NRL Report 8443

Some Error Probabilities for the Association of Passive DF Measurements with Radar Returns

J. O. Coleman

Radar Analysis Branch
Radar Division

October 21, 1980

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Report 8443	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SOME ERROR PROBABILITIES FOR THE ASSOCIATION OF PASSIVE DF MEASUREMENTS WITH RADAR RETURNS		5. TYPE OF REPORT & PERIOD COVERED Interim report on a continuing NRL problem. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. O. Coleman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, DC 20375		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62712N; ZF12151001 NRL Problem 53-0612-0-0
11. CONTROLLING OFFICE NAME AND ADDRESS Department of the Navy Chief of Naval Material Washington, DC 20360		12. REPORT DATE October 21, 1980 13. NUMBER OF PAGES 57
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sensor integration
Multisensor correlation system
Classification error rate
Sensor correlation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An important problem in multiple-sensor integration is the correct association of data from different sensors with the same target. This report considers a family of four discriminants for associating a set of measurements from a passive DF system with one of the radar targets being tracked by a surveillance radar system. This association involves choosing the radar target that will minimize a discriminant which is a function of the DF data and data on that radar target. The discriminants considered here are sums of squared differences of independent unbiased Gaussian

(Continues)

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (Continued)

measurements of bearings to stationary targets. Because data generally arrive from the different sensors asynchronously, it is becoming common practice with this type of system to smooth (average together) the radar data, the DF data, or both prior to use in the discriminant. This report attempts to provide a basis for deciding which smoothing option is most appropriate by giving the error probabilities for each of these three smoothing options and for the option of no smoothing.

CONTENTS

SOME ERROR PROBABILITIES FOR THE ASSOCIATION OF PASSIVE DF MEASUREMENTS WITH RADAR RETURNS

INTRODUCTION

A great deal of attention is currently focused on multiple-sensor surveillance systems. One important aspect of such systems is the correct association of data from different sensors with the same target. Of particular interest here is the problem of associating a set of measurements from a passive direction-finding (DF) system with the correct target in a surveillance radar system. This can be viewed as a problem in statistical pattern recognition in which a set of DF measurements is to be assigned to a class whose characteristics are determined from measurements of a radar target. This classification is performed by choosing the radar target that will minimize a discriminant which is a function of the DF data and the data on the radar target.

One difficulty inherent in this problem is due to the asynchronous operation of the radar and DF systems. Because the targets may be in motion, it is necessary to smooth or extrapolate in time the DF measurements, the radar data, or both in order that information from both sensors refer to the same set of target positions. It is not immediately obvious, however, which set of data should be smoothed. Gerlach [1] assumed simultaneous measurements and used no smoothing. Coleman [2] smoothed the radar data only. Bath [3] smoothed both the radar and DF data to some extent. No claim was made by any of these authors that his choice was optimum, and in fact the question of which data to smooth was not explicitly addressed. The purpose of this report is to bound the performance obtainable under each of the four possible smoothing options listed in Table 1. In the interest of conciseness the mnemonic abbreviations shown in Table 1 for these options will be used freely. Using some simplifying and unifying assumptions outlined subsequently, exact expressions for the probability of classification error are obtained for three of the options: RDS, RS, and NS. The probability of error with DS is evaluated with an importance-sampling simulation. Due to the nature of the assumptions, these error probabil-

Table 1 — Mnemonic Abbreviations for the Smoothing Options

Mnemonic	Smoothing Option
RDS	Radar and DF smoothing
RS	Radar smoothing only
DS	DF smoothing only
NS	No smoothing

Manuscript submitted July 15, 1980.

ities represent lower bounds on realistically obtainable performance. These results for the RDS and RS options have been described previously [2] but are reviewed here in somewhat more detail. The results for the remaining options are new.

ASSUMPTIONS

A number of assumptions about the problem are made. Those based on earlier work [2] include the following:

- Starting with a set of passive measurements and asking which set of radar detections it should be paired with. This assumption assures that each DF target is associated with only one radar target (with the implicit requirement that radar data on the appropriate target be available). Several DF targets are allowed to be paired with a single radar target; for example, paired with an aircraft target may be separate sets of DF measurements to each of several transmitters on board.

- Associating an entire passive measurement set as a unit (rather than measurement by measurement). This assumption implies an ability to correctly associate a new DF measurement with existing DF measurements of the same target. It should allow more accurate DF-to-radar association than if the association proceeded on a single DF measurement.

- Basing the association decision on minimizing a sum of squared passive-to-radar bearing differences. This assumption is based on a generalized likelihood classifier for Gaussian random variables [4] and has been implicit in all of the radar/DF association schemes known to this author.

Further assumptions are made to simplify the problem. There are two targets, and both are stationary. This allows classification without smoothing to be a viable option. Target 1 is under both radar and DF observation, while target 2 is observed only by radar. Due to the symmetry in the problem this assumption can be made without loss of generality. Radar and DF equipments are colocated. Therefore the only radar measurement of interest is azimuth. The targets are separated in azimuth by μ_a (defined as the azimuth of target 2 minus the azimuth of target 1). The set of DF measurements and the two sets of radar measurements contain n measurements each. All measurements are unbiased and independent. The measurement errors are Gaussian, with the DF errors having variance σ^2_{DF} and the radar errors having variance σ^2_R. Since there are only two targets, the decision procedure is to evaluate the discriminant function for each, form the difference, and compare to zero. Specifically, the procedure is to compute

$$d = \left[\frac{1}{n} \sum_{i=1}^{n} \left(\hat{\theta}_{\mathrm{DF}i} - \hat{\theta}_{2i} \right)^2 \right] - \left[\frac{1}{n} \sum_{i=1}^{n} \left(\hat{\theta}_{\mathrm{DF}i} - \hat{\theta}_{1i} \right)^2 \right] \tag{1}$$

and associate the DF measurements with target 1 if $d > 0$ (correct decision) and with target 2 if $d \leq 0$ (incorrect decision). In this equation $\hat{\theta}_{\mathrm{DF}i}$ is the ith DF-based azimuth estimate. If DF smoothing is not used (RS and NS), $\hat{\theta}_{\mathrm{DF}i}$ is just $\theta_{\mathrm{DF}i}$, the ith DF measurement. If DF smoothing is used (RDS and DS),

$$\hat{\theta}_{\mathrm{DF}i} = \hat{\theta}_{\mathrm{DF}} = \frac{1}{n} \sum_{k=1}^{n} \theta_{\mathrm{DF}k},$$

the average of the DF measurements. Similarly $\hat{\theta}_{1i}$ is the ith radar-based estimate of the azimuth of target 1. With no radar smoothing $\hat{\theta}_{1i}$ is θ_{1i}, the ith radar azimuth measurement of target 1. With radar smoothing

$$\hat{\theta}_{1i} = \hat{\theta}_{1} = \frac{1}{n} \sum_{k=1}^{n} \theta_{1k},$$

the average of the radar azimuth measurements of target 1. Similar relationships hold between $\hat{\theta}_{2i}$, $\hat{\theta}_2$, θ_{2i} and the smoothing options.

Because the sign of d in Eq. (1) is unaffected by a scale change in azimuth, the means and standard deviations of the densities are normalized by σ_{DF} throughout the subsequent development except where noted. The notation

$$\mu = \mu_a/\sigma_{\mathrm{DF}}$$

and

$$\sigma = \sigma_R/\sigma_{\mathrm{DF}}$$

will be used.

In the next four sections the probability of classification error P_e is evaluated for the four smoothing options under the assumptions just outlined.

P_e WITH RADAR AND DF SMOOTHING (RDS)

When both the radar data and the DF data are smoothed, the estimates $\hat{\theta}_{\mathrm{DF}i}$, $\hat{\theta}_{1i}$, and $\hat{\theta}_{2i}$ are all independent of i. Equation (1) can therefore be rewritten as

$$\begin{aligned} d &= (\hat{\theta}_{\mathrm{DF}} - \hat{\theta}_2)^2 - (\hat{\theta}_{\mathrm{DF}} - \hat{\theta}_1)^2 \\ &= 2\,\hat{\theta}_1\hat{\theta}_{\mathrm{DF}} - 2\,\hat{\theta}_2\hat{\theta}_{\mathrm{DF}} + \hat{\theta}_2^2 - \hat{\theta}_1^2 \qquad (2) \\ &= (\hat{\theta}_2 - \hat{\theta}_1)\,[(\hat{\theta}_2 + \hat{\theta}_1) - 2\,\hat{\theta}_{\mathrm{DF}}]. \end{aligned}$$

Both of the factors in this expression are linear combinations of Gaussian random variables and are therefore Gaussian. The mean of the first factor is

$$E\{\hat{\theta}_2 - \hat{\theta}_1\} = E\left\{\frac{1}{n}\sum_{i=1}^{n} (\theta_{2i} - \theta_{1i})\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\{\theta_{2i} - \theta_{1i}\}$$

$$= \mu.$$

Similarly

$$E\{\hat{\theta}_2 - \hat{\theta}_{\mathrm{DF}}\} = \mu$$

and

$$E\{\hat{\theta}_1 - \hat{\theta}_{\mathrm{DF}}\} = 0.$$

These last two equations imply

$$E\left\{(\hat{\theta}_2 + \hat{\theta}_1) - 2\,\hat{\theta}_{\mathrm{DF}}\right\} = \mu.$$

The variances are just as simply calculated:

$$\mathrm{var}\{\hat{\theta}_2 - \hat{\theta}_1\} = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{var}\{\theta_{2i} - \theta_{1i}\}$$

$$= \frac{2\sigma^2}{n}$$

and

$$\mathrm{var}\{(\hat{\theta}_2 + \hat{\theta}_1) - 2\hat{\theta}_{\mathrm{DF}}\} = \frac{2\sigma^2}{n} + 4\,\mathrm{var}\{\hat{\theta}_{\mathrm{DF}}\}$$

$$= \frac{2\sigma^2}{n} + \frac{4}{n}.$$

From independence of the θ_{DFi}, θ_{1i}, and θ_{2i} it follows that the $(\theta_{2i} - \theta_{1i})$ and $(\theta_{2i} + \theta_{1i})$ are independent. Consequently, the two factors in Eq. (2) are independent, and the product d is less than zero if and only if its two independent Gaussian factors are of opposite sign. The probability of error can therefore be expressed as

$$P_e = \text{Prob}\,(d \leqslant 0) = \text{Prob}\,(\hat{\theta}_2 - \hat{\theta}_1 \leqslant 0)\ \text{Prob}\left[(\hat{\theta}_2 + \hat{\theta}_1) - 2\theta_{DF} \geqslant 0\right]$$

$$+ \text{Prob}\,(\hat{\theta}_2 - \hat{\theta}_1 \geqslant 0)\ \text{Prob}\left[(\hat{\theta}_2 + \hat{\theta}_1) - 2\theta_{DF} \leqslant 0\right].$$

Or, if $\Phi(x)$ is defined as the cumulative distribution function for a zero-mean unit-variance Gaussian random variable,

$$P_e = \Phi\left(\frac{-\mu}{\sqrt{2\sigma^2/n}}\right)\left[1 - \Phi\left(\frac{-\mu}{\sqrt{\frac{2}{n}(2+\sigma^2)}}\right)\right]$$

$$+ \left[1 - \Phi\left(\frac{-\mu}{\sqrt{2\sigma^2/n}}\right)\right]\Phi\left(\frac{-\mu}{\sqrt{\frac{2}{n}(2+\sigma^2)}}\right).$$

This result can be more concisely expressed by using (from Ref. 5)

$$\Phi(x) = \frac{1}{2}\,\text{erfc}\left(\frac{-x}{\sqrt{2}}\right),$$

where

$$\text{erfc}\,(x) = 1 - \text{erf}\,(x)\,,$$

to obtain

$$P_e = \frac{1 - \text{erf}\left(\dfrac{\mu\sqrt{n}}{2\sigma}\right)\text{erf}\left(\dfrac{\mu\sqrt{n}}{2\sqrt{2+\sigma^2}}\right)}{2}\,. \tag{3}$$

This function is plotted in Fig. 1 versus $\mu\sqrt{n}$ with σ as a parameter.

Several observations can be made regarding this result. First, the curve for $\sigma = 0$ reflects performance with perfect radar measurements and, in effect, is obtained using true positions instead of estimates for the $\hat{\theta}_{1i}$ and $\hat{\theta}_{2i}$ in Eq. (1). Therefore this curve gives the performance of an idealized maximum-liklihood detector. As σ increases, the probability of error increases slowly until σ approaches unity. This suggests that the accuracy of the radar (σ_R^2) is not important as long as it is significantly better than the accuracy of the DF equipment (σ_{DF}^2).

Further insight can be gained into the behavior of this function by using a different normalization. After substituting the definitions

$$\mu = \frac{\mu_a}{\sigma_{DF}}$$

and

$$\sigma = \frac{\sigma_R}{\sigma_{DF}}$$

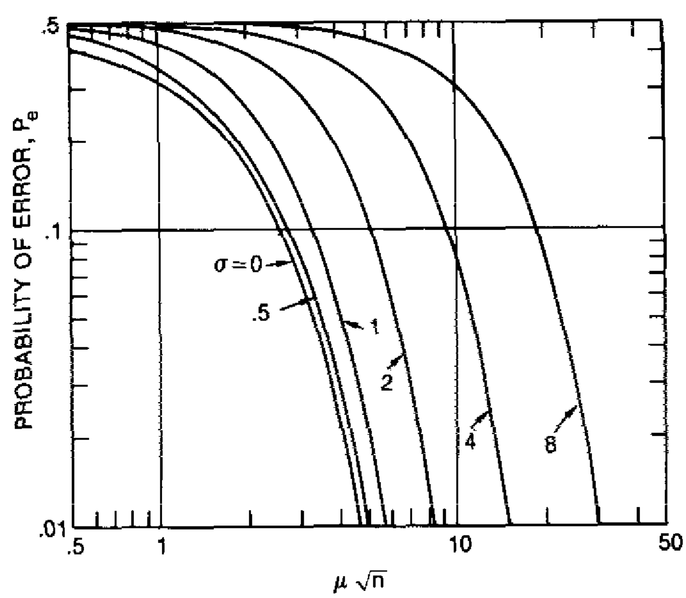

Fig. 1 — Performance with radar and DF smoothing (RDS)

into Eq. (3), the probability of error becomes

$$P_e = \frac{1 - \operatorname{erf}\left(\dfrac{1/2}{\sigma_R/(|\mu_a|\sqrt{n})}\right)\operatorname{erf}\left(\dfrac{1/2}{\sqrt{2\left(\dfrac{\sigma_{DF}}{\mu_a\sqrt{n}}\right)^2 + \left(\dfrac{\sigma_R}{\mu_a\sqrt{n}}\right)^2}}\right)}{2} \tag{4}$$

after some manipulation*. Figure 2 shows contours of constant $\log_{10} P_e$ plotted versus $\log_{10} |\sigma_{DF}/(\mu_a\sqrt{n})|$ and $\log_{10} |\sigma_R/(\mu_a\sqrt{n})|$.† It can be thought of as a log-log-log plot of P_e versus the two standard deviations σ_{DF} and σ_R witn $|\mu_a|\sqrt{n}$ fixed to unity. It is apparent from this plot that not only is σ_R unimportant when $\sigma_R \ll \sigma_{DF}$ but σ_{DF} is unimportant when $\sigma_{DF} \ll \sigma_R$.

The question of which variance is more important can be considered more rigorously using sensitivities. The sensitivity of the function P_e with respect to σ_R is defined as

$$S_{\sigma_R}^{P_e} = \frac{\sigma_R}{P_e}\frac{\partial P_e}{\partial \sigma_R} .$$

Similarly

$$S_{\sigma_{DF}}^{P_e} = \frac{\sigma_{DF}}{P_e}\frac{\partial P_e}{\partial \sigma_{DF}} .$$

*As (from Ref. 5) erf $(-x) = -$erf (x), Eq. (3) shows P_e to be an even function of μ. In the manipulation to arrive at Eq. (4) the sign of μ is lost as it is brought under the radical and the symmetry of P_e with respect to μ is destroyed. This situation was corrected by replacing μ_a by $|\mu_a|$ in Eq. (4).

†Figure 2 was produced with the aid of MACSYMA, a large symbolic manipulation program developed at the MIT Laboratory for Computer Science and supported by the National Aeronautics and Space Administration under grant MSG 1323, by the Office of Naval Research under grant N00014-77-C-0841, by the U.S. Department of Energy under grant ET-78-C-02-4687, and by the U.S. Air Force under grant F49620-79-C-020.

It is not difficult to show that these can be defined equivalently as

$$S_{\sigma_R}^{P_e} = \frac{\partial(\log_{10} P_e)}{\partial(\log_{10} \sigma_R)}$$

$$S_{\sigma_{DF}}^{P_e} = \frac{\partial(\log_{10} P_e)}{\partial(\log_{10} \sigma_{DF}} \, .$$

These two quantities are equal in Fig. 2 wherever the contours are at 45° with respect to the axis. The locus of such points is shown in Fig. 2 and represents the boundary between the region where a small percentage change in σ_R is more significant than the same percentage change in σ_{DF} and the region where the opposite is true.

