeta) \tag{338}$$

where $a(\xi, \zeta)$ is the point spread function of the optical system.

Derivation of Point Spread Functions

The Fourier transform pair in two dimensions is

$$T(f_\sigma, f_\gamma) = \iint_{-\infty}^{\infty} t(\xi, \zeta)\ e^{-i(\xi\omega_\sigma + \zeta\omega_\gamma)}\ d\xi d\zeta, \tag{339}$$

$$t(\xi,\zeta) = \iint_{-\infty}^{\infty} T(f_\sigma, f_\gamma)\, e^{i(\xi\omega_\sigma + \zeta\omega_\gamma)}\, df_\sigma\, df_\gamma. \tag{340}$$

Rectangular Aperture

Consider the rectangular aperture with sides a and b. For this case the aperture function is

$$T(\sigma,\gamma) = \sqcap(\sigma/a)\sqcap(\gamma/b) \tag{341}$$

where $T(\sigma) = \sqcap(\sigma/a)$ is defined by

$$T(\sigma) = \begin{cases} 0, & |\sigma| > a/2 \\ \tfrac{1}{2}, & |\sigma| = a/2 \\ 1, & |\sigma| < a/2. \end{cases} \tag{342}$$

This can be rewritten in spatial-frequency coordinates by defining $f_\sigma = \sigma/\nu R$ and $f_\sigma = \gamma/\nu R$. Hence,

$$T(f_\sigma, f_\gamma) = \sqcap(f_\sigma \nu R/a)\sqcap(f_\gamma \nu R/b). \tag{343}$$

The point spread function of intensity is the square of the two-dimensional Fourier transform of $T(f_\sigma, f_\gamma)$.

$$t(\xi,\zeta) = \int_{-\infty}^{\infty}\int \sqcap(f_\sigma \nu R/a) \sqcap(f_\gamma \nu R/b)\, e^{i(\xi\omega_\sigma + \zeta\omega_\gamma)}\, df_\sigma\, df_\gamma \tag{344}$$

$$= \left[\int_{-\infty}^{\infty} \sqcap(f_\sigma \nu R/a)\, e^{i2\pi f_\sigma \xi} df_\sigma\right]\left[\int_{-\infty}^{\infty} \sqcap(f_\gamma \nu R/b)\, e^{i2\pi f_\gamma \zeta} df_\gamma\right]$$

$$= \frac{ab}{\nu^2 R^2} \text{ sinc } (a\xi/\nu R) \text{ sinc } b\zeta/\nu R).$$

Normalizing this expression we have

$$\frac{t(\xi,\zeta)}{t(0,0)} = \text{sinc } (a\xi/\nu R) \text{ sinc } (b\zeta/\nu R) \tag{345}$$

or for the normalized intensity we have

$$\left[\frac{t(\xi,\zeta)}{t(0,0)}\right]^2 = \text{sinc}^2 (a\xi/\nu R) \text{ sinc}^2 (b\zeta/\nu R). \tag{346}$$

This is the point spread function or impulse response due to a point source.

Slit Aperture

Investigate the point spread function of an infinite slit and an infinite line source. This is the same as considering a one dimensional case where we have a point source and a one-dimensional aperture. The aperture function for this case is

$$T(\sigma) = \sqcap(\sigma/D) \tag{347}$$

where D is the aperture width. Written in spatial-frequency coordinates (347) becomes

$$T(f_\sigma) = \sqcap(f_\sigma \nu R/D). \tag{348}$$

The point spread function is the square of the Fourier transform of $T(f_\sigma)$.

$$t(\xi) = \int_{-\infty}^{\infty} \sqcap(f_\sigma \nu R/D)\, e^{i\xi\omega_\sigma} df_\sigma = \int_{-D/\nu R}^{D/\nu R} e^{i2\pi f_\sigma \xi} df_\sigma \tag{349}$$

$$= (D/\nu R)\,\mathrm{sinc}\,(D\xi/\nu R).$$

The normalized intensity is then

$$\left[\frac{t(\zeta)}{t(0)}\right]^2 = \mathrm{sinc}^2(D\xi/\lambda R). \tag{350}$$