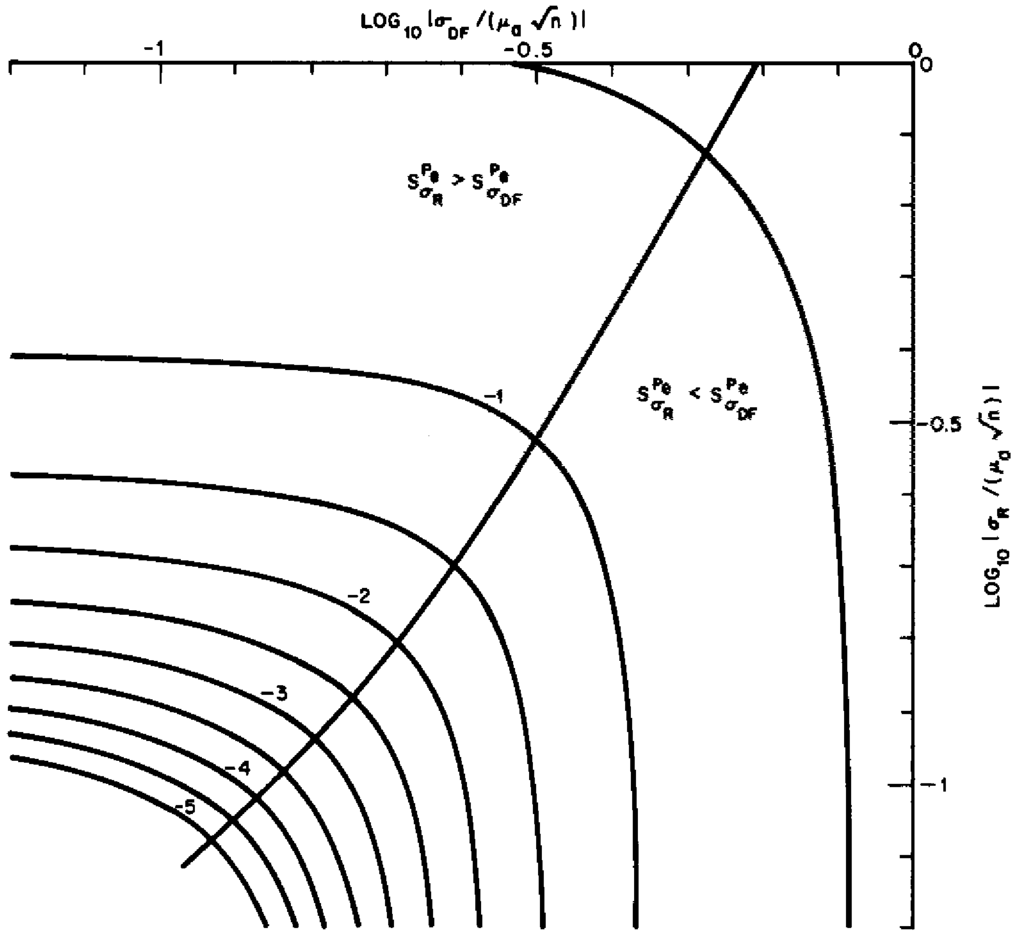

Fig. 2 — RDS contours of constant $\log_{10}(P_e)$

P_e WITH RADAR SMOOTHING ONLY (RS)

When only the radar data are smoothed, Eq. (1) becomes

$$d = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\theta_{\mathrm{DF}i} - \hat{\theta}_2\right)^2\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\left(\theta_{\mathrm{DF}i} - \hat{\theta}_1\right)^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(2\hat{\theta}_1\theta_{\mathrm{DF}i} - 2\hat{\theta}_2\theta_{\mathrm{DF}i} + \hat{\theta}_2^2 - \hat{\theta}_1^2\right)$$

$$= (\hat{\theta}_2 - \hat{\theta}_1)\left[(\hat{\theta}_2 + \hat{\theta}_1) - \frac{2}{n}\sum_{i=1}^{n}\theta_{\mathrm{DF}i}\right]$$

$$= (\hat{\theta}_2 - \hat{\theta}_1)\left[(\hat{\theta}_2 + \hat{\theta}_1) - 2\hat{\theta}_{\mathrm{DF}}\right]. \tag{5}$$

This discriminant is identical to Eq. (2). The performace is therefore exactly the same as when both radar and DF data are smoothed. This equivalence is shown under a rather restrictive set of assumptions and is not expected to hold in a more general case such as when the target is in motion.

P_e WITH DF SMOOTHING ONLY (DS)

With DF smoothing only, Eq. (1) becomes

$$d = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_{\mathrm{DF}} - \theta_{2i}\right)^2\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_{\mathrm{DF}} - \theta_{1i}\right)^2\right]. \tag{6}$$

After an unsuccessful attempt to determine analytically the probability that this quantity is less than zero, a Monte Carlo simulation was conducted using importance sampling to reduce the number of trials required. (Importance sampling is briefly described in Appendix A.) To make effective use of the importance-sampling method, it seemed best to use one approach for $\sigma < 1$ and another for $\sigma \geqslant 1$. For $\sigma < 1$ the parameter modified by importance sampling was σ_{DF}. This led to an estimate of probability of error of the form

$$\hat{P}_e = \frac{1}{N_T} \sum_{k=1}^{N_T} \frac{\sigma'_{\mathrm{DF}}}{\sigma_{\mathrm{DF}}} \exp\left[\frac{-\hat{\theta}^2_{\mathrm{DF}k}}{2} \left(\frac{1}{\sigma^2_{\mathrm{DF}}} - \frac{1}{\sigma'^2_{\mathrm{DF}}} \right) \right] U(-d') , \tag{7}$$

where $\hat{\theta}_{\mathrm{DF}k}$ is the sample value of $\hat{\theta}_{\mathrm{DF}}$ on the kth trial, N_T is the number of trials, σ'_{DF} is the modified σ_{DF} (the value actually used in the simulation), and the function $U(x)$ is defined by

$$U(x) = 1, \text{ if } x \geqslant 0,$$

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

The variable d' is the computed discriminant with $\hat{\theta}_{\mathrm{DF}}$ drawn from the modified density. The value of σ_R was set to unity with no loss of generality other than the $\sigma_R = 0$ case treated separately below. Based on only the assumption that P_e curves with DF smoothing would not be drastically different from Fig. 1 (which turned out to be correct), σ'_{DF} was chosen to bring $(\mu_a/\sigma'_{\mathrm{DF}})\sqrt{n}$ to a value (constant for each value of σ_{DF} and chosen by hand) such that Fig. 1 indicated a P_e in the range 0.15 to 0.25. The results of this simulation for $N_T = 10{,}000$ are shown in Figs. 3a, 3b, and 3c. Figures 3d, 3e and 3f are the results of a different simulation; the reasons for and description of this simulation are as follows.

As σ is increased to values above unity, it becomes impractical to obtain an increase in simulated errors by varying σ_{DF}. To see this, first note that $\sigma < 1$ implies $\sigma_{\mathrm{DF}} > \sigma_R$ as sketched in Fig. 4a. Most errors occur when $\hat{\theta}_{\mathrm{DF}}$ falls above θ_{2i}. Increasing σ_{DF} by a modest amount will increase the probability of this type of error significantly. Figure 4b shows the type of situation implied by $\sigma > 1$ or $\sigma_R > \sigma_{\mathrm{DF}}$. Here errors are usually due to θ_{1i} falling above θ_{2i}. Increasing σ_{DF} by a modest factor will have very little effect on the probability of error. Increasing σ_{DF} enough to significantly affect P_e would in fact lead to the nonzero terms in Eq. (7) (ratio of the true to the modified probability density function for $\hat{\theta}_{\mathrm{DF}}$) being usually very small but occasionally very large. This implies an undesirably large variance for $\hat{P}_e$. (This was in fact demonstrated.)

For the special case in which $\sigma = \sigma_R = 0$ the simulation is not necessary, because RDS and DS become identical. (It is immaterial whether or not radar measurements are smoothed when they are exact.) From examination of Eq. (3) in the limit as σ goes to 0

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\frac{\mu\sqrt{n}}{2\sqrt{2}} \right)$$

or

$$P_e = \Phi\left(\frac{-\mu\sqrt{n}}{2} \right).$$

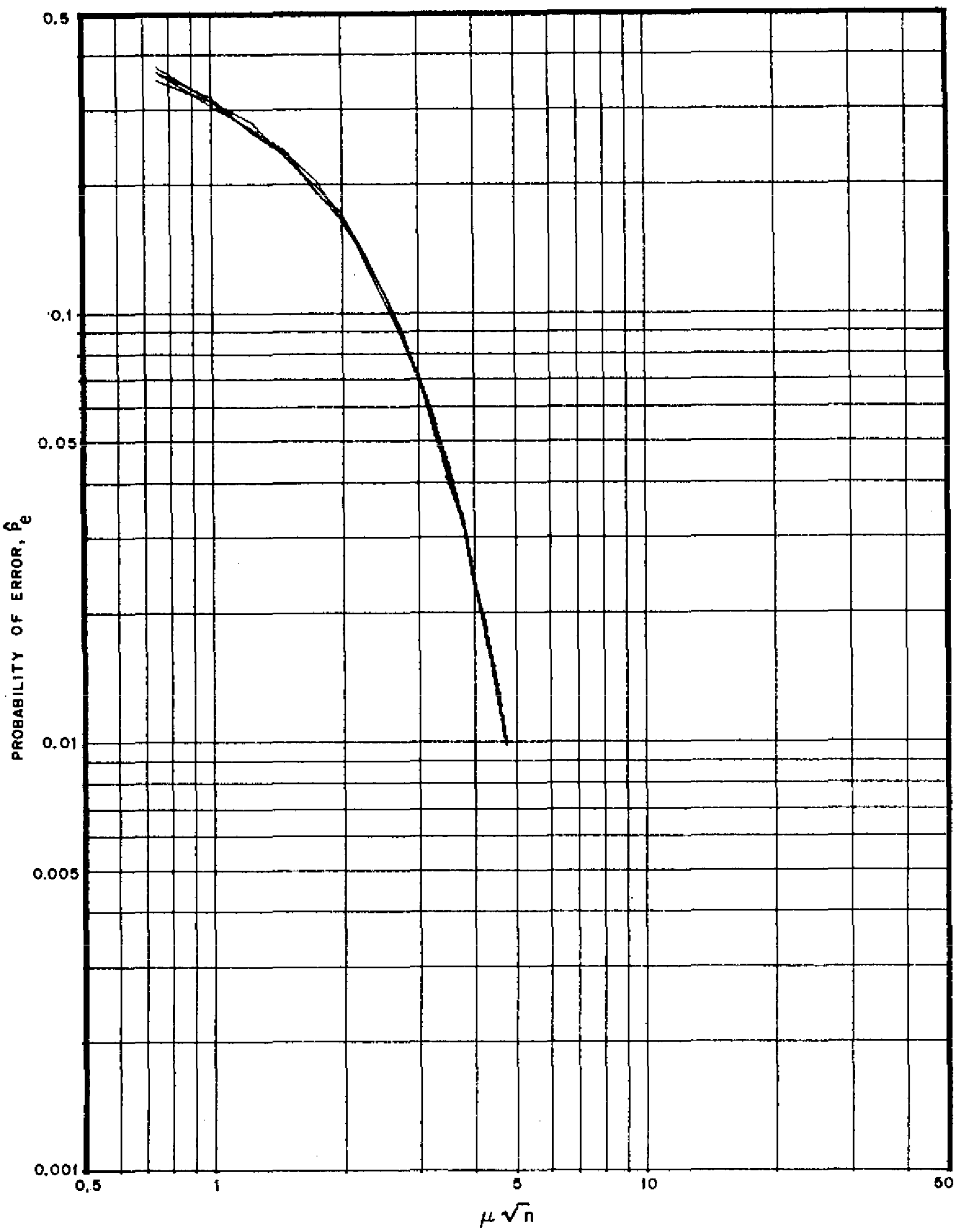

Fig. 3a — Performance with DF smoothing only (DS): 10,000 trials with $(\mu_a/\sigma'_{DF})\sqrt{n}$ set to 2 and $\sigma = 0.25$, with the four curves being for n = 1, 2, 4, and 8

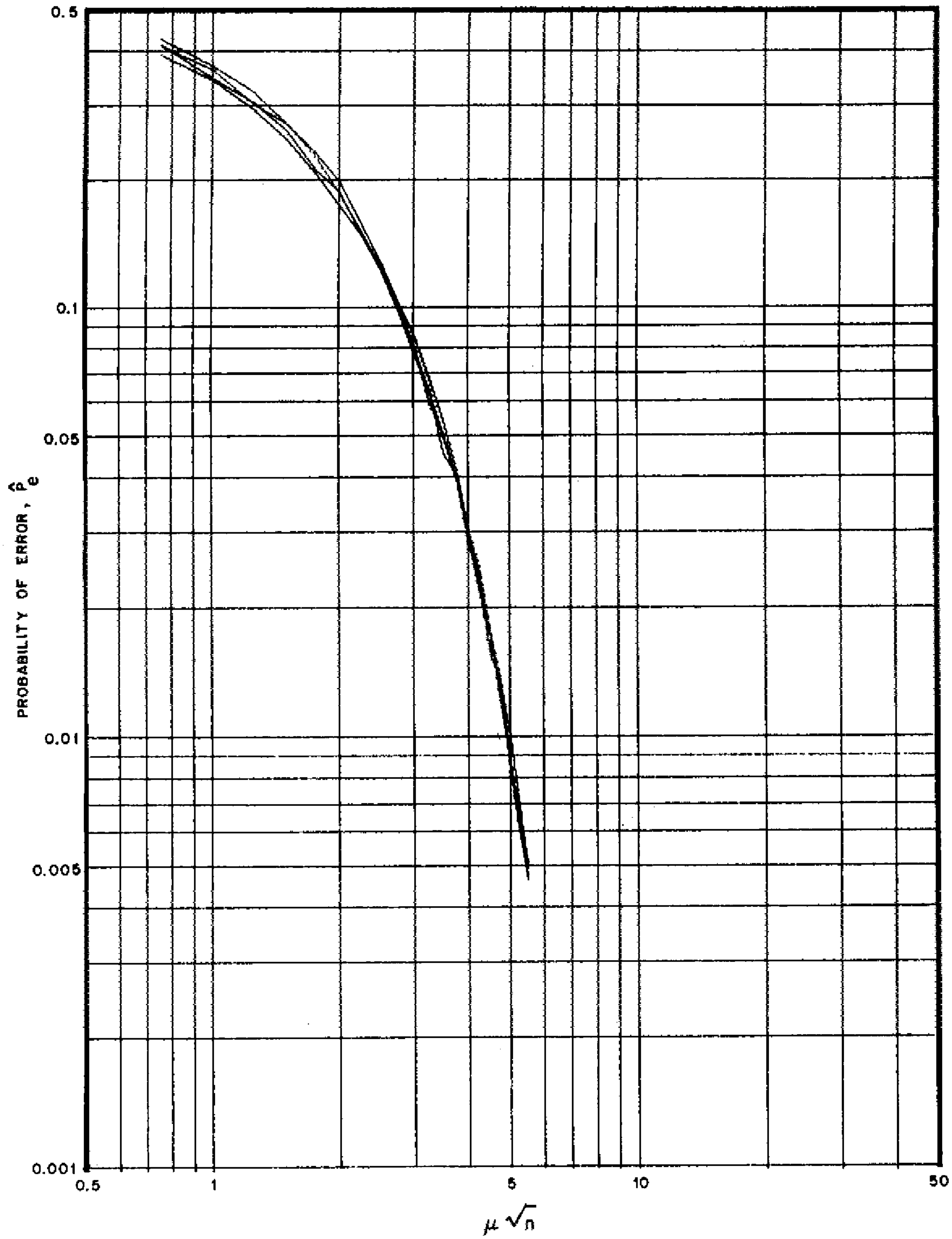

Fig. 3b — Performance with DF smoothing only: 10,000 trials with $(\mu_\sigma/\sigma'_{DF})/\sqrt{n}$ set to 2 and $\sigma = 0.5$, with $n = 1, 2, 4$, and 8

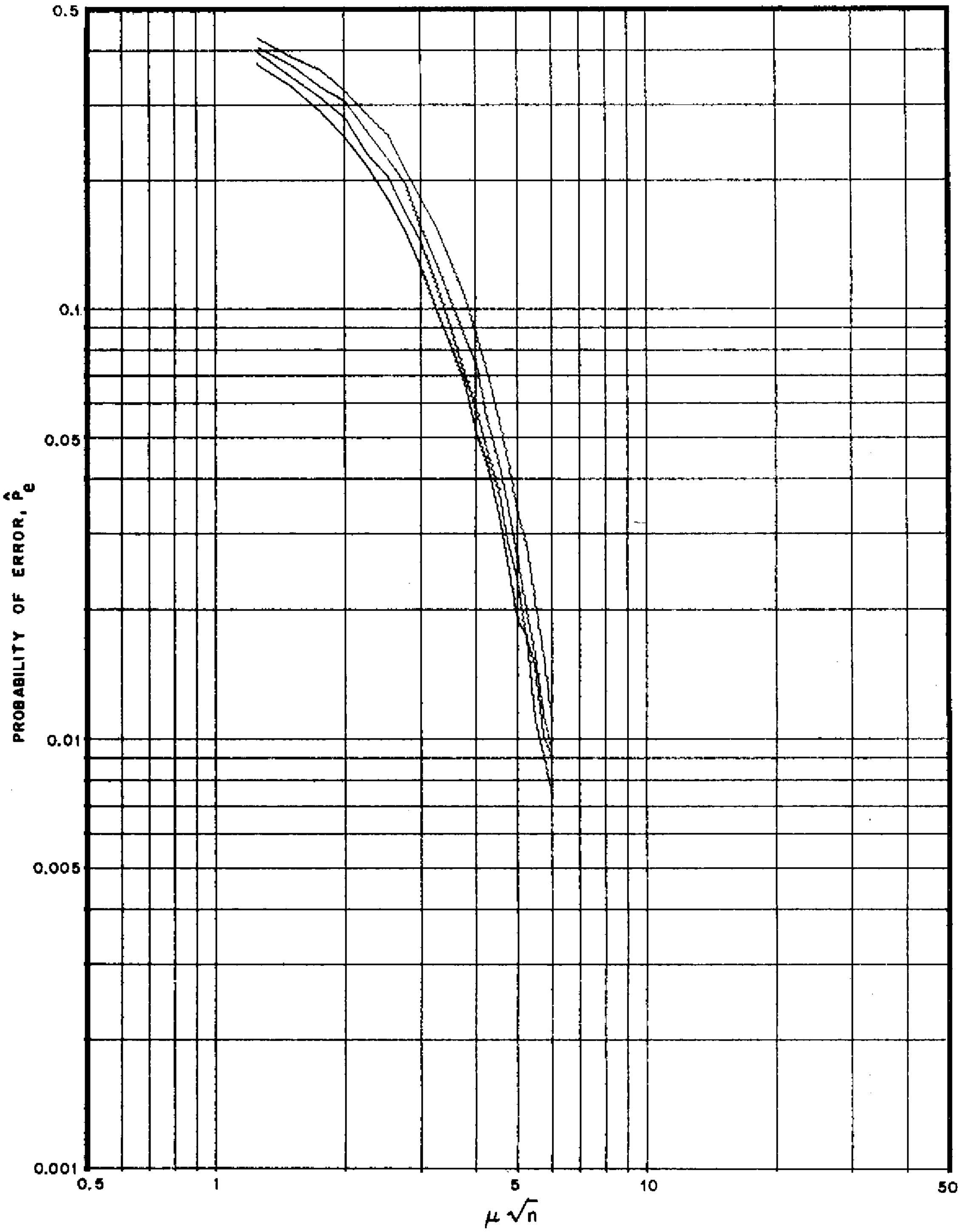

Fig. 3c — Performance with DF smoothing only: 10,000 trials with $(\mu_a/\sigma'_{DF})\sqrt{n}$ set to 2.75 and $\sigma = 1$, with $n = 1, 2, 4$, and 8