Circular Aperture

For mathematical convenience rewrite the two-dimensional Fourier transform pair in polar coordinates. Let

$$\begin{aligned} f_\sigma &= f\cos\theta, \quad \xi = r\cos\phi, \\ f_\gamma &= f\sin\theta, \text{ and } \zeta = r\sin\phi. \end{aligned} \tag{351}$$

To make the transformation, use the relationships (Olmsted, 1956)

$$\iint f(\xi, \zeta)\,d\xi\,d\zeta = \iint g(r,\phi)\left|J(r,\phi)\right| dr\,d\phi \tag{352}$$

where $g(r, \phi) = f(\xi(r, \phi), \zeta(r, \phi))$ and

$$\iint h(f_\sigma, f_\gamma)\,df_\sigma\,df_\gamma = \iint k(f, \theta)\left|J(f, \theta)\right| df\,d\theta \tag{353}$$

where $h(f, \theta) = h(f_\sigma(f, \theta), f_\gamma(f, \theta))$.

$$J(r, \phi) = \frac{\partial(\xi, \eta)}{\partial(r, \phi)} = \begin{vmatrix} \dfrac{\partial r \cos\phi}{\partial r} & \dfrac{\partial r \cos\phi}{\partial \phi} \\ & \\ \dfrac{\partial r \sin\phi}{\partial r} & \dfrac{\partial r \sin\phi}{\partial \phi} \end{vmatrix} = r. \tag{354}$$

$$J(f, \theta) = \frac{\partial(f_\sigma, f_\gamma)}{\partial(f, \theta)} = \begin{vmatrix} \dfrac{\partial f \cos\theta}{\partial f} & \dfrac{\partial f \cos\theta}{\partial \theta} \\ & \\ \dfrac{\partial f \sin\theta}{\partial f} & \dfrac{\partial f \sin\theta}{\partial \theta} \end{vmatrix} = f. \tag{355}$$

Making these substitutions yields:

$$t(r, \phi) = \int_0^\infty \int_0^{2\pi} T(f, \theta)\, e^{i(r\omega\cos\phi\cos\theta + \omega r\sin\phi\sin\theta)} f\,df\,d\theta, \tag{356}$$

$$T(f, \theta) = \int_0^\infty \int_0^{2\pi} t(r, \phi)\, e^{-i(\omega r\cos\theta\cos\phi + \omega r\sin\theta\sin\phi)} r\,dr\,d\phi. \tag{357}$$

If $T(f, \theta)$ and $t(r, \phi)$ are symmetrical and have no dependency upon θ and ϕ we can write

$$t(r) = \int_0^\infty T(f)\left[\int_0^{2\pi} e^{i\omega r \cos(\phi-\theta)} d\theta\right] f df \qquad (358)$$

$$= 2\pi \int_0^\infty T(f)\, J_o(\omega r) f df$$

and

$$T(f) = 2\pi \int_0^\infty t(r)\, J_o(\omega r) r dr. \qquad (359)$$

The point spread function of intensity is the square of the two-dimensional Fourier transform of T(f).

Let a be the radius of the aperture and ρ be the radial coordinate in the aperture plane. The spatial frequency is $f = \rho / \nu R$.

$$t(r) = 2\pi \int_0^{a/\nu R} J_o(\omega r) f df = \frac{1}{2\pi r^2} \int_0^{2\pi r a/\nu R} J_o(x) x dx$$

$$= \frac{\pi a^2}{\nu^2 R^2}\left[\frac{2 J_1(2\pi r a/\nu R)}{2\pi r a/\nu R}\right]^2 . \qquad (360)$$

The normalized intensity is then

$$\left[\frac{t(r)}{t(o)}\right]^2 = \left[\frac{2\,J_1(2\pi ra/\nu R)}{2\pi ra/\nu R}\right]^2. \tag{361}$$