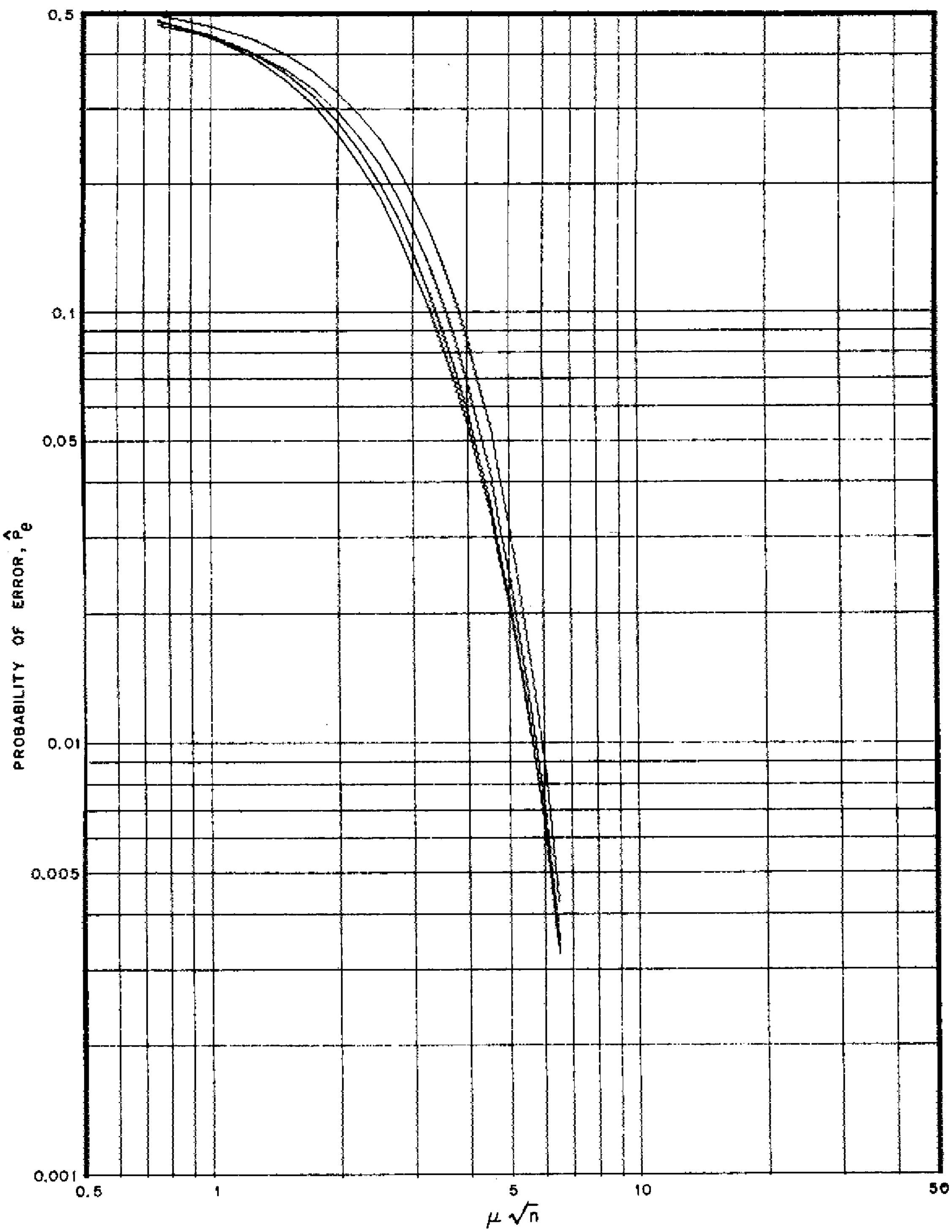

Fig. 3d — Performance with DF smoothing only: 10,000 trials with X' set to 2.75 and $\sigma = 1$, with $n = 1, 2, 4$, and 8

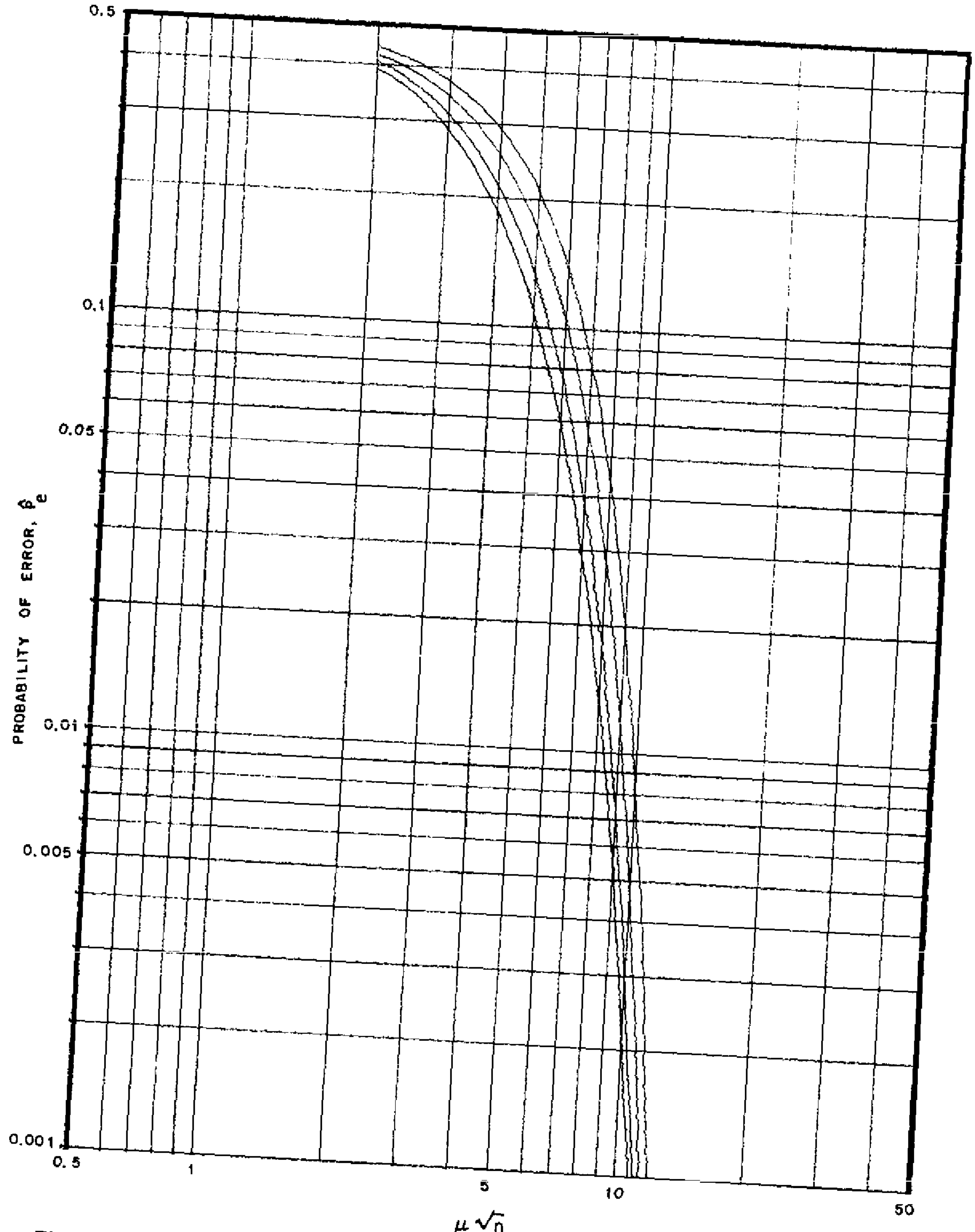

Fig. 3e — Performance with DF smoothing only: 10,000 trials with X' set to 3.8 and $\sigma = 2$, with $n = 1, 2, 4$, and 8

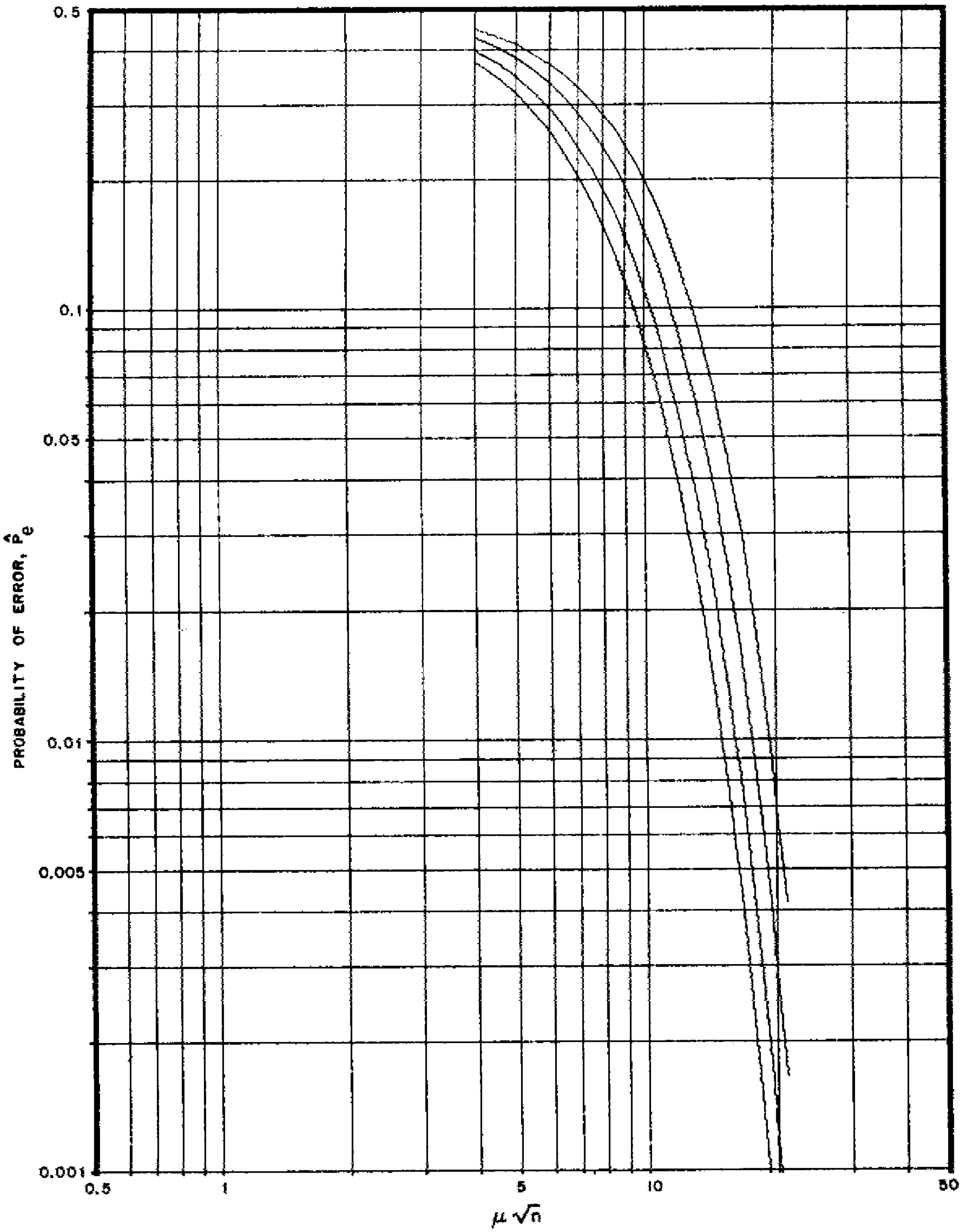

Fig 3f — Performance with DF smoothing only: 10,000 trials with X' set to 7 and $\sigma = 4$, with $n = 1, 2, 4$, and 8

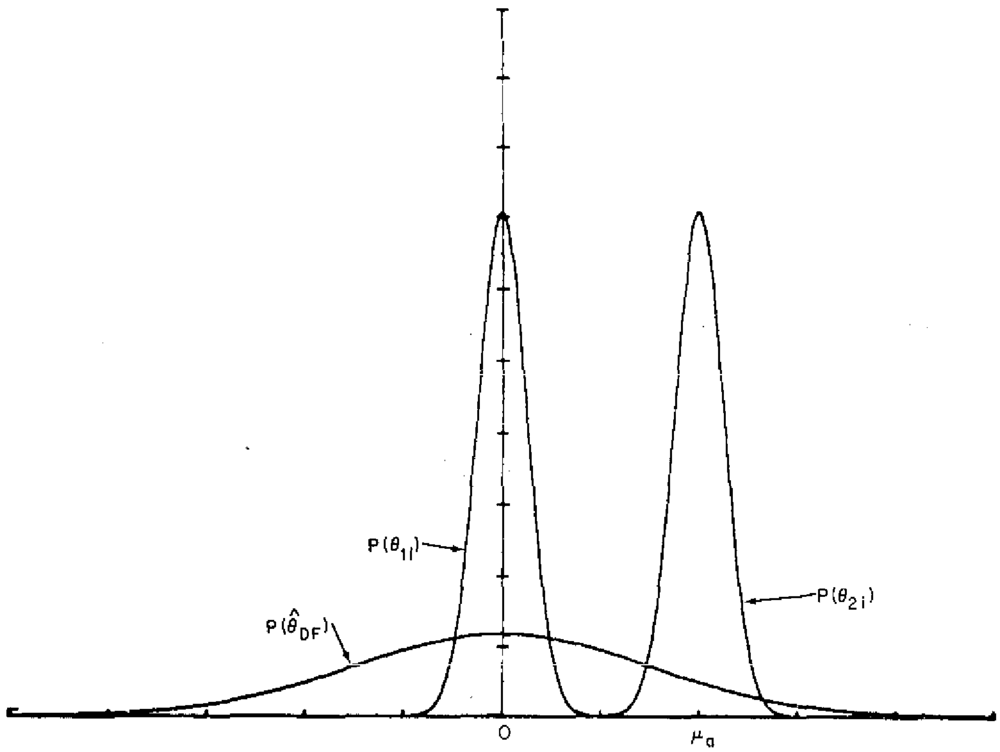

Fig. 4a — Situation in which $\sigma_{\mathrm{DF}} \gg \sigma_R$, causing errors when $\theta_{\mathrm{DF}} > \theta_{2i}$

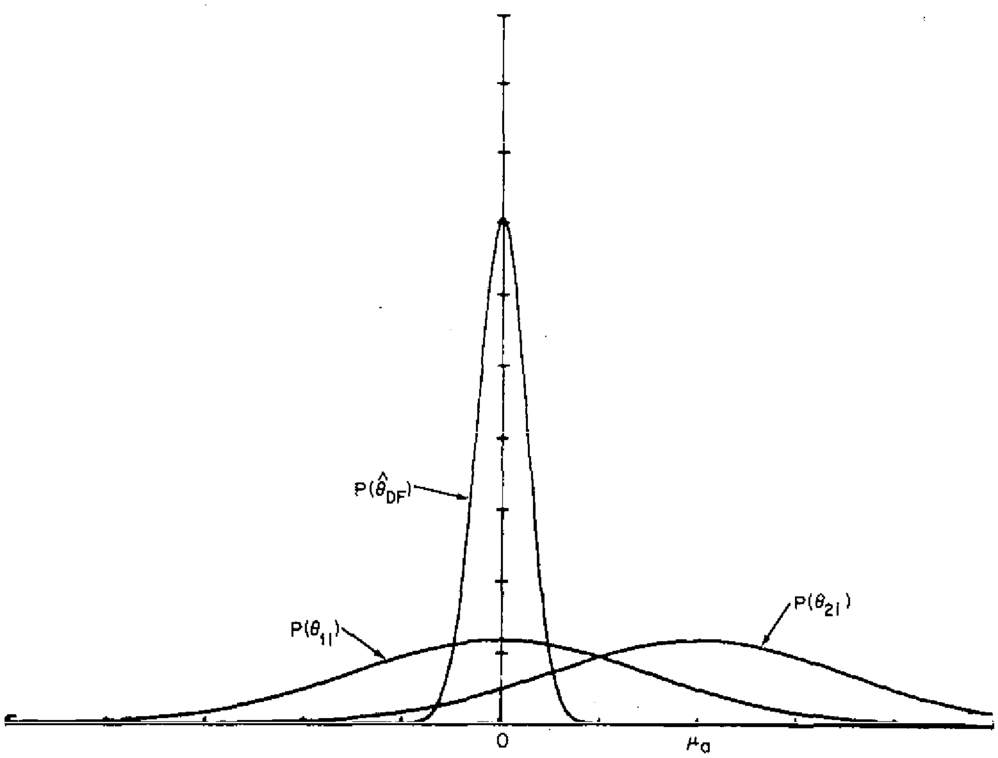

Fig. 4b — Situation in which $\sigma_{\mathrm{DF}} \ll \sigma_R$, causing errors when $\theta_{1i} > \theta_{2i}$

For $\sigma > 1$ the quantity varied in the importance-sampling formulation was μ_a, the true separation in azimuth of the two radar targets. The primary type of classification error for $\sigma > 1$ is, as described earlier, that for which $\theta_{1i} > \theta_{2i}$. It should be apparent from Fig. 4b that the frequency of this type of error should increase significantly as μ_a becomes smaller. Also, for $\sigma < 1$ as in Fig. 4a changing μ_a enough to significantly effect P_e would result in the ratio of the true to the modified probability density function becoming usually very small but occasionally very large. Therefore, to hold down the variance of the estimate of P_e, varying μ_a is restricted to $\sigma > 1$.