Derivation of Optimum Sampling Scheme

Consider the case of white noise where the point sources are separated by the Rayleigh criterion distance. We will assume that the number of measurements ℓ is equal to a multiple of the number of point sources km. Our problem is to find a sampling procedure to minimize

$$\mathrm{tr}\,(A'A)^{-1}. \tag{362}$$

If ℓ measurements are made at the peak of each point spread function then

$$A'A = \begin{pmatrix} k & & 0 \\ & \ddots & \\ 0 & & k \end{pmatrix} \tag{363}$$

and

$$(A'A)^{-1} = \begin{pmatrix} 1/k & & 0 \\ & \ddots & \\ 0 & & 1/k \end{pmatrix}. \tag{364}$$

Therefore, for m point sources

$$\text{tr}\,(A'A)^{-1} = m/k = m^2/_{\ell}\,. \qquad (365)$$

We want to show that these are the conditions necessary to minimize $\text{tr}(A'A)^{-1}$ when we have white noise and Rayleigh criterion distances between point sources.

To prove the above we need to consider several lemmas.

Lemma 1

A'A is a symmetric matrix.

Proof:

$$(A'A)' = A'(A')' = A'A\,. \qquad (366)$$

Lemma 2

A'A is a positive definite matrix.

Proof: By definition, A'A is a positive definite matrix if $\langle x, A'Ax\rangle > 0$ (Zadeh and Desoer, 1963) where $\langle\ \rangle$ denotes an inner product. Since A'A is a symmetric matrix with real valued elements, it is a Hermitian matrix. Hence,

$$(A'A)^{\dagger} \triangleq [(A'A)^{*}]' = (A'A)' = A'A. \qquad (367)$$

Consider Schwartz's inequality

$$\langle x, x\rangle\langle y, y\rangle \geq \langle x, y\rangle\langle y, x\rangle = |\langle x, y\rangle|^2 \tag{368}$$

where equality holds if and only if y = cx where c is a scalar constant. Hence, we have

$$\langle x, x\rangle\langle y, y\rangle \geq \langle x, y\rangle\langle y, x\rangle \geq 0. \tag{369}$$

For any intelligent estimation procedure

$$\langle x, x\rangle > 0 \text{ and } |c| > 0. \tag{370}$$

For y = cx we have

$$\langle x, x\rangle\langle y, y\rangle = \langle x, y\rangle\langle x, y\rangle = |c|^2 \langle x, x\rangle\langle x, x\rangle > 0 \tag{371}$$

which implies that

$$\langle y, y\rangle > 0. \tag{372}$$

Let y = Ax, then

$$\langle y, y\rangle = \langle Ax, Ax\rangle = \langle x, A^{+}Ax\rangle = \langle x, A'Ax\rangle > 0. \tag{373}$$

Now consider the case of $y \neq cx$. We now have

$$\langle x, x\rangle\langle y, y\rangle > \langle x, y\rangle\langle x, y\rangle \geq 0 \quad (374)$$

hence,

$$\langle y, y\rangle > 0. \quad (375)$$

Let $y = Ax$, then

$$\langle y, y\rangle = \langle Ax, Ax\rangle = \langle x, A^{\dagger}Ax\rangle = \langle x, A'Ax\rangle > 0. \quad (376)$$

Hence, A'A is a positive definite matrix.

<u>Lemma 3</u>

Let $A'A = B$.

$$|B| = b_{ii}\,|B_{ii}| - \sum_{j,\,k=1}^{m} b_{ji} b_{ik} |B_{ii \cdot jk}| \quad (377)$$

where $|B_{ii.\,jk}|$ is the cofactor of b_{jk} in B_{ii} and $|B_{ii}|$ is the cofactor of b_{ii} in B.