A significant amount of computation was saved by transforming Eq. (6) to an equivalent form

$$d' = \gamma\left(z_1^2 - z_{n+1}^2\right) + \sum_{k=2}^{n} \left(z_k^2 - z_{n+k}^2\right), \tag{8}$$

where

$$\gamma = \sqrt{1 + \frac{2}{\sigma^2}}$$

and where the z_k are independent and Gaussian with unity variance. All the z_k except z_1 and z_{n+1} have zero mean, and

$$E\{z_1\} = Xm_1$$

and

$$E\{z_{n+1}\} = Xm_2 ,$$

where

$$X = \mu\sqrt{n} = \frac{\mu_a}{\sigma_{\mathrm{DF}}}\sqrt{n} ,$$

$$m_1 = \frac{1+\gamma}{2\sigma\gamma} ,$$

and

$$m_2 = \frac{1 - \gamma}{2\sigma\gamma}.$$

The quantity X is the quantity against which P_e was plotted in Figs. 1 and 3. This equivalent formulation is derived in Appendix B. Varying μ_a in Eq. (6) is equivalent to varying X (which is proportional to μ_a) here. The computational advantage is obtained because only two random variables are affected now, whereas varying μ_a in Eq. (6) affected $2n$ random variables.

The error-rate estimate is now

$$\hat{P}_e = \frac{1}{N_T} \sum_{k=1}^{N_T} \exp \left\{ -\frac{1}{2} \left[(z_1 - Xm_1)^2 + (z_{n+1} - Xm_2)^2 \right. \right.$$
$$\left. \left. - (z_1 - X'm_1)^2 - (z_{n+1} - X'm_2)^2 \right] \right\} \quad U(-d'),$$

where X' is the modified value of X.

The results of this simulation are shown in Figs. 3d, 3e, and 3f. The $\sigma = 1$ case was included to allow a comparison of the two simulations. The smoothness of the curves in Figs. 3d, 3e, and 3f does not reflect superior accuracy compared to Figs. 3a, 3b, and 3c, but instead reflects the fact that the points on the curves are not independent. Because the parameter modified in the importance-sampling procedure is the parameter used as the independent variable in the plots, and because the value to which it is modified (selected as before for an actual error rate of 0.15 to 0.25) depends only on σ, the actual set of N_T trials was carried out only once for each combination of n and σ and was not repeated for different values of X. The only things that changed with X were the weights given to the outcomes of the trials. In fact when a calculation of d' at a higher value of n followed a calculation using a smaller n (but with other parameters unchanged), only those terms of d' corresponding to the difference between the two n values needed to be calculated. These observations greatly reduced the amount of computation from that required with all points on the curves independently calculated. They also imply that both the points within a curve and the curves within a plot in Figs. 3d, 3e, and 3f are dependent. For example, comparison with the curves for other values of σ indicates that Fig. 3c may be a better representation of true P_e than Fig. 3d, even though Fig. 3d appears neater.

P_e WITH NO SMOOTHING (NS)

When no smoothing is used, Eq. (1) becomes

$$d = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\theta_{\mathrm{DF}i} - \theta_{2i}\right)^2\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\left(\theta_{\mathrm{DF}i} - \theta_{1i}\right)^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(-2\theta_{\mathrm{DF}i}\theta_{2i} + 2\theta_{\mathrm{DF}i}\theta_{1i} + \theta_{2i}^2 - \theta_{1i}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\theta_{2i} - \theta_{1i}\right)\left[\left(\theta_{2i} + \theta_{1i}\right) - 2\theta_{\mathrm{DF}i}\right].$$

With definition of the variables

$$v_i = \frac{\theta_{2i} - \theta_{1i}}{\sqrt{2}\sigma}$$

and

$$w_i = \frac{\theta_{2i} + \theta_{1i} - 2\theta_{\mathrm{DF}i}}{\sqrt{2}\sqrt{\sigma^2 + 2}}$$

d becomes

$$d = \frac{2\sigma\sqrt{\sigma^2 + 2}}{n}\sum_{i=1}^{n} v_i w_i \,.$$

It is easily shown that v_i and w_i are independent Gaussian random variables with unit variance. The calculation of their means is postponed until needed. After the substitutions

$$v_i = \frac{y_i - x_i}{\sqrt{2}}$$

and

$$w_i = \frac{y_i + x_i}{\sqrt{2}}$$

d becomes

$$d = \frac{\sigma\sqrt{\sigma^2 + 2}}{n}\left[\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} x_i^2\right)\right]. \tag{9}$$

Solving for x_i and y_i in terms of v_i and w_i results in

$$x_i = \frac{w_i - v_i}{\sqrt{2}}$$

and

$$y_i = \frac{w_i + v_i}{\sqrt{2}}$$

It is readily seen that the x_i and y_i are Gaussian variables independent of each other with unit variance. The parenthesized terms in Eq. (9) are therefore the independent noncentral chi-squared [6] variates defined by

$$x^2 = \sum_{i=1}^{n} x_i^2$$

and

$$y^2 = \sum_{i=1}^{n} y_i^2$$

with noncentrality parameters a^2 and b^2 respectively, where

$$a^2 = \sum_{i=1}^{n} \left(E\{x_i\}\right)^2$$

and

$$b^2 = \sum_{i=1}^{n} \left(E\{y_i\}\right)^2 .$$

Substituting for x_i and y_i yields

$$a^2 = \sum_{i=1}^{n} \left(\frac{E\{w_i\} - E\{v_i\}}{\sqrt{2}} \right)^2$$

and

$$b^2 = \sum_{i=1}^{n} \left(\frac{E\{w_i\} + E\{v_i\}}{\sqrt{2}} \right)^2 ,$$

where

$$E\{v_i\} = \frac{\mu}{\sqrt{2}\sigma}$$

and

$$E\{w_i\} = \frac{\mu}{\sqrt{2}\sqrt{\sigma^2 + 2}} .$$

Consequently

$$a = \frac{\sqrt{n}\mu}{2}\left(\frac{1}{\sigma} - \frac{1}{\sqrt{\sigma^2 + 2}}\right)$$

and

$$b = \frac{\sqrt{n}\mu}{2}\left(\frac{1}{\sigma} + \frac{1}{\sqrt{\sigma^2 + 2}}\right).$$

The probability density functions for x^2 and y^2 are [6]

$$p_{x^2}(x^2) = \frac{1}{2}\left(\frac{x}{a}\right)^{(n-2)/2} \exp\left(-\frac{a^2}{2} - \frac{x^2}{2}\right) I_{n/2-1}(ax)$$

and

$$p_{y^2}(y^2) = \frac{1}{2}\left(\frac{y}{b}\right)^{(n-2)/2} \exp\left(-\frac{b^2}{2} - \frac{y^2}{2}\right) I_{n/2-1}(by) .$$

Rewriting Eq. (9) as

$$d = \frac{\sigma\sqrt{\sigma^2 + 2}}{n}\left(y^2 - x^2\right)$$

shows the probability that $d < 0$ is the same as the probability that $y^2 < x^2$, or

$$P_e = \int_0^\infty p_{x^2}(x^2)\, d(x^2) \int_0^{x^2} p_{y^2}(y^2)\, d(y^2) .$$

Substituting for the density functions and carrying out a change of variables results in

$$P_e = (ab)^{-\nu} \int_0^{\infty} x^{\nu+1} e^{-(x^2+a^2)/2} I_\nu(ax)\, dx \int_0^{x} y^{\nu+1} e^{-(y^2+b^2)/2} I_\nu(by)\, dy \,,$$

where $\nu = (n-2)/2$ and n is assumed even. This integral was evaluated by Price in Appendix C of Ref. 7, which is reproduced in this report as Appendix C (with comments and errata for Appendix C being given in Appendix D). From Eqs. (C-37) and (C-50) the probability of error is

$$P_e = Q\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right) + e^{-(a^2+b^2)/4} \sum_{m=0}^{\nu} D_m I_m\left(\frac{ab}{2}\right),$$

where

$$D_m = -\frac{1}{2}\delta_{m0} + \left[\left(\frac{b}{a}\right)^m - \left(\frac{b}{a}\right)^{-m}\right] \sum_{j=m}^{\nu} \binom{\nu+j}{\nu+m} 2^{-j-\nu-1} .$$

This result is expressed in terms of the Marcum Q-function, defined in Eq. (C-4), and I_m, the modified Bessel function of the first kind. The factor δ_{m0} is the Kronecker delta function, which is equal to one when its subscripts are equal and is otherwise equal to zero. P_e can be derived in closed form for odd n by expressing it in terms of the doubly noncentral F distribution [8], but the resulting expression is so complex as to make its evaluation in the present context unjustified. For the special case $n = 1$ the resulting P_e is identical with that for the other smoothing options.

P_e was evaluated with the aid of the techniques outlined in Ref. 9 for the evaluation of the Q-function and Ref. 10 for the Bessel functions. The results are plotted in Figs. 5a through 5f. Not surprisingly, NS turns out to be the poorest performer of the four smoothing options.

BOUNDING PERFORMANCE AS RESTRICTIONS ARE REMOVED

Under the assumptions outlined at the beginning of this report the best performance is obtained (Fig. 1) when the radar measurements are smoothed prior to their use in a discriminant. As long as the radar measurements are smoothed, performance is not affected by smoothing the DF measurements. This is somewhat surprising, because in the absence of

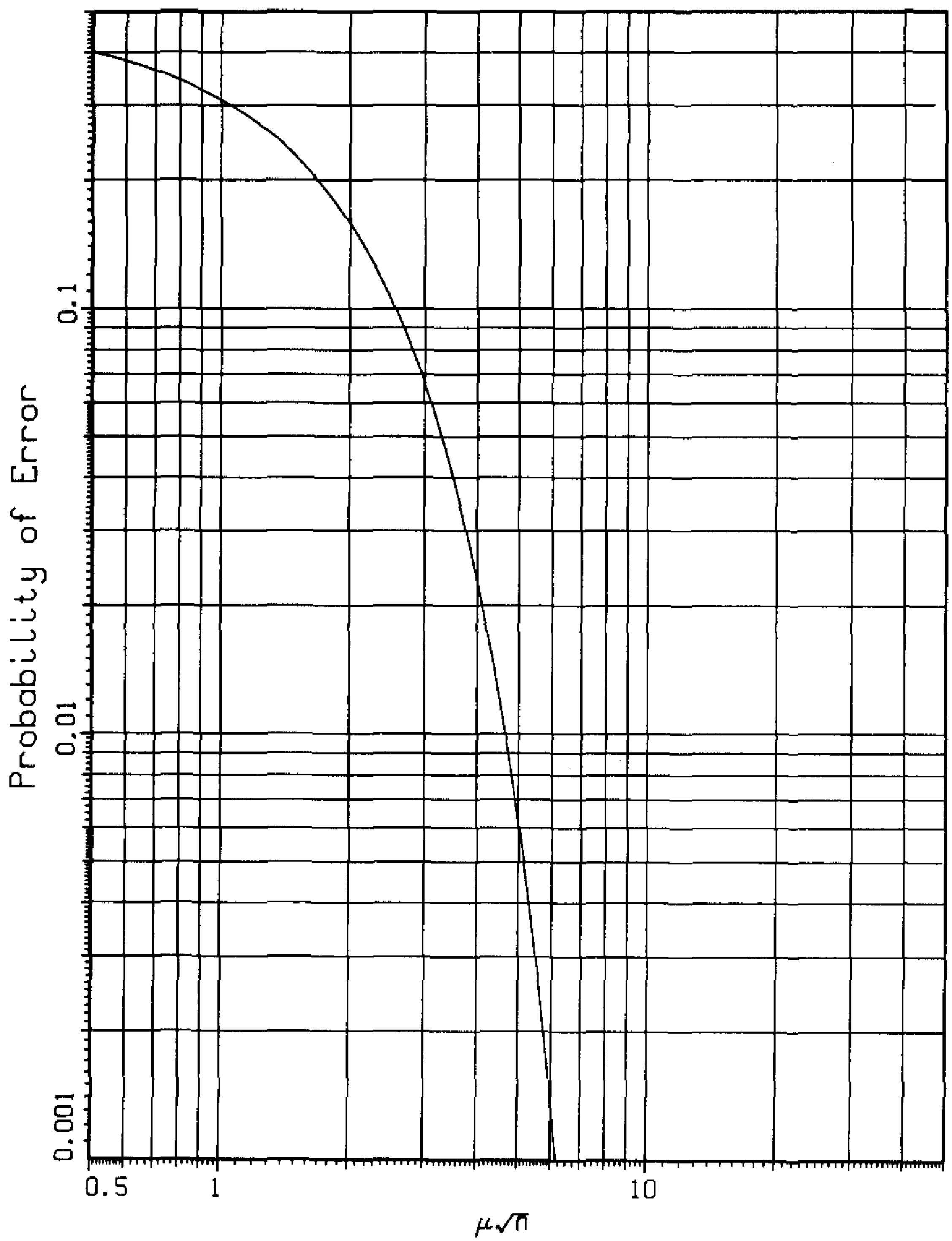

Fig. 5a — Performance with no smoothing: $\sigma = 0$, with $n = 1, 2, 4, 8$, and 16

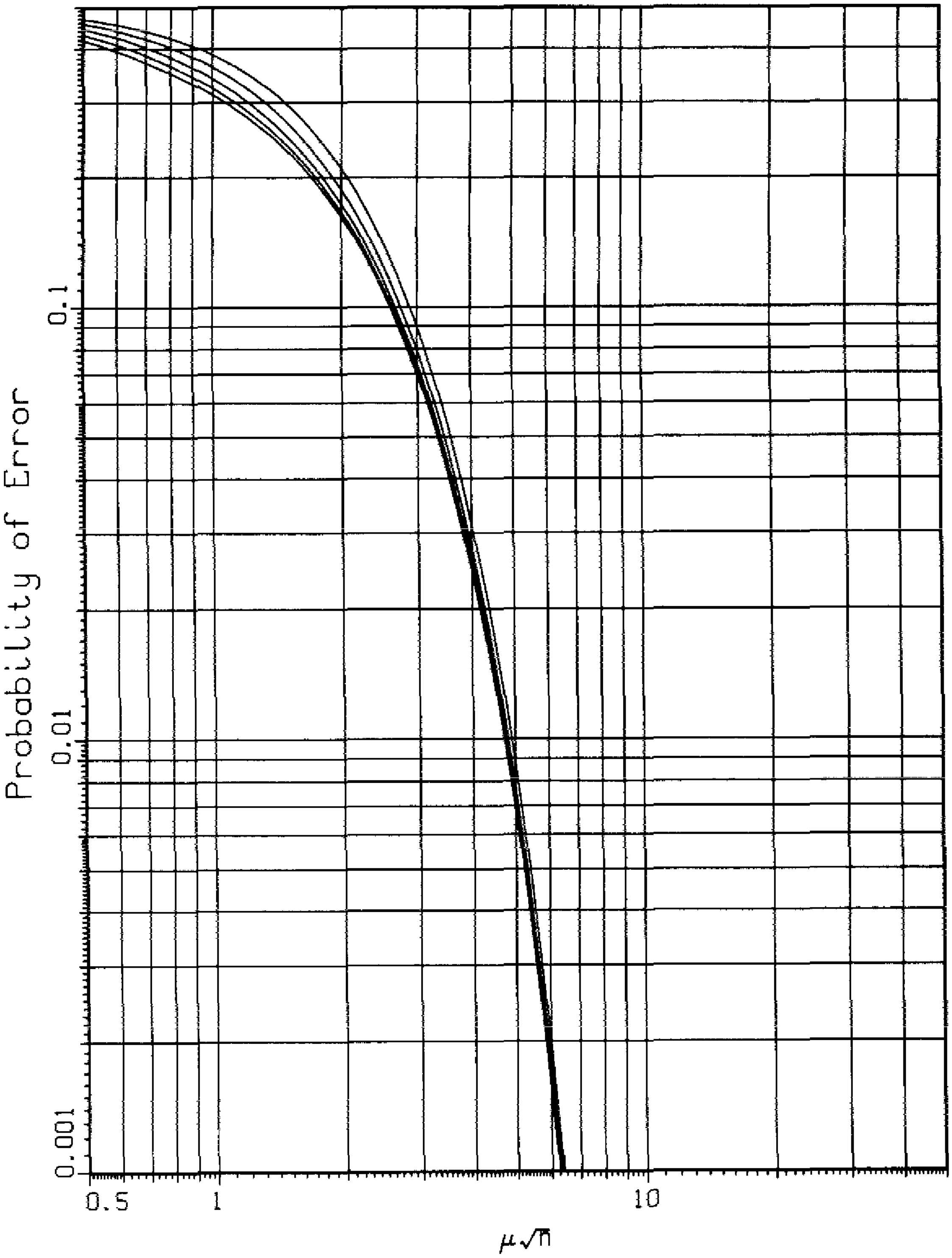

Fig. 5b — Performance with no smoothing: $\sigma = 0.25$, with n = 1, 2, 4, 8, and 16