Proof:

$$|B| = \sum_{j=1}^{m} b_{ji} |B_{ji}| = b_{ii}|B_{ii}| + \sum_{\substack{j=1 \\ j\neq i}}^{m} b_{ji}|B_{ji}|. \quad (378)$$

Hence, we need to show

$$b_{ji}\,|B_{ji}| = -b_{ji}\sum_{\substack{k=1\\k\neq i}}^{m} b_{ik}\,|B_{ii\cdot jk}|\,. \qquad (379)$$

By definition,

$$|B_{ij}| = (-1)^{i+j}\,M_{ij} \qquad (380)$$

and

$$|B_{ii\cdot jk}| = (-1)^{i+i+j+k}\,M_{ii\cdot jk} \qquad (381)$$

where M_{ij} and $M_{ii\cdot jk}$ are complementary minors (Wylie, 1960). For $j < i$ we can write

$$\begin{aligned} b_{ji}|B_{ji}| &= b_{ji}[-b_{i1}(-1)^{i+i+j+1}M_{ii\cdot j1} - b_{i2}(-1)^{i+i+j+2}M_{ii\cdot j2} \\ &\quad -\ldots-(-1)b_{ii-1}(-1)^{i+i+j+(i-1)}M_{ii\cdot ji-1} + b_{ii+1}(-1)^{i+i+j+(i+1)}M_{ii\cdot ji+1} \\ &\quad +\ldots+b_{in}(-1)^{i+i+j+m}M_{ii\cdot jm}] \\ &= b_{ji}[-b_{i1}|B_{ii\cdot j1}| - b_{i2}|B_{ii\cdot j2}| - \ldots - b_{ii-1}|B_{ii\cdot ji-1}| \\ &\quad - b_{ii+1}|B_{ii\cdot ji+1}| - b_{in}|B_{ii\cdot jm}| = -b_{ji}\sum_{\substack{k=1\\k\neq i}}^{m} b_{ik}|B_{ii\cdot jk}|\,. \end{aligned} \qquad (382)$$

For $j > i$

$$b_{ji}|B_{ji}| = b_{ji}\,[b_{i1}(-1)^{i+i+j+1}M_{ii\cdot j1} + b_{i2}(-1)^{i+i+j+2}M_{ii\cdot j2}$$

$$+ \ldots + b_{ii-1}(-1)^{i+i+j+(i-1)}M_{ii\cdot ji-1} - b_{ii+1}(-1)^{i+i+j+(i+1)}M_{ii\cdot ji+1}$$

$$- \ldots - b_{in}(-1)^{i+i+j+m}\,M_{ii\cdot jm}] \tag{383}$$

$$= b_{ji}[-b_{i1}|B_{ii\cdot j1}| - b_{i2}|B_{ii\cdot j2}| - \ldots - b_{ii-1}|B_{ii\cdot ji-1}|$$

$$- b_{ii+1}|B_{ii\cdot ji+1}| - b_{in}|B_{ii\cdot jm}| = -b_{ji}\sum_{\substack{k=1\\k\neq}}^{m} b_{ik}|B_{ii\cdot jk}|.$$

Hence,

$$b_{ji}|B_{ji}| = -b_{ji}\sum_{\substack{k=1\\k\neq i}}^{m} b_{ik}|B_{ii\cdot jk}| \tag{384}$$

and therefore

$$|B| = b_{ii}|B_{ii}| - \sum_{\substack{j,\,k=1\\j\neq i\\k\neq i}}^{m} b_{ji}b_{ik}|B_{ii\cdot jk}|. \tag{385}$$

Lemma 4

If $B = A'A$ is a symmetric, real, and positive definite matrix then B^{-1} is positive definite.

Proof: Given $\langle x, Bx \rangle > 0$. Let $x = B^{-1}y$, then

$$\langle x, Bx \rangle = \langle B^{-1}y, BB^{-1}y \rangle = \langle B^{-1}y, y \rangle = \langle y, B^{-1}y \rangle > 0. \tag{386}$$

Hence, B^{-1} is a positive definite matrix.