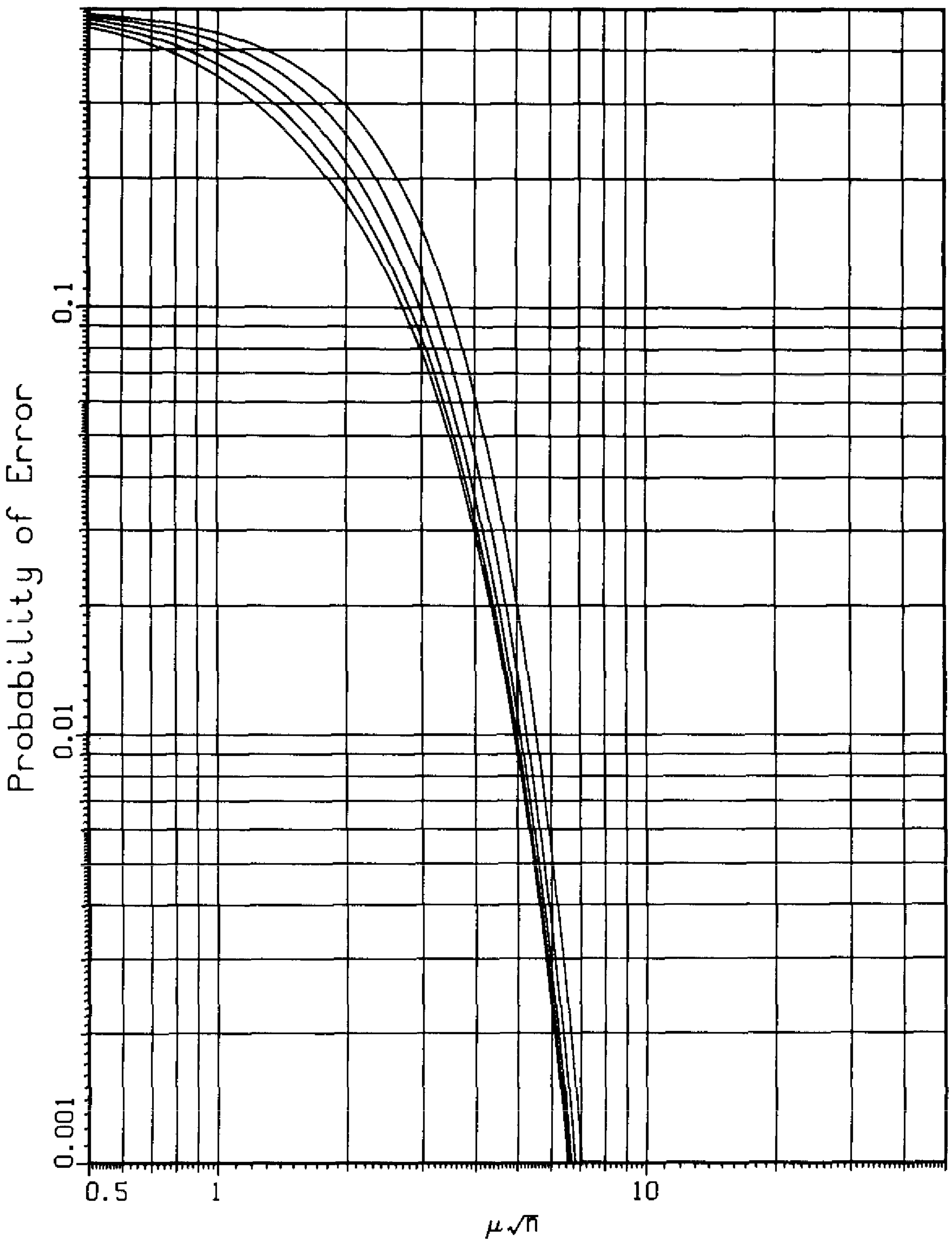

Fig. 5c — Performance with no smoothing: σ = 0.5, with n = 1, 2, 4, 8, and 16

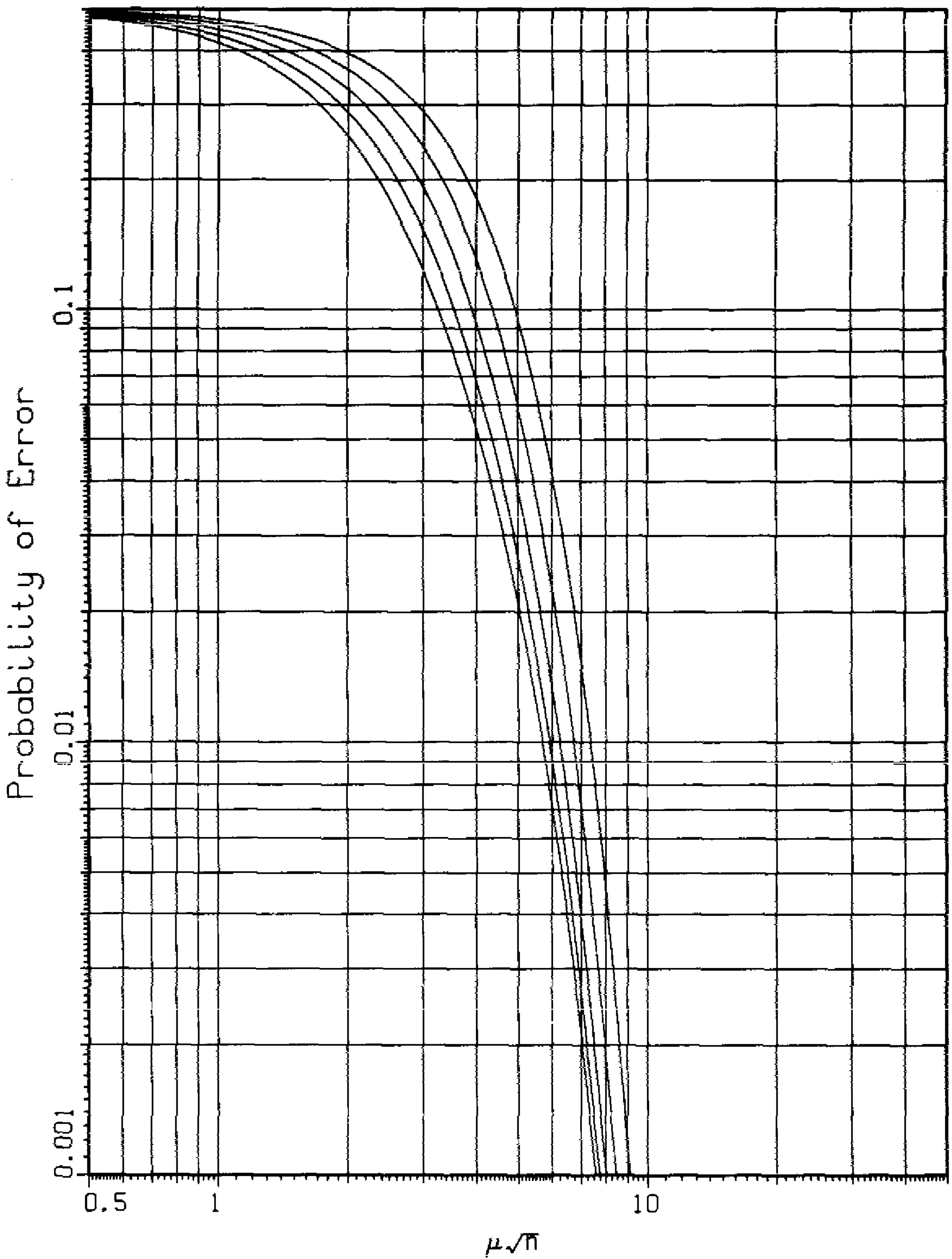

Fig. 5d — Performance with no smoothing: $\sigma = 1$, with $n = 1, 2, 4, 8$, and 16

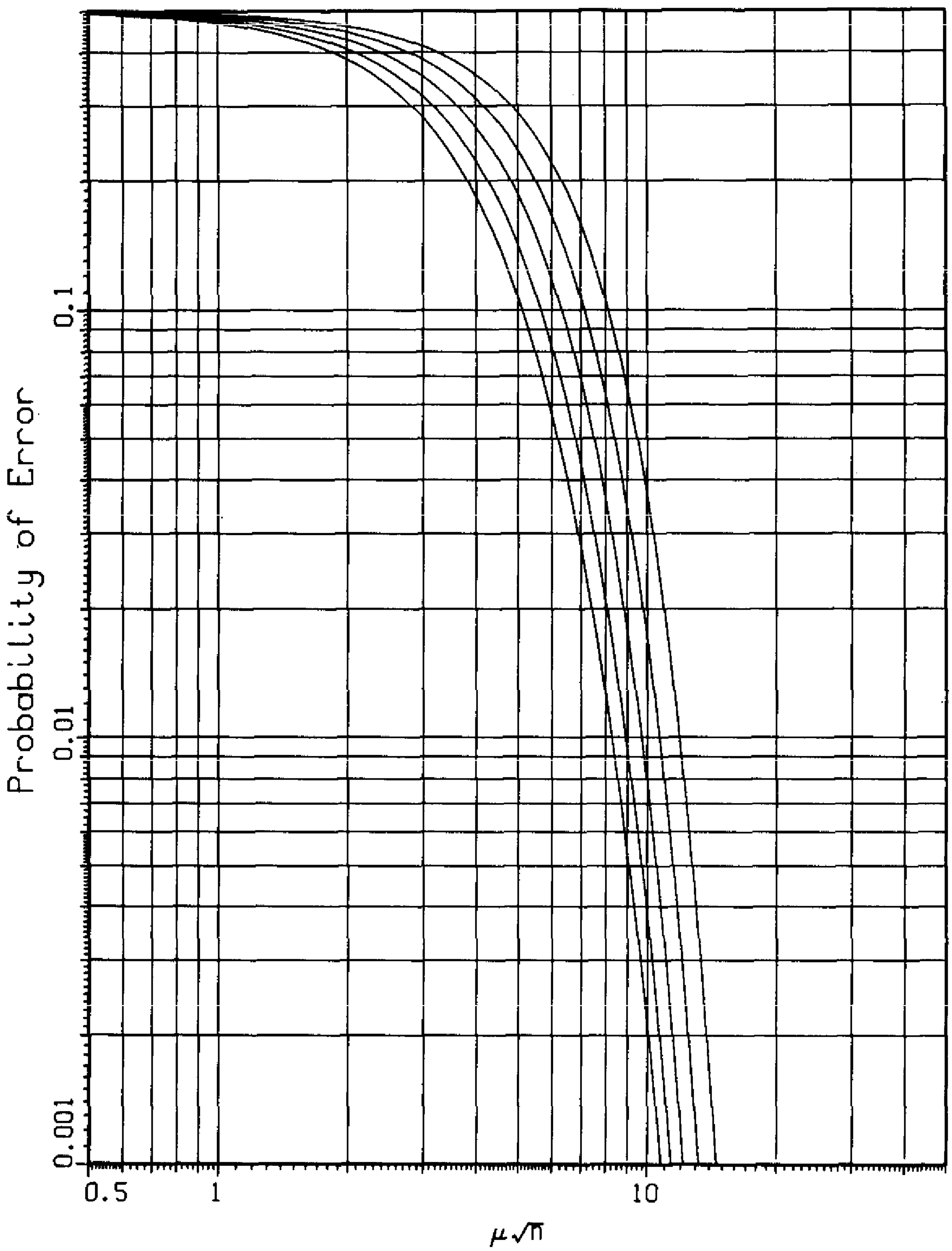

Fig. 5e — Performance with no smoothing: σ = 2, with n = 1, 2, 4, 8, and 16

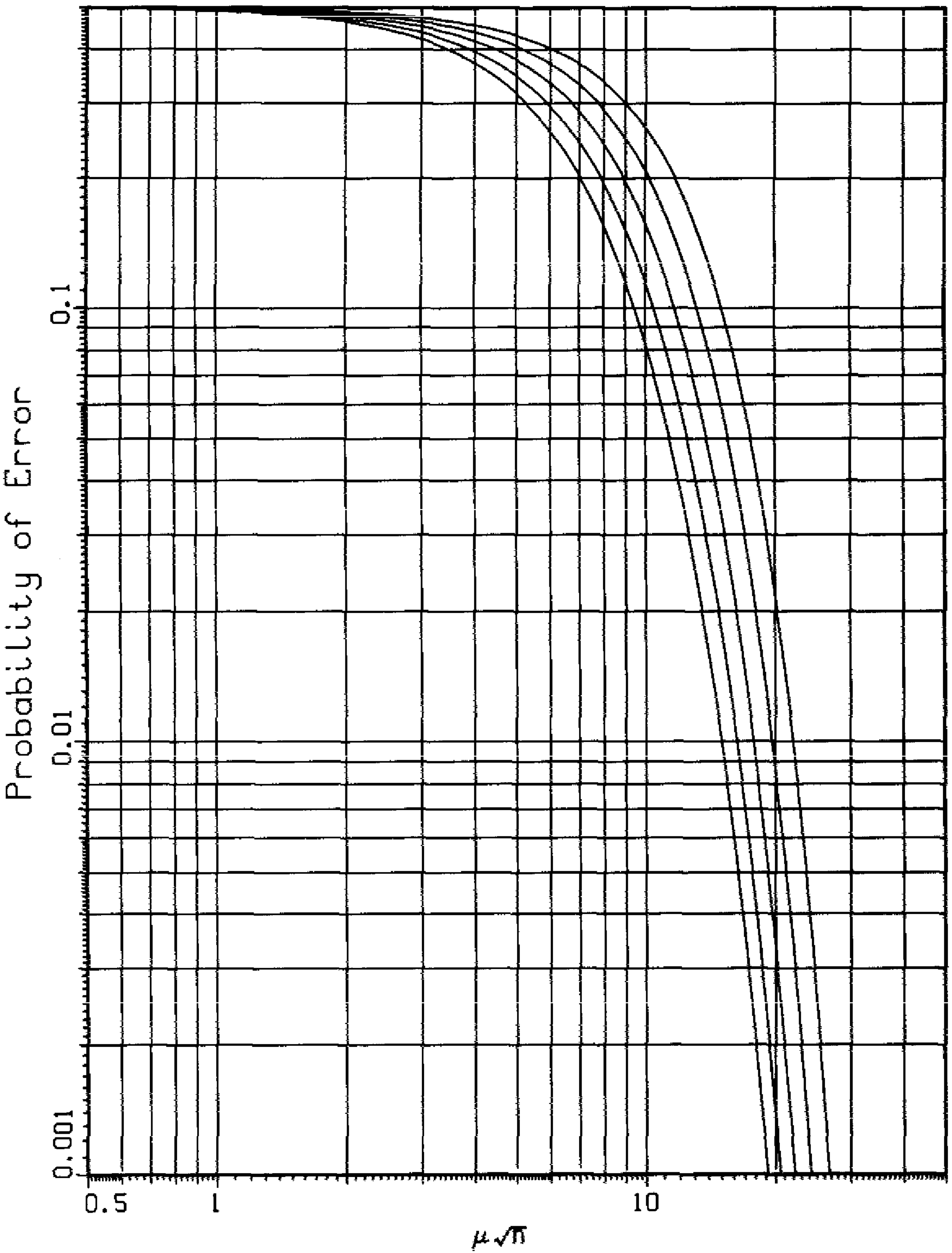

Fig. 5f — Performance with no smoothing: $\sigma = 4$, with $n = 1, 2, 4, 8$, and 16 smoothing of the radar data performance is improved by smoothing the DF measurements (Figs. 3 and 5). Although the probability of error derivations given earlier do not generalize easily to multiple (more than two) radar targets, these same conclusions are expected to hold.

With nonstationary targets the situation is expected to be somewhat different. The simple averaging used for smoothing with stationary targets would have to be replaced by a significantly more complex procedure. Examples of suitable procedures are α–β tracking [11] and Kalman filtering [12]. Suppose the radar and DF systems are synchronized in time so that when a DF measurement is made, the radar system simultaneously measures the positions (in two or three dimensions for best smoothing performance) of all of the targets. Even under these rather ideal conditions, the quality of a smoothed estimate would be reduced from the stationary target case, and hence the remaining performance gain due to smoothing would be diminished. The amount of the remaining performance gain would depend on the number of degrees of freedom in the motion of the targets. If the targets were constrained to move in a straight line (two degrees of freedom if the motion is confined to a plane), performance might be expected to approach that of the nonstationary case (with a large number of measurements). At the other extreme, if no constraints whatever were placed on the motion of the targets, positions at different instances in time would be completely independent and smoothing would contribute nothing (fortunately this is not realistic). The probability-of-error curves in Figs. 1 and 5 therefore represent lower and upper bounds respectively on the probability of error with target motion, synchronous measurements, and radar smoothing only. With DF smoothing only and conditions otherwise as just described the probability of error would be bounded by Figs. 3 as the lower bound and Figs. 5 as the upper bound. (The assumption of independent unbiased Baussian measurement errors remains in force throughout this discussion.)

If the radar and DF measurements were made asynchronously and target motion were still permitted (there being no need to synchronize the measurements if targets are guaranteed stationary), additional difficulties would be introduced. Subtracting an estimated azimuth (as is done under two of the four options discussed) is meaningful only when the estimate refers to the time the measurement was made. Comparing radar and DF measurements directly in this context is not possible, and some sort of smoothing or extrapolating procedure would be necessary. If only the radar data were smoothed, they would have to be extrapolated to the times of the DF measurements. This type of extrapolation introduces additional error, as it must be based on inexact velocity estimates. Due to this additional error, the error rate can never be expected to equal the lower bound given by Fig. 1. No upper bound on the probability of error in this situation is available.