Lemma 5

$$\sum_{\substack{j,k=1 \\ j \neq i \\ k \neq i}}^{m} b_{ji}\, b_{ik}\, |B_{ii \cdot jk}| > 0 \tag{387}$$

for b_{ji} and b_{ik} not equal to zero.

Proof: The matrix $B_{ii} = (A'_{\neq i} A_{\neq i})$ is a positive definite matrix since the matrix B_{ii} is a submatrix whose principle diagonal lies along the principle diagonal of B. Therefore, $|B_{ii}|$ is positive. This also implies that B_{ii}^{-1} is positive definite. We can write

$$\sum_{\substack{j,k=1\\ j\neq i\\ k\neq i}}^{m} b_{ji}\, b_{ik} \left| B_{ii\cdot jk} \right| = \left| B_{ii} \right| \sum_{\substack{j,k=1\\ j\neq i\\ k\neq i}}^{m} \frac{\left| B_{ii\cdot jk} \right|}{\left| B_{ii} \right|}\, b_{ji} b_{ik}$$

$$= \left| B_{ii} \right| \sum_{\substack{j,k=1\\ j\neq i\\ k\neq i}}^{m} b_{ji} \left| B_{ii}^{-1} \right| b_{ik} = \left| B_{ii} \right| < b_i, B_{ii}^{-1} b_i > \qquad (388)$$

The quantity $<b_i, B_{ii}^{-1} b_i> \; > 0$ for $b_i \neq 0$ since B_{ii}^{-1} is a positive definite matrix. Hence,

$$\sum_{\substack{j,k=1\\ j\neq i\\ k\neq i}}^{m} b_{ji}\, b_{ik} \left| B_{ii\cdot jk} \right| > 0 \qquad (389)$$

for b_{ji} and b_{ik} not equal to zero.

Lemma 6

If $B = (A'A)$ is a symmetric positive definite matrix:

(a) $|B| = b_{ii} \, |B_{ii}|$ for each i if B is diagonal,

(b) $|B| < b_{ii} \, |B_{ii}|$ for at least one value of i if B is not diagonal.

Proof:

(a) If B is diagonal, then an expansion by cofactors along the i^{th}

row will give $b_{ii}|B_{ii}|$. Hence,

$$|B| = b_{ii} \, |B_{ii}| \; . \tag{390}$$

(b) Using lemmas 3 and 5 we can write (assuming that b_{ji} and b_{ik} are the off-diagonal elements which are not equal to zero)

$$|B| = b_{ii} \, |B_{ii}| - \sum_{\substack{j,k=1 \\ j \neq i}}^{m} b_{ji} b_{ij} |B_{ii.jj}| < b_{ii} |B_{ii}| \; . \tag{391}$$

Hence,

$$|B| < b_{ii} \, |B_{ii}| \; . \tag{392}$$

Note: By hypotheses B is not a diagonal matrix; hence, there are at least two non-zero symmetrically located off-diagonal elements (since B is symmetrical). Assume that these two non-zero elements are b_{ji} and b_{ij}.

Lemma 7

B is a symmetric, positive definite matrix with diagonal elements k_i and with arbitrary off-diagonal elements such that the matrix is positive definite. The elements along the principle diagonal of B^{-1} will be a minimum of $1/k_i$ if and only if B is a diagonal matrix.

Proof: If B is a diagonal matrix then

$$B^{-1} = (1/k_i)\ \delta_{ij} = \begin{pmatrix} 1/k_1 & & 0 \\ & \ddots & \\ 0 & & 1/k_m \end{pmatrix} . \qquad (393)$$

From lemma 6

$$0 < |B| \le b_{ii}\ |B_{ii}| \qquad (394)$$

or

$$\frac{1}{b_{ii}} \le \frac{|B_{ii}|}{|B|} \qquad (395)$$

since $b_{ii} > 0$ and $|B| > 0$. The expression

$$\left\{b_{ij}^{-1}\right\} = B^{-1} = \left\{\frac{|B_{ij}|}{|B|}\right\} \qquad (396)$$

implies

$$b_{ii}^{-1} = \frac{|B_{ii}|}{|B|} . \qquad (397)$$

Since

$$b_{ii}^{-1} = \frac{|B_{ii}|}{|B|} \tag{398}$$

then

$$b_{ii}^{-1} \geq \frac{1}{b_{ii}} = \frac{1}{k_i} . \tag{399}$$

If B is not a diagonal matrix then

$$b_{ii}^{-1} = \frac{1}{b_{ii}} = \frac{1}{k_i} . \tag{400}$$

Hence, if B is not a diagonal matrix then

$$|B| < b_{ii} |B_{ii}| \tag{401}$$

for some $i = i_o$ which implies that

$$b_{i_o i_o}^{-1} = \frac{|B_{i_o i_o}|}{|B|} > \frac{1}{b_{i_o i_o}} = \frac{1}{k_{i_o}} . \tag{402}$$

Lemma 8

$\sum_{i=1}^{m} 1/a_i$ is a minimum when the a's are equal (i.e., $a_1 = a_2 = \dots = a_m = \ell/m$) given that $\sum_{i=1}^{m} a_i = \ell$.

Proof: We will use dynamic programming to show this proof. Given the function

$$f(a_1, a_2, \dots, a_m) = \frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_m} . \quad (403)$$

we want to minimize $f(a_1, \dots, a_m)$ for $a_1 \geq 0$, $a_2 \geq 0, \dots,$ and $a_m \geq 0$ and subject to the constraint $a_1 + a_2 + \dots + a_m = \ell$.

Hence we can write

$$\min_{0 \leq a_m \leq \ell_m} \min_{\substack{a_1 \geq 0 \\ \vdots \\ a_{m-1} \geq 0}} \left[\frac{1}{a_1} + \dots + \frac{1}{a_{m-1}} + \frac{1}{a_m} \right]$$

$$= \min_{0 \leq a_m \leq \ell_m} \left[\frac{1}{a_m} + \left[\min_{\substack{a_1 \geq 0 \\ \vdots \\ a_{m-1} \geq 0}} \left(\frac{1}{a_1} + \dots + \frac{1}{a_{m-1}} \right) \right] \right] \quad (404)$$

subject to $a_1 + a_2 + \dots + a_{m-1} = \ell - a_m = \ell_{m-1}$. If we define

$$f_m(\ell_m) \triangleq \min_{0 \leq a_m \leq \ell_m} \left[\frac{1}{a_m} + f_{m-1} (\ell_m - a_m) \right] \quad (405)$$

we can write

$$f_1(\ell_1) = \min_{0 \le a_1 \le \ell_1} \frac{1}{a_1} = \frac{1}{\ell_1} \tag{406}$$

which implies that $a_1 = \ell_1$ and $\ell_1 = \ell_2 - a_2$. Also,

$$f_2(\ell_2) = \min_{0 \le a_2 \le \ell_2} \left[\frac{1}{a_2} + f_1(\ell_2 - a_2) \right] = \min_{0 \le a_2 \le \ell_2} \left[\frac{\ell_2}{\ell_2 a_2 - a_2^2} \right] . \tag{407}$$

To minimize (407) we require that

$$\frac{\partial f_2}{\partial a_2} = 0 = -\ell_2 \; (\ell_2 - 2a_2) \tag{408}$$

and

$$\frac{\partial^2 f_2}{\partial a_2^2} = 2\ell_2 > 0 \tag{409}$$

(sufficiency condition for minimum) which implies that $a_2 = \ell_2/2$. Hence, $f_2(\ell_2) = 4/\ell_2$ and $\ell_1 = \ell_2 - a_2 = \ell_2 - \ell_2/2 = \ell_2/2$ which implies that $a_1 = \ell_2/2 = a_2$. We will finish the proof by using mathematical introduction. Assume lemma 8 is true for $m-1$ (i.e., $a_1 = a_2 = \ldots = a_{m-1}$) and prove true for m (i.e., $a_1 = a_2 = \ldots = a_{m-1} = a_m$). Assuming lemma 8 is true for m-1 implies that