With asynchronous operation and target motion, smoothing the DF data without smoothing the radar data introduces problems. With radar measurements of different targets taken at different times, to what time should the DF data be extrapolated? To extrapolate to the time of each radar measurement in turn could require an unrealistic amount of computation. Realistic or not, Figs. 3 provide the lower bound on error rate. No upper bound is available.

Table 2 identifies the figures representing the bounds described in this section.

Table 2 — Summary of Figures Bounding the Probability of Error

Target Motion Allowed?	Synchronous Measurements	Figure Numbers of P_e Bounds					
		RS		DS		RDS	
		Lower	Upper	Lower	Upper	Lower	Upper
No	Irrelevant	1,2	1,2	3	3	1,2	1,2
Yes	Yes	1,2	5	3	5	1,2	NA
Yes	No	1,2	NA	3	NA	1,2	NA

CONCLUSIONS

From the discussion thus far it is apparent that in the general target-motion-with-asynchronous-operation environment, smoothing of the radar data should be used. As discussed earlier in this report, the use of DF smoothing in addition to radar-data smoothing does not necessarily improve performance. (Equation (5) showed that with stationary targets no improvement is obtained.) DF smoothing can be used in a limited way, however, to reduce the total computational burden in certain types of environments. For example, the DF data can be preprocessed by breaking up the data into batches in such a way that all measurements within a batch are taken over a short enough time interval that they can be averaged together and submitted to further processing as a single measurement. If the environment is such that several independent DF measurements can be made in a time interval in which total angular motion of a target would never approach the standard deviation of the resulting average, error-rate performance should not be degraded. This can be understood (at least in terms of the stationary-target case) by noting that if DF measurements are batched into groups of n_b in the manner described, then in Eq. (4) σ_{DF} must be replaced by $\sigma_{\mathrm{DF}}/\sqrt{n_b}$ (the standard deviation of the average) and n must be replaced by n/n_b. To prevent this change of n from inappropriately raising the variance of $\hat{\theta}_1$ and $\hat{\theta}_2$ by n_b, σ_R must be replaced by $\sigma_R/\sqrt{n_b}$. The effects of these substitutions exactly cancel. The reduction of n in Eq. (5) will more than compensate for the computing time required for the preprocessing.

Now that the desirability of smoothing the radar data has been established, how much (or what method of) smoothing is appropriate? Comparison of Eqs. (1), (3), and (5) shows that the performance difference between the best options (RDS or RS) and each of the other is zero if $\sigma = 0$ or $n = 1$ and increases as either σ or n becomes greater. This is consistent with the intuitively satisfying notion that the details of the smoothing technique will not significantly affect performance as long as the variance of the smoothed radar data is less than the variance of the DF bearing against which it is compared. Both because of radar data smoothing and because this DF bearing may be the result of averaging together a number of DF measurements, the relationship of these two variances may be significantly different than the relationship between the variances of the measurements themselves.

REFERENCES

1. K. R. Gerlach, "Autocorrelator for a Radar/Passive Detection System," NRL Report 8243, Aug. 21, 1978.
2. J. O. Coleman, "Discriminants for Assigning Passive Bearing Observations to Radar Targets", IEEE 1980 International Radar Conference Record.
3. W. G. Bath, "Automatic Tracking of Radar and ESM Returns for Point Defense Application", Johns Hopkins Univ., Applied Physics Lab, Report F3A-77-2-002, July 1977.
4. H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I, Wiley, New York, 1968, pp. 92-96.
5. M. A. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, No. 55, June 1964, Eqs. 7.1.2, 7.1.9, and 7.1.22.
6. A. D. Whalen, *Detection of Signals in Noise*, Academic Press, New York, 1971, pp. 113-114.
7. R. Price, "Error Probabilities for Adaptive Multichannel Reception of Binary Signals", MIT Lincoln Lab. Tech. Rept. 258, 23 July 1962.
8. R. Price, "Some Non-central F-distributions Expressed in Closed Form", Biometrika 51 (1 and 2), 107-122 (1964).
9. A. R. Didonato and M. P. Jarnagin, "A Method for Computing the Circular Coverage Function", Mathematics of Computation 16, 347-355 (July 1962).
10. M. A. Abramowitz and I. A. Stegun, op. cit., Eqs. 9.3.9, 9.6.26, and 9.7.7.
11. T. R. Benedict and G. W. Bordner, "Synthesis of an Optimal Set of Radar Track-While-Scan Smoothing Equations", IRE Trans. on Automatic Control AC-7 (No. 4), 27-32 (1962).
12. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME, J. Basic Eng., Ser. D, 82, 35-45 (Mar. 1960).

Appendix A

IMPORTANCE SAMPLING

Accuracy in probability estimation by Monte Carlo simulation is obtained by making the number of experimental trials sufficiently high that the event whose probability is being estimated occurs many times. For low probability events, the number of trials required for a desired accuracy can be impractically large. Importance sampling [A1, A2] is a simulation technique which can often reduce the number of trials required by modifying the probability density of the random variable(s) in the simulation to increase the number of occurrences of the event of interest. The estimate is then adjusted in such a way as to remove the bias this would otherwise induce in the estimate. The explanation which follows is patterned after Mitchell [A1].

First consider the ordinary Monte Carlo simulation procedure depicted in Fig. A1. For each trial a Bernoulli random variable y representing the event of interest is generated as a function of an input random vector x characterized by a density function $p(x)$. That is, $y = F(x)$. The variable y is equal to one if the event of interest occurred and zero if it did not. The probability estimate is obtained as the average of y over many trials.

For the simulation using importance sampling shown in Fig. A2 the density $p(x)$ is replaced with a modified density $p_m(x)$ which will cause the event of interest to occur more

Fig. A1 — Ordinary Monte Carlo simulation

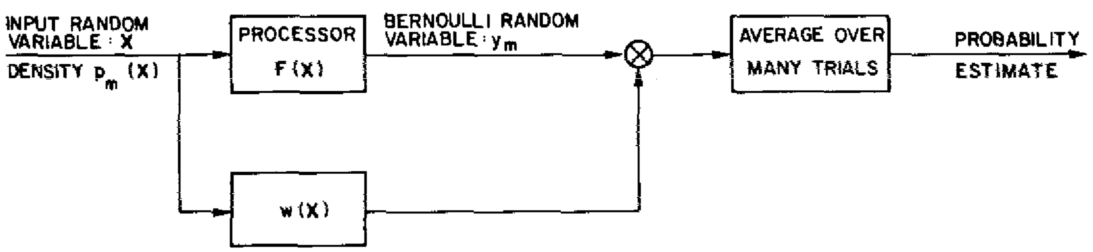

Fig. A2 — Simulation with importance sampling frequently. The processor output is multiplied by a weighting function $w(x)$ to compensate for the change from $p(x)$ to $p_m(x)$. The weighting function must be chosen so that $E\{y\} = E\{w(x)y_m\}$ or

$$\int F(x)p(x)\,dx = \int F(x)w(x)p_m(x)\,dx\,.$$

Therefore the weighting function must be

$$w(x) = \frac{p(x)}{p_m(x)}\,.$$

The weight to be applied for a particular trial is just the ratio of the original to the modified density function evaluated at the specific value of x used for that trial.

The variance of the estimate can be expressed in terms of the first and second moments of the estimate. Because the estimate has been designed to be unbiased, the first moment is equal to the probability being estimated. Obviously in any case of real interest this will not be available in a suitable mathematical form or there would be no need to perform the simulation. The second moment tends to be at least as elusive. In the absence of analysis of the variance of the estimate, the choice of a modified density $p_m(x)$ becomes heuristic. It can be shown, however, that in cases involving simple (analytically tractable) functions $p(x)$ and $F(x)$, importance sampling can dramatically improve estimation accuracy [A1, A2].

REFERENCES

A1. R. L. Mitchell, "Importance Sampling Applied to Simulation of False Alarm Statistics", to be published in IEEE Trans. on Aerospace and Electronic Systems (around January 1981).

A2. V. G. Hansen, "Importance Sampling in Computer Simulation of Signal Processors", Comput. and Elect. Engrg. 1 (No. 4), 545-550 (1974).

Appendix B

DERIVATION OF EQ. (8)

Equation (6) can be rewritten using vectors rather than summations as

$$d = \frac{1}{n} \left\| \theta_{DF} - \theta_2 \right\|^2 - \frac{1}{n} \left\| \theta_{DF} - \theta_1 \right\|^2$$

by defining the n-vectors

$$\Theta_{DF} = \begin{bmatrix} \hat{\theta}_{DF} \\ \hat{\theta}_{DF} \\ \cdots \\ \hat{\theta}_{DF} \end{bmatrix}, \quad \Theta_1 = \begin{bmatrix} \theta_{11} \\ \theta_{12} \\ \cdots \\ \theta_{1n} \end{bmatrix}, \quad \Theta_2 = \begin{bmatrix} \theta_{21} \\ \theta_{22} \\ \cdots \\ \theta_{2n} \end{bmatrix}.$$

The vectors θ_{DF}, θ_1, and θ_2 are independent Gaussian-distributed random vectors with covariances M, $\sigma^2 I$, and $\sigma^2 I$ respectively, where M is an n-by-n matrix with every element equal to $1/n$ and I is the n-by-n identity matrix.

By defining (for this appendix only) a partitioned vector

$$x = \left[\begin{array}{c} (\Theta_{DF} - \Theta_2) \\ \hdashline (\Theta_{DF} - \Theta_1) \end{array} \right]$$

and a partitioned matrix

$$Q = \left[\begin{array}{c:c} I & 0 \\ \hdashline 0 & -I \end{array} \right]$$

the discriminant equation can be written as

$$nd = x^T Q x .$$

The Gaussian random vector x has a mean

$$E\{x\} = \left[\begin{array}{c} \mu c \\ \hdashline 0 \end{array}\right],$$

where c is defined to be an n-vector whose elements are all unity. The covariance of x, denoted by K, is found to be

$$K = \left[\begin{array}{c:c} \sigma^2 I + M & M \\ \hdashline M & \sigma^2 I + M \end{array}\right].$$

From this point on, linear transformations on x are used to proceed toward the desired discriminant form.

It is simple to demonstrate that the matrix K is positive definite as long as σ is nonzero. Because it is both positive definite and symmetric, it can be factored into symmetric positive-definite square roots:

$$K = SS.$$

It is straightforward to verify that S is given by

$$S = \left[\begin{array}{c:c} \sigma I + \left(\dfrac{\sqrt{\sigma^2 + 2} - \sigma}{2}\right) M & \left(\dfrac{\sqrt{\sigma^2 + 2} - \sigma}{2}\right) M \\ \hdashline \left(\dfrac{\sqrt{\sigma^2 + 2} - \sigma}{2}\right) M & \sigma I + \left(\dfrac{\sqrt{\sigma^2 + 2} - \sigma}{2}\right) M \end{array}\right].$$

Because S is positive definite, its inverse must exist and is obtained as

$$S^{-1} = \left[\begin{array}{c:c} \dfrac{1}{\sigma} I - \dfrac{1}{2}\left(\dfrac{1}{\sigma} - \dfrac{1}{\sqrt{\sigma^2 + 2}}\right) M & -\dfrac{1}{2}\left(\dfrac{1}{\sigma} - \dfrac{1}{\sqrt{\sigma^2 + 2}}\right) M \\ \hdashline -\dfrac{1}{2}\left(\dfrac{1}{\sigma} - \dfrac{1}{\sqrt{\sigma^2 + 2}}\right) M & \dfrac{1}{\sigma} I - \dfrac{1}{2}\left(\dfrac{1}{\sigma} - \dfrac{1}{\sqrt{\sigma^2 + 2}}\right) M \end{array}\right]$$

and used to transform x to a new variable

$$y = S^{-1}x \ .$$

The covariance of y is

$$E\left\{(y - E\{y\})(y - E\{y\})^T\right\} = S^{-1}E\left\{(x - E\{x\})(x - E\{x\})^T\right\}S^{-1}$$

$$= S^{-1}KS^{-1}$$

$$= I$$

and the mean of y is

$$E\{y\} = S^{-1}\left[\begin{array}{c}\mu c \\ \hdashline 0\end{array}\right]$$

$$= \frac{\mu}{2}\left[\begin{array}{c}\left(\dfrac{1}{\sqrt{\sigma^2 + 2}} + \dfrac{1}{\sigma}\right)c \\ \hdashline \left(\dfrac{1}{\sqrt{\sigma^2 + 2}} - \dfrac{1}{\sigma}\right)c\end{array}\right] .$$

The discriminant equation can now be written in terms of the transformed variable y as

$$nd = x^TQx = y^TSQSy = y^TTy \ ,$$

where

$$T = SQS \ .$$

Performing the multiplications and simplifying yields

$$T = \sigma\left[\begin{array}{c:c}\sigma I + \left(\sqrt{\sigma^2 + 2} - \sigma\right)M & 0 \\ \hdashline 0 & -\sigma I - \left(\sqrt{\sigma^2 + 2} - \sigma\right)M\end{array}\right] .$$

The next step in the derivation requires that T first be diagonalized. Matrices U and D are required such that

$$T = U^T D U ,$$

where U is orthogonal ($U^T = U^{-1}$) and D is diagonal (all off-diagonal elements are zero). The matrix U satisfying these requirements is in this case of the form

$$U = \left[\begin{array}{c|c} F & 0 \\ \hline 0 & F \end{array}\right] ,$$

where F is any n-by-n orthogonal matrix whose first row is composed of identical elements. It is not necessary to specify F further in order to complete the derivation. For convenience the value in the first row will be assumed positive, which implies a value of $1/\sqrt{n}$. A familiar example of such an F matrix is a discrete-Fourier-transform (DFT) matrix.

Solving for the diagonal matrix D results in

$$D = \sigma^2 \left[\begin{array}{ccccc|ccccc} \sqrt{1 + 2/\sigma^2} & 0 & 0 & \cdots & 0 & & & & & \\ 0 & 1 & 0 & \cdots & 0 & & & & & \\ 0 & 0 & 1 & \cdots & \cdots & & & 0 & & \\ \cdots & \cdots & \cdots & \cdots & 0 & & & & & \\ 0 & 0 & \cdots & 0 & 1 & & & & & \\ \hline & & & & & -\sqrt{1 + 2/\sigma^2} & 0 & 0 & \cdots & 0 \\ & & & & & 0 & -1 & 0 & \cdots & 0 \\ & & 0 & & & 0 & 0 & -1 & \cdots & \cdots \\ & & & & & \cdots & \cdots & \cdots & \cdots & 0 \\ & & & & & 0 & 0 & \cdots & 0 & -1 \end{array}\right] .$$

A new random vector can now be defined as

$$z = Uy ,$$

and the discriminant equation becomes

$$nd = y^T T y = (U^T z)^T (U^T D U)(U^T z) = z^T D z .$$

The covariance of the new variable is the same as the old:

$$\operatorname{cov}\{z\} = \operatorname{cov}\{Uy\} = U \operatorname{cov}\{y\} U^T = UIU^T = I.$$

The mean of z is

$$E\{z\} = UE\{y\}$$

or

$$E\{z\} = \frac{\mu\sqrt{n}}{2} \begin{bmatrix} \left(\dfrac{1}{\sqrt{\sigma^2+2}} + \dfrac{1}{\sigma}\right) \\ 0 \\ \cdots \\ 0 \\ \hline \left(\dfrac{1}{\sqrt{\sigma^2+2}} - \dfrac{1}{\sigma}\right) \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

If z is expanded to

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_{2n} \end{bmatrix},$$

the discriminant equation can be written in nonmatrix form as

$$\frac{nd}{\sigma^2} = \sqrt{1 + \frac{2}{\sigma^2}}\left(z_1^2 - z_{n+1}^2\right) + \sum_{k=2}^{n}\left(z_k^2 - z_{n+k}^2\right),$$

where the z_k are independent Gaussian random variables with unit variance. Defining

$$d' = \frac{nd}{\sigma^2}$$

and noting that the means of the z values here match those in the text completes the derivation.

APPENDIX C

APPENDIX C [Reproduction of Appendix C of Ref. 7] REDUCTION OF A DOUBLE INTEGRAL INVOLVING BESSEL FUNCTIONS*

R. Price, *MIT Lincoln Laboratory*

I. INTRODUCTION

Maximon[29] has studied the following integral, which for $\nu = 0$ arises in trying to find the probability that the envelope y of a sine-wave-plus-narrow-band-noise process is less than r times the envelope x of another such process, the noises being assumed independent and both of unit variance.

$$P_{-\nu}(a,b,r) = a^{\nu} b^{-\nu} \int_0^{\infty} x^{-\nu+1} \exp[-(x^2 + a^2)/2]\, I_{\nu}(ax)\, dx \cdot \int_0^{xr} y^{\nu+1} \exp[-(y^2 + b^2)/2]\, I_{\nu}(by)\, dy \qquad \text{(C-1)}$$

Here $I_{\nu}(z)$ is the modified Bessel function of nonnegative integral order ν, and a and b are parameters which, for $\nu = 0$, are respectively equal to the amplitudes of the sine waves in the processes whose envelopes are x and y. (See Rice[30] for the derivation of the envelope probability density functions contained in Eq. (C-1), with $\nu = 0$.) Maximon obtains P in terms of a Neumann series of modified Bessel functions:

$$P_{-\nu}(a,b,r) = \left(\frac{a}{b}\right)^{2\nu} \exp\left[-\frac{1}{2}(a^2 r^2 + b^2)/(1 + r^2)\right] \sum_{m=0}^{\infty} \epsilon_m \left(\frac{ar}{b}\right)^{\nu+m} I_{\nu+m}\left(\frac{abr}{1 + r^2}\right) \qquad \text{(C-2)}$$

where

$$\epsilon_m = \begin{cases} r^2/(1 + r^2) & \text{(if } m = 0) \\ 1 & \text{(if } m > 0) \qquad (\nu \text{ an integer} \geqslant 0) \end{cases} \qquad \text{(C-3)}$$

The series (C-2) can be considered in closed form in terms of Lommel's functions of two variables,[31] but tabulations by which the series may be evaluated exist chiefly in terms of the "Q-function" examined by Marcum.[32] This function is the cumulative probability distribution of the envelope of a sine-wave-plus-narrow-band-noise process:

$$\begin{aligned} Q(u,v) &= \int_v^{\infty} x \exp[-(x^2 + u^2)/2]\, I_0(ux)\, dx \\ &= 1 - \exp[-(u^2 + v^2)/2] \sum_{m=1}^{\infty} \left(\frac{v}{u}\right)^m I_m(uv) \\ &= \exp[-(u^2 + v^2)/2] \sum_{m=0}^{\infty} \left(\frac{u}{v}\right)^m I_m(uv) \end{aligned} \qquad \text{(C-4)}$$

---

* Jones[33] has recently and independently reduced the double integral [Eq. (C-6)] for $\nu = 1 = r$, obtaining the particular case of the solution [Eq. (C-35)] where $m = \nu = 1$ [given also by Eq. (C-41)].

The first series is due to Bennett, and the second, which is connected with the reciprocation properties of Lommel's functions of two variables,[31] is quoted by Helstrom.[23]

In this Appendix we deal with a double integral similar to Eq. (C-1), which arises in finding the probability that the sum Y of a number $(\nu + 1)$ of squared envelopes of sine-wave-plus-narrow-band-noise processes is less than r^2 times another such sum X where all the noises are independent and have unit variance.* The probability density function of X is[14]

$$p(X) = \frac{1}{2}(X/\alpha)^{\nu/2} \exp[-(X+\alpha)/2]\, I_\nu \sqrt{X\alpha} \tag{C-5}$$

and that of Y is identical to Eq. (C-5) with β in place of α. Here α and β are proportional to the sums of the sine-wave powers (squared amplitudes) in the processes yielding X and Y, respectively. The double integral for the probability that Y is less than r^2 times X is thus, after making the changes of variables $X = x^2$, $Y = y^2$ and $\alpha = a^2$, $\beta = b^2$:

$$P_\nu(a,b,r) = (ab)^{-\nu} \int_0^\infty x^{\nu+1} \exp[-(x^2 + a^2)/2]\, I_\nu(ax)\, dx \cdot \int_0^{xr} y^{\nu+1} \exp[-(y^2+b^2)/2]\, I_\nu(by)\, dy \quad . \tag{C-6}$$

II. DEVELOPMENT OF A RECURSION RELATION FOR $P_\nu(a, b, r)$

The first step in attacking Eq. (C-6) is to integrate by parts:

$$\exp[(a^2 + b^2)/2]\ (ab)^\nu P_\nu(a,b,r) = -\exp[-x^2/2]\left[x^\nu I_\nu(ax) \int_0^{xr} (\ )\, dy\right]\Big|_0^\infty + \int_0^\infty e^{-x^2/2} \frac{d}{dx}\left[x^\nu I_\nu(ax) \int_0^{xr} y^{\nu+1} e^{-y^2/2} I_\nu(by)\, dy\right] dx \quad , \tag{C-7}$$

where, for $\nu \geqslant 1$, the first term on the right of Eq. (C-7) vanishes. We now make use of the Bessel-function property[34]

$$\frac{1}{z}\frac{d}{dz}[z^\nu I_\nu(z)] = z^{\nu-1} I_{\nu-1}(z) \quad , \tag{C-8}$$

which yields

$$\frac{d}{dx}[x^\nu I_\nu(cx)] = cx^\nu I_{\nu-1}(cx) \quad . \tag{C-9}$$

Expanding the derivative in Eq. (C-7) and applying Eq. (C-9), we obtain

$$\exp[(a^2+b^2)/2](ab)^\nu P_\nu(a,b,r) = a\int_0^\infty x^\nu \exp[-x^2/2] I_{\nu-1}(ax) \left[\int_0^{xr} y^{\nu+1} \exp[-y^2/2] I_\nu(by)\, dy\right] dx + r^{\nu+2} \int_0^\infty x^{2\nu+1} \exp[-x^2(1+r^2)/2] I_\nu(ax) I_\nu(brx)\, dx \quad . \tag{C-10}$$

Again integrating by parts in the inner integral of Eq. (C-10), and using Eq. (C-9),

---

*** In terms of classical statistics, this is equivalent to finding the probability distribution function of the ratio of the lengths of two $(2\nu + 2)$-dimensional random vectors, all of whose components are mutually independent and Gaussian with unit variance, and whose two mean vectors are of specified lengths. It appears that the method of solution followed here could be generalized to vectors of different dimensionalities.**

$$\int_0^{xr} y^{\nu+1} \exp[-y^2/2]\, I_\nu(by)\, dy = -\exp[-y^2/2]\, y^\nu I_\nu(by) \Big|_0^{xr}$$

$$+ b \int_0^{xr} y^\nu \exp[-y^2/2]\; I_{\nu-1}(by)\, dy \qquad \text{(C-11)}$$

which, when substituted into Eq. (C-10) yields, for $\nu \geqslant 1$,

$$\exp\{(a^2+b^2)/2\}(ab)^\nu P_\nu(a,b,r) = ab \int_0^\infty x^\nu \exp[-x^2/2]\, I_{\nu-1}(ax)\, dx \int_0^{xr} y^\nu \exp[-y^2/2]\, I_{\nu-1}(by)\, dy$$

$$+ r^{\nu+2} \int_0^\infty x^{2\nu+1} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_\nu(brx)\, dx$$

$$-ar^\nu \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_{\nu-1}(ax)\, I_\nu(brx)\, dx \quad . \qquad \text{(C-12)}$$

In order to cast Eq. (C-12) in a more symmetric form for later convenience, we note that, integrating by parts and using Eq. (C-9) once again, with $\nu \geqslant 1$,

$$r^{\nu+2}\int_0^\infty x^{2\nu+1} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_\nu(brx)\, dx = -\frac{r^{\nu+2}\, x^{2\nu}}{1+r^2} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_\nu(brx) \Big|_0^\infty$$

$$+ \frac{r^{\nu+2}}{1+r^2} \int_0^\infty \exp\{-x^2(1+r^2)/2\} \frac{d}{dx} \{[x^\nu I_\nu(ax)]\, [x^\nu I_\nu(brx)]\}\, dx$$

$$= \frac{ar^{\nu+2}}{1+r^2} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_{\nu-1}(ax)\, I_\nu(brx)\, dx$$

$$+ \frac{br^{\nu+3}}{1+r^2} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_{\nu-1}(brx)\, dx \quad . \qquad \text{(C-13)}$$

Combining Eqs. (C-12) and (C-13), we have

$$P_\nu(a,b,r) = (ab)^{1-\nu} \int_0^\infty x^\nu \exp[-x^2/2]\, I_{\nu-1}(ax)\, dx \int_0^{xr} y^\nu \exp[-y^2/2]\, I_{\nu-1}(by)\, dy$$

$$+ \frac{\exp[-(a^2+b^2)/2]}{1+r^2} \Big\{ r^{\nu+3}\, a^{-\nu} b^{1-\nu} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_{\nu-1}(brx)\, dx$$

$$- r^\nu a^{1-\nu} b^{-\nu} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_{\nu-1}(ax)\, I_\nu(brx)\, dx \Big\} \quad ,$$

thus giving the recursion

$$P_\nu(a,b,r) = P_{\nu-1}(a,b,r) + \frac{\exp[-(a^2+b^2)/2]}{1+r^2}$$

$$\cdot \Big\{ r^{\nu+3}\, a^{-\nu} b^{1-\nu} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_\nu(ax)\, I_{\nu-1}(brx)\, dx$$

$$- r^\nu a^{1-\nu} b^{-\nu} \int_0^\infty x^{2\nu} \exp[-x^2(1+r^2)/2]\, I_{\nu-1}(ax)\, I_\nu(brx)\, dx \Big\} \quad . \qquad \text{(C-14)}$$

Equation (C-14) is a reduction formula by which we can obtain the general result if we know $P_o(a,b,r)$ as given by Maximon.[29] It is necessary, however, to evaluate the pair of integrals in Eq. (C-14), and this is where the complication lies.* The integrals are solved by using another Bessel-function property[35] that is similar to Eq. (C-8):

$$\left(\frac{1}{z}\frac{d}{dz}\right)^{\nu} z^{-m} I_m(z) = z^{-\nu-m} I_{\nu+m}(z) \tag{C-15}$$

which leads to

$$\left(\frac{1}{c}\frac{d}{dc}\right)^{\nu} (cx)^{-m} I_m(cx) = x^{\nu-m} c^{-\nu-m} I_{\nu+m}(cx) \tag{C-16}$$

and as a special case

$$\left(\frac{1}{c}\frac{d}{dc}\right)^{\nu} I_o(cx) = x^{\nu} c^{-\nu} I_{\nu}(cx) \quad . \tag{C-17}$$

Therefore,

$$a^{-\nu}(br)^{1-\nu}\int_o^{\infty} x^{2\nu} \exp[-x^2(1+r^2)/2] I_{\nu}(ax) I_{\nu-1}(brx)\, dx = \left(\frac{1}{a}\frac{d}{da}\right)^{\nu}\left(\frac{1}{br}\frac{d}{d(br)}\right)^{\nu-1}$$
$$\times\int_o^{\infty} x \exp[-x^2(1+r^2)/2] I_o(ax) I_o(brx)\, dx \tag{C-18}$$

and similarly for the other integral in Eq. (C-14), a and (br) simply being interchanged.

The integral in the right member of Eq. (C-18) is*

$$\int_o^{\infty} x \exp[-x^2(1+r^2)/2] I_o(ax) I_o(brx)\, dx = (1+r^2)^{-1} \exp\left[\frac{a^2+b^2r^2}{2(1+r^2)}\right] I_o\left[\frac{abr}{1+r^2}\right] \quad . \tag{C-19}$$

Combining Eqs. (C-14), (C-18) and (C-19),

$$P_{\nu}(a,b,r) = P_{\nu-1}(a,b,r) + \frac{\exp[-(a^2+b^2)/2]}{(1+r^2)^2}\left[\frac{1}{a}\frac{d}{da}\right]^{\nu-1}\left[\frac{1}{br}\frac{d}{d(br)}\right]^{\nu-1}$$
$$\cdot\left\{\left[r^{2\nu+2}\left(\frac{1}{a}\frac{d}{da}\right) - r^{2\nu}\left(\frac{1}{br}\frac{d}{d(br)}\right)\right]\exp\left[\frac{a^2+b^2r^2}{2(1+r^2)}\right] I_o\left[\frac{abr}{1+r^2}\right]\right\}$$
$$= P_{\nu-1}(a,b,r) + \frac{r^{2\nu}\exp[-(a^2+b^2)/2]}{(1+r^2)^3}\left[\frac{1}{a}\frac{d}{da}\right]^{\nu-1}\left[\frac{1}{br}\frac{d}{d(br)}\right]^{\nu-1}$$
$$\cdot\left(\exp\left[\frac{a^2+b^2r^2}{2(1+r^2)}\right]\right)\left([r^2-1]\, I_o\left[\frac{abr}{1+r^2}\right] + \left[\frac{br^3}{a} - \frac{a}{br}\right] I_1\left[\frac{abr}{1+r^2}\right]\right) \tag{C-20}$$

so long as $\nu \geqslant 1$.

III. FINAL SOLUTION FOR $P_{\nu}(a,b,r)$

Iterating the recursion formula [Eq. (C-20)] and reverting to the variables $\alpha = a^2$, $\beta = b^2$, we obtain

*Such integrals are treated by Watson (Ref. 31, pp. 395-396) in connection with Weber's second exponential integral, of which Eq. (C-19) is an example; with regard to integrals of the type appearing in Eq. (C-14) Watson seems unduly pessimistic.

$$P_\nu(\sqrt{\alpha}, \sqrt{\beta}, r) = P_o(\sqrt{\alpha}, \sqrt{\beta}, r) + \frac{\exp[-(\alpha+\beta)]}{4(r+r^{-1})^3}$$

$$\cdot \sum_{j=1}^{\nu} 2^{2j} \left(\frac{d}{d\alpha}\right)^{j-1} \left(\frac{d}{d\beta}\right)^{j-1} \left(\exp\left[\frac{\alpha+\beta r^2}{2(1+r^2)}\right]\right)$$

$$\cdot \left\{(r - r^{-1}) I_o[\sqrt{\alpha\beta}\, r/(1+r^2)] + \left[r^2\sqrt{\frac{\beta}{\alpha}} - r^{-2}\sqrt{\frac{\alpha}{\beta}}\right] I_1[\sqrt{\alpha\beta}\, r/(1+r^2)]\right\} \qquad \text{(C-21)}$$

which, expanding according to the binomial formula for the derivative of a product, differentiating $\exp[(\alpha+\beta r^2)/2(1+r^2)]$, and returning to the variables a, b, yields

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})^3} \sum_{j=1}^{\nu} \sum_{k=0}^{j-1} \sum_{\ell=0}^{j-1} (1+r^2)^{k+\ell+2-2j}\, r^{2j-2\ell-2}$$

$$\cdot \binom{j-1}{k}\binom{j-1}{\ell} \left(\frac{1}{a}\frac{d}{da}\right)^k \left(\frac{1}{b}\frac{d}{db}\right)^\ell \left\{(r-r^{-1}) I_o\left[\frac{abr}{1+r^2}\right]\right.$$

$$\left. + \left[r^2\left(\frac{b}{a}\right) - r^{-2}\left(\frac{a}{b}\right)\right] I_1\left[\frac{abr}{1+r^2}\right]\right\} \quad , \qquad \text{(C-22)}$$

where

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!} \quad \left(\text{if } m < 0 \text{ or } m > n,\ \binom{n}{m} = 0\right) \qquad \text{(C-23)}$$

is the binomial coefficient.

We now need to find the result of operating on the term in the braces of Eq. (C-22). Using Eq. (C-16) with $c = b$, $x = ar/(1+r^2)$, $\nu = \ell$ and $m = 0$ or -1 [note: $I_m(z) = I_{-m}(z)$], leads to

$$\left(\frac{1}{b}\frac{d}{db}\right)^\ell \left\{\frac{b}{a} I_1\left[\frac{abr}{1+r^2}\right]\right\} = \left(\frac{r}{1+r^2}\right)^\ell \left(\frac{b}{a}\right)^{1-\ell} I_{\ell-1}\left[\frac{abr}{1+r^2}\right]$$

$$\left(\frac{1}{b}\frac{d}{db}\right)^\ell I_o\left[\frac{abr}{1+r^2}\right] = \left(\frac{r}{1+r^2}\right)^\ell \left(\frac{b}{a}\right)^{-\ell} I_\ell\left[\frac{abr}{1+r^2}\right] \qquad \text{(C-24)}$$

and further application, this time with $c = a$, $x = br/(1+r^2)$, $\nu = k$, and $m = 1-\ell$ or $-\ell$ yields

$$\left(\frac{1}{a}\frac{d}{da}\right)^k \left(\frac{1}{b}\frac{d}{db}\right)^\ell \left\{\frac{b}{a} I_1\left[\frac{abr}{1+r^2}\right]\right\} = \left(\frac{r}{1+r^2}\right)^{k+\ell} (b/a)^{k-\ell+1}\, I_{k-\ell+1}\left[\frac{abr}{1+r^2}\right] \qquad \text{(C-25)}$$

$$\left(\frac{1}{a}\frac{d}{da}\right)^k \left(\frac{1}{b}\frac{d}{db}\right)^\ell \left\{I_o\left[\frac{abr}{1+r^2}\right]\right\} = \left(\frac{r}{1+r^2}\right)^{k+\ell} (b/a)^{k-\ell}\, I_{k-\ell}\left[\frac{abr}{1+r^2}\right] \quad . \qquad \text{(C-26)}$$

Similarly, we find

$$\left(\frac{1}{a}\frac{d}{da}\right)^k \left(\frac{1}{b}\frac{d}{db}\right)^\ell \left\{\frac{a}{b} I_1\left[\frac{abr}{1+r^2}\right]\right\} = \left(\frac{r}{1+r^2}\right)^{k+\ell} (a/b)^{\ell-k+1}\, I_{\ell-k+1}\left[\frac{abr}{1+r^2}\right] \quad . \qquad \text{(C-27)}$$

Substituting Eqs. (C-25) through (C-27) in Eq. (C-22), and interchanging the dummy variables k and ℓ in the part of Eq. (C-22) corresponding to the term $r^{-2}\,(a/b)\, I_1\,[abr/(1+r^2)]$, we are led to

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \sum_{j=1}^{\nu} (r+r^{-1})^{-2j} \sum_{k=0}^{j-1} \sum_{\ell=0}^{j-1} \binom{j-1}{k}\binom{j-1}{\ell}$$

$$\cdot \left\{(r - r^{-1}) \left(\frac{br}{a}\right)^{k-\ell} I_{k-\ell}\left[\frac{abr}{1+r^2}\right] + \left[r\left(\frac{br}{a}\right)^{k-\ell+1} - r^{-1}\left(\frac{a}{br}\right)^{k-\ell+1}\right]\right.$$

$$\left. \times I_{k-\ell+1}\left[\frac{abr}{1+r^2}\right]\right\} \quad . \tag{C-28}$$

Letting $m = k - \ell + 1$, and noting that k and hence m can range from $-\infty$ to $+\infty$ without changing the value of Eq. (C-28) because of the benavior of $\binom{j-1}{k}$ as defined by Eq. (C-23),

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \sum_{m=-\infty}^{+\infty} \left\{\left[r\left(\frac{br}{a}\right)^m - r^{-1}\left(\frac{a}{br}\right)^m\right] I_m\left[\frac{abr}{1+r^2}\right]\right.$$

$$\left. + (r - r^{-1})\left(\frac{br}{a}\right)^{m-1} I_{m-1}\left[\frac{abr}{1+r^2}\right]\right\} \sum_{j=1}^{\nu} (r+r^{-1})^{-2j} \sum_{\ell=0}^{j-1} \binom{j-1}{\ell}\binom{j-1}{m+\ell-1} . \tag{C-29}$$