$$a_{m-1} = \frac{\ell_{m-1}}{m-1} = \frac{\ell_m - a_m}{m-1} \tag{410}$$

since $a_1 = a_2 = \ldots = a_{m-1}$ and $a_1 + a_2 + \ldots + a_{m-1} = \ell_{m-1} = \ell_m - a_m$.

We can write

$$\begin{aligned} f_m(\ell_m) &= \min_{0 \le a_m \le \ell_m} \left[\frac{1}{a_m} + f_{m-1}(\ell_m - a_m)\right] \\ &= \min_{0 \le a_m \le \ell_m} \left[\frac{\ell_m - m^2 a_m - 2m a_m}{a_m \ell_m - a_m^2}\right]. \end{aligned} \tag{411}$$

For minimization we require that

$$\frac{\partial f_m}{\partial a_m} = -\ell_m^2 + 2 a_m \ell_m + (m^2 - 2m)\, a_m^2 = 0 \tag{412}$$

and

$$\frac{\partial^2 f_m}{\partial a_m^2} = 2\ell_m + (m^2 - 2m)\, 2 a_m > 0 \text{ for } n \ge 2 \tag{413}$$

(sufficiency condition for minimum). Hence,

$$a_m = \frac{\ell_m[1 \pm (m-1)]}{-m(m-2)} . \tag{414}$$

For the plus sign we have $a_m = -\ell_m/(m-2) < 0$ which does not meet the requirement of $a_m \geq 0$. For the minus sign we have $a_m = \ell_m/m > 0$. We can then write

$$a_1 + a_2 + \ldots + a_{m-1} = (m-1)a_{m-1} = \ell_m - a_m = ma_m - a_m = (m-1)a_m \tag{415}$$

which implies that $a_{m-1} = a_m$. Hence,

$$a_1 = a_2 = \ldots = a_{m-1} = a_m. \tag{416}$$

From the above lemmas we can see that in order to minimize the trace of $B^{-1} = (A'A)^{-1}$ we must make the off-diagonal elements equal to zero. This implies that we make all our measurements at peaks of point spread functions. In doing this for A'A the diagonal elements become larger as the off-diagonal elements go to zero; hence, the trace is further reduced. From the last lemma, we see that each of the diagonal elements of A'A must be equal which implies that the same number of measurements be made at the peak of each point spread function.

LITERATURE CITED

Born, Max, and Emil Wolf. 1964. Principles of Optics. 2d (revised) ed. The Macmillian Company, New York, N. Y. 808 p.

Bussgang, J. J., and D. Middleton. 1955. Optimum Sequential Detection of Signals in Noise. IRE Transactions on Information Theory 1:5-18.

Deutsch, Ralph. 1965. Estimation Theory. Prentice-Hall, Inc., Englewood Cliffs, N. J. 269 p.

Farrell, Edward J. 1966. Information Content of Photoelectric Star Images. Journal of the Optical Society of America 56: 578-587.

Goodman, J. W. 1965. Some Effects of Target-Induced Scintillation on Optical Radar Performance. Proceedings of the IEEE 53: 1688-1700.

Harris, J. L. 1964. Diffraction and Resolving Power. Journal of the Optical Society of America 54: 931-936.

Harris, Richard W., and C. K. Rushforth. 1966. Numerical Restoration of Optical Objects Obscured by Diffraction and Noise. Scientific Report No. 8, Contract No. AF 19(628)-3825. Electro-Dynamics Laboratories, Utah State University, Logan, Utah.

Helstrom, Carl W. 1964. The Detection and Resolution of Optical Signals. IEEE Transactions on Information Theory 10: 275-287.