We now inquire about the final sum in Eq. (C-29). We have

$$\begin{aligned}
\sum_{q=0}^{2j-2} \binom{2j-2}{q} x^q &= (1+x)^{2j-2} \\
&= (1+x)^{j-1}(1+x)^{j-1} \\
&= \sum_{\ell=0}^{j-1}\sum_{n=0}^{j-1} \binom{j-1}{\ell}\binom{j-1}{n} x^{\ell+j-1-n} \quad .
\end{aligned} \tag{C-30}$$

Comparing powers of x, with $n = m + \ell - 1$, we deduce

$$\sum_{\ell=0}^{j-1} \binom{j-1}{\ell}\binom{j-1}{m+\ell-1} = \binom{2j-2}{j-m} \quad . \tag{C-31}$$

By Eq. (C-31) and the behavior of $\binom{2j-2}{j-m}$ as defined by Eq. (C-23), Eq. (C-29) now becomes

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \sum_{j=1}^{\nu} (r+r^{-1})^{-2j} \sum_{m=-j+2}^{j} \binom{2j-2}{j-m}$$

$$\cdot \left\{\left[r\left(\frac{br}{a}\right)^m - r^{-1}\left(\frac{a}{br}\right)^m\right] I_m\left[\frac{abr}{1+r^2}\right] + (r - r^{-1})\left(\frac{br}{a}\right)^{m-1} I_{m-1}\left[\frac{abr}{1+r^2}\right]\right\}$$

$$= P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \left\{\frac{\left(\frac{br^2}{a} - \frac{a}{br^2}\right) I_1\left[\frac{abr}{1+r^2}\right]}{(r+r^{-1})^2}\right.$$

$$+\frac{(r-r^{-1})}{(r+r^{-1})^2} I_o\left[\frac{abr}{1+r^2}\right] + \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \left(\sum_{m=-j+2}^{j} \binom{2j-2}{j-m} \right.$$

$$\left. \left. \cdot \left[r(\tfrac{br}{a})^m - r^{-1}(\tfrac{a}{br})^m\right] I_m\left[\frac{abr}{1+r^2}\right] + (r-r^{-1}) \sum_{n=-j+1}^{j-1} \binom{2j-2}{j-n-1} (\tfrac{br}{a})^n I_n\left[\frac{abr}{1+r^2}\right] \right) \right\}$$

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \left\{ \frac{\left(\frac{br^2}{a} - \frac{a}{br^2}\right) I_1[h]}{(r+r^{-1})^2} \right.$$

$$+ \frac{(r-r^{-1})}{(r+r^{-1})^2} I_o[h] + \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \binom{2j-2}{j} (r-r^{-1}) I_o[h]$$

$$+ \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \sum_{m=1}^{j} \left(\left[r\binom{2j-2}{j-m} - r^{-1}\binom{2j-2}{j+m} \right] (\tfrac{br}{a})^m \right.$$

$$\left. - \left[r^{-1}\binom{2j-2}{j-m} - r\binom{2j-2}{j+m} \right] (\tfrac{a}{br})^m \right) I_m[h]$$

$$+ (r-r^{-1}) \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \binom{2j-2}{j-1} I_o[h]$$

$$\left. + (r-r^{-1}) \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \sum_{n=1}^{j-1} \binom{2j-2}{j-n-1} \left[(\tfrac{br}{a})^n + (\tfrac{a}{br})^n\right] I_n[h] \right\} \qquad \text{(C-32)}$$

where we have used the fact that $I_m(z) = I_{-m}(z)$, and $h = abr/(1+r^2)$. Gathering terms,

$$P_\nu(a,b,r) = P_o(a,b,r) + \frac{\exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \left\{ (r-r^{-1}) I_o[h] \sum_{j=1}^{\nu} \frac{\binom{2j-1}{j}}{(r+r^{-1})^{2j}} \right.$$

$$+ (r+r^{-1})^{-2} \left(\frac{br^2}{a} - \frac{a}{br^2}\right) I_1[h]$$

$$+ (r-r^{-1}) \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \sum_{n=1}^{j-1} \binom{2j-2}{j-n-1} \left[(\tfrac{br}{a})^n + (\tfrac{a}{br})^n\right] I_n[h]$$

$$+ \sum_{j=2}^{\nu} (r+r^{-1})^{-2j} \sum_{m=1}^{j} \left(\left[r\binom{2j-2}{j-m} - r^{-1}\binom{2j-2}{j+m} \right] (\tfrac{br}{a})^m \right.$$

$$\left. \left. - \left[r^{-1}\binom{2j-2}{j-m} - r\binom{2j-2}{j+m} \right] (\tfrac{a}{br})^m \right) I_m[h] \right\}$$

$$= P_o(a,b,r) + \frac{\exp\left[-\dfrac{a^2r^2+b^2}{2(1+r^2)}\right]}{(r+r^{-1})} \Bigg\{ (r - r^{-1})\, I_o[h] \sum_{j=1}^{\nu} \frac{\binom{2j-1}{j}}{(r+r^{-1})^{2j}}$$

$$+ \sum_{j=1}^{\nu} (r+r^{-1})^{-2j} \sum_{m=1}^{j} \left(\left[r\binom{2j-1}{j-m} - r^{-1}\binom{2j-1}{j+m} \right] (\tfrac{br}{a})^m \right.$$

$$\left. - \left[r^{-1}\binom{2j-1}{j-m} - r\binom{2j-1}{j+m} \right] (\tfrac{a}{br})^m \right) I_m[h] \Bigg\} \quad . \tag{C-33}$$

Rearranging the double summation in Eq. (C-33) so that summing on m occurs first, we have, finally,

$$P_\nu(a,b,r) = P_o(a,b,r) + \exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right] \Bigg\{ (r - r^{-1})\, I_o\left[\frac{abr}{1+r^2}\right] \sum_{j=1}^{\nu} \frac{\binom{2j-1}{j}}{(r+r^{-1})^{2j+1}}$$

$$+ \sum_{m=1}^{\nu} I_m\left[\frac{abr}{1+r^2}\right] \left((\tfrac{br}{a})^m \sum_{j=m}^{\nu} \left[r\binom{2j-1}{j-m} - r^{-1}\binom{2j-1}{j+m} \right] (r+r^{-1})^{-2j-1} \right.$$

$$\left. - (\tfrac{a}{br})^m \sum_{j=m}^{\nu} \left[r^{-1}\binom{2j-1}{j-m} - r\binom{2j-1}{j+m} \right] (r+r^{-1})^{-2j-1} \right) \Bigg\}$$

$$= P_o(a,b,r) + \exp\left[-\frac{a^2r^2+b^2}{2(1+r^2)}\right] \Bigg\{ \frac{(r-r^{-1})}{2} \sum_{m=0}^{\nu} [(\tfrac{br}{a})^m + (\tfrac{a}{br})^m]\, I_m\left[\frac{abr}{1+r^2}\right]$$

$$\cdot \sum_{j=m+\delta_{mo}}^{\nu} \binom{2j-\delta_{mo}}{j+m} (r+r^{-1})^{-2j-1}$$

$$+ \sum_{m=1}^{\nu} m[(\tfrac{br}{a})^m - (\tfrac{a}{br})^m]\, I_m\left[\frac{abr}{1+r^2}\right] \sum_{j=m}^{\nu} \frac{(2j-1)!\,(r+r^{-1})^{-2j}}{(j+m)!\,(j-m)!} \Bigg\} \quad . \tag{C-34}$$

Equation (C-34) is valid for $\nu \geqslant 1$, $a \neq 0$, and $b \neq 0$, and δ_{mo} is the Kronecker delta-function, $\delta_{mo} = 1$ for $m = 0$, $\delta_{mo} = 0$ otherwise. When $r = 1$, Eq. (C-34) reduces to

$$P_\nu(a,b,1) = P_o(a,b,1) + \exp[-(a^2+b^2)/4]$$

$$\cdot \sum_{m=1}^{\nu} m[(\tfrac{b}{a})^m - (\tfrac{a}{b})^m]\, I_m(ab/2) \sum_{j=m}^{\nu} \frac{(2j-1)!\,2^{-2j}}{(j+m)!\,(j-m)!} \quad . \tag{C-35}$$

There is an alternate expression for the final summation in Eq. (C-35) that may sometimes be preferable. Using the identity

$$\sum_{j=m}^{\nu} \frac{(2j-1)!\,2^{-2j}}{(j+m)!\,(j-m)!} = \frac{(2m)^{-1}}{(\nu+m)!} \sum_{j=m}^{\nu} \frac{(\nu+j)!\,2^{-\nu-j}}{(j-m)!} \qquad (m \geqslant 1) \tag{C-36}$$

we obtain in lieu of Eq. (C-35)

$$P_\nu(a,b,1) = P_o(a,b,1) + 2^{-\nu-1} \exp[-(a^2+b^2)/4]$$

$$\cdot \sum_{m=1}^{\nu} \frac{1}{(\nu+m)!} [(\tfrac{b}{a})^m - (\tfrac{a}{b})^m] I_m(ab/2) \sum_{j=m}^{\nu} \frac{(\nu+j)!\, 2^{-j}}{(j-m)!} \quad . \tag{C-37}$$

To prove Eq. (C-36), it suffices to establish that it is true for $\nu = m$, which can be seen immediately because the summations of Eq. (C-36) are then just single terms, and that the difference in the right side between the summation for $\nu = M+1$ and that for $\nu = M$ is equal to a like difference in the left side. Taking these differences, we have to prove

$$\frac{(2M+1)!\, 2^{-2M-2}}{(M+1+m)!\,(M+1-m)!} = \frac{(2M+2)!\, 2^{-2M-3}}{m(M+1+m)!\,(M+1-m)!}$$

$$+ \frac{2^{-M-2}}{m} \sum_{j=m}^{M} \frac{(M+j)!\, 2^{-j}(j-M-2m-1)}{(M+1+m)!\,(j-m)!} \tag{C-38}$$

which becomes, multiplying Eq. (C-38) through by $2^{M+2} m(M+1+m)!$ and taking the difference between the single terms

$$\sum_{j=m}^{M} \frac{(M+j)!\, 2^{-j}(M+1+2m-j)}{(j-m)!} = \frac{(2M+1)!\, 2^{-M}}{(M-m)!} \quad . \tag{C-39}$$

The left member of Eq. (C-39) may be written

$$\sum_{j=m}^{M} \frac{(M+j)!\, 2^{-j}}{(j-m)!} [(M+j+1) - 2(j-m)] = \sum_{j=m}^{M} \frac{(M+j+1)!}{(j-m)!} 2^{-j}$$

$$- \sum_{j=m+1}^{M} \frac{(M+j)!\, 2^{-j+1}}{(j-m-1)!} \tag{C-40}$$

and changing the subscript j to $j' = j-1$ in the second summation of the right-hand side of Eq. (C-40), we obtain a sum identical to the first sum on the right-hand side, except that it reaches only to $M-1$ rather than M. Thus Eq. (C-40) is equal to $(M+j+1)!\, 2^{-j}/(j-m)!$ for $j = M$, and Eq. (C-39) is thereby proved, substantiating Eq. (C-36).

Although Eqs. (C-35) and (C-37) are general formulas, it may on occasion be preferable to express the solution in terms of $P_o(a,b,1)$ and a single pair of Bessel functions, by reducing the sum of Bessel functions through Bessel-function recursion relations.[36] We have done this for ν of 1 through 5, inclusive, and find that the minimum number of algebraic terms is obtained when reduction is made in terms of $I_\nu(ab/2)$ and $I_{\nu-1}(ab/2)$. However, the number of algebraic terms appears to grow like $1 + [\nu(\nu-1)]/2$, so that it is difficult to go much beyond $\nu = 5$:

$$P_1(a,b,1) = P_o(a,b,1) + 2^{-4}(b^2-a^2) \exp[-(a^2+b^2)/4]\; q^{-1} I_1(q) \tag{C-41}$$

$$P_2(a,b,1) = P_o(a,b,1) + 2^{-7}(b^2 - a^2)\exp[-(a^2+b^2)/4]$$

$$\cdot\ [12q^{-1} I_1(q) + q^{-2} I_2(q)\ (s)] \tag{C-42}$$

$$P_3(a,b,1) = P_o(a,b,1) + 2^{-10}(b^2 - a^2)\exp[-(a^2+b^2)/4]$$

$$\cdot\ [16q^{-2} I_2(q)\ (29 + s) + q^{-3} I_3(q)\ (28p + s^2)] \tag{C-43}$$

$$P_4(a,b,1) = P_o(a,b,1) + 2^{-13}(b^2 - a^2)\exp[-(a^2+b^2)/4]$$

$$\cdot\ [4q^{-3} I_3(q)\ (195 \cdot 2^5 + 69 \cdot 2^2 s + 60p + 5s^2)$$

$$+ q^{-4} I_4(q)\ (65 \cdot 2^4 p + 44ps + s^3)] \tag{C-44}$$

$$P_5(a,b,1) = P_o(a,b,1) + 2^{-16}(b^2 - a^2)\exp[-(a^2+b^2)/4]$$

$$\cdot\ [8q^{-4} I_4(q)\ (843 \cdot 2^8 + 87 \cdot 2^7 s + 3104p + 268s^2 + 52ps + 3s^3)$$

$$+ q^{-5} I_5(q)\ (843 \cdot 2^6 p + 87 \cdot 2^5 ps + 64ps^2 + 496p^2 + s^4)] \tag{C-45}$$

where $p = a^2b^2$, $q = ab/2$, and $s = a^2 + b^2$.

IV. SIMPLE DERIVATION OF $P_o(a,b,r)$

For completeness we present here a short derivation* of the relation between $P_o(a,b,r)$ and Marcum's Q-function, which is quoted in Eqs. (36) and (37) and which can be obtained from Maximon's result [Eqs. (C-2) and (C-3)] and the second series of Eq. (C-4) of this appendix. The following derivation given here appears to be considerably less complicated than that employed by Maximon.

Referring to Eq. (C-6) of this appendix and to Eqs. (28) and (32) of the text (generalized from $r = 1$), with $\alpha = a^2$, $\beta = b^2$,

$$P_o(\sqrt{\alpha},\sqrt{\beta},r) = \frac{1}{4}\int_o^\infty \exp[-(X+\alpha)/2]\, I_o(\sqrt{X\alpha})\, dX \int_o^{Xr^2} \exp[-(Y+\beta)/2]\, I_o(\sqrt{Y\beta})\, dY \tag{C-46}$$

The Laplace transform of $P_o(\sqrt{\alpha},\sqrt{\beta},r)$ with respect to β is

$$L(\alpha,r,s) = \int_o^\infty P_o(\sqrt{\alpha},\sqrt{\beta},r)\, e^{-\beta s}\, d\beta$$

$$= \frac{1}{4}\int_o^\infty \exp[-(X+\alpha)/2]\, I_o(\sqrt{X\alpha}\, dX \int_o^{Xr^2} \exp[-Y/2]\, dY \int_o^\infty \exp[-\beta(s+\tfrac{1}{2}]\ I_o(\sqrt{Y\beta})\, d\beta$$

$$= \frac{1}{4}\int_o^\infty \exp[-(X+\alpha)/2]\, I_o(\sqrt{X\alpha})\, dX \int_o^{Xr^2} \frac{\exp[-Y/2 + Y/(4s+2)]}{(s+\frac{1}{2})}\, dY$$

---

* The Laplace transform method given here was apparently anticipated by J. E. Storer in unpublished work performed in 1960. (See the remarks of Jones[33] which introduce his Appendix B.)

$$= \frac{1}{2s} \int_0^\infty e^{-(X+\alpha)/2}\, I_o(\sqrt{X\alpha}) \left[1 - \exp[-sXr^2/(1+2s)]\right] dX$$

$$L(\alpha, r, s) = \frac{1}{s} - \frac{\exp\left\{-\frac{\alpha r^2}{2(1+r^2)}\right\}}{(1+r^2)} \cdot \frac{\exp\left\{\frac{\alpha r^2}{4(r^2+1)^2 [s + 2^{-1}(1+r^2)^{-1}]}\right\}}{s + 2^{-1}(1+r^2)^{-1}}$$

$$- \frac{\exp-\left\{\frac{\alpha r^2}{2(1+r^2)}\right\}}{2(1+r^2)} \cdot \frac{\exp\left\{\frac{\alpha r^2}{4(r^2+1)^2[s+2^{-1}(1+r^2)^{-1}]}\right\}}{s[s+2^{-1}(1+r^2)^{-1}]} \tag{C-47}$$

where we have used the Laplace transform formula[37]

$$\int_0^\infty \exp[-\gamma U]\, I_o\,(2\sqrt{\eta U})\, dU = \gamma^{-1} \exp(\eta/\gamma) \quad . \tag{C-48}$$

Examining Eq. (C-47), we see that it is the Laplace transform of a sum of three terms; the first is unity, the second is, by Eq. (C-48),

$$- \frac{\exp-\left\{\frac{\alpha r^2}{2(1+r^2)}\right\}}{(1+r^2)} \cdot \exp\left\{-\frac{\beta}{2(1+r^2)}\right\} I_o\,[\sqrt{\alpha\beta}\, r/(1+r^2)]$$

and the third is

$$- \frac{\exp\left\{-\frac{\alpha r^2}{2(1+r^2)}\right\}}{2(1+r^2)} \cdot \int_0^\beta \exp\left\{-\frac{\beta}{2(1+r^2)}\right\} I_o\,[\sqrt{\alpha\beta}\, r/(1+r^2)]\, d\beta \quad . \tag{C-49}$$

If the upper limit of the integral in Eq. (C-49) were infinity, the entire expression would be unity, again by Eq. (C-48), so that we conclude

$$P_o(\sqrt{\alpha}, \sqrt{\beta}, r) = \frac{1}{2(1+r^2)} \int_\beta^\infty \exp