Kailath, T. 1963. Adaptive Matched Filters. p. 109-140. In R. Bellman, ed, Mathematical Optimization Techniques, 1963. University of California Press, Berkeley, California.

Olmsted, John M. H. 1956. Advanced Calculus. Appleton-Century-Crofts, Inc., New York, N. Y. 706 p.

O'Neill Edward L. 1963. Introduction to Statistical Optics. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts. 179 p.

Parzen, Emanuel. 1962. Stochastic Processes. Holden-Day, Inc., San Francisco, California. 324 p.

Rushforth, Craig K. 1965. Restoration of Optical Patterns. Scientific Report No. 3, Contract No. AF 19(628)-3825. Electro-Dynamics Laboratories, Utah State University, Logan, Utah.

Stone, John M. 1963. Radiation and Optics. McGraw-Hill Book Company, Inc., New York, N. Y. 544 p.

Wald, Abraham. 1947. Sequential Analysis. John Wiley and Sons, Inc., New York, N. Y. 212 p.

Wylie, C. R., Jr. 1960. Advanced Engineering Mathematics. 2d ed. McGraw-Hill Book Company, Inc., New York, N. Y. 696 p.

Zadeh, Lotfi A., and Charles A. Desoer. 1963. Linear System Theory. McGraw-Hill Book Company, Inc., New York, N. Y. 628 p.

ACKNOWLEDGEMENTS

I wish to express sincere appreciation to Dr. Craig K. Rushforth for his help and guidance in this study. I also appreciate helpful discussions with Mr. Richard W. Harris. The use of the computer facilities of the Western Data Processing Center in Los Angeles is gratefully acknowledged. This study was supported in part by the Air Force Cambridge Research Laboratories, Bedford, Massachusetts, under Contract No. AF19(628)-3825. Part of this study was supported by a three-year traineeship from the National Aeronautics and Space Administration.

Mark C. Austin

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Electro-Dynamics Laboratories Utah State University, Logan, Utah	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

ESTIMATION AND DETECTION OF OPTICAL SIGNALS DISTORTED BY DIFFRACTION, BACKGROUND NOISE AND DETECTION NOISE

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Scientific Report, Interim

5. AUTHOR(S) *(Last name, first name, initial)*

Austin, Mark C.
Rushforth, Craig K.

6. REPORT DATE 31 December 1966	7a. TOTAL NO. OF PAGES 229	7b. NO. OF REFS 17
8a. CONTRACT OR GRANT NO. AF19(628)-3825 b. PROJECT AND TASK NO. 8663 — ARPA Order No. 450 c. DOD ELEMENT 62503015 d. DOD SUBELEMENT N/A	9a. ORIGINATOR'S REPORT NUMBER(S) Scientific Report No. 9 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* AFCRL-67-0071	

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES Prepared for Hq. AFCRL, OAR (CRO) United States Air Force L. G. Hanscom Fld, Bedford, Mass	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency

13. ABSTRACT

Estimation and detection of optical signals distorted by diffraction, additive background noise, and multiplicative (detection) noise are studied. Assuming that the output of the detector is a Poisson process, that the signal and noise are additive, and that they have prescribed means and covariance matrices, the optimum linear estimate of the optical signal or object is obtained. In the physical detection process, the interaction between the incident radiation and the detector produces an effect called multiplicative noise which must be taken into account in obtaining the optimum linear estimate. The performance of the estimation procedure is evaluated for several special cases. Both white and colored noise are considered in the estimation problem. The problem of discriminating between optical signals is considered. Optimum procedures are derived for detecting known and unknown optical using fixed-sample detectors. The properties of sequential detectors which are optimum for the detection of random or unknown optical signals are investigated. A comparison is made of the average test lengths of these optimum random signal detectors with those of a detector designed for particular optical signals. The test lengths of the fixed-sample detector and sequential detector are compared for a particular example.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Image restoration Pattern recognition Sequential detection Fixed-sample detection Fourier optics Adaptive data processing						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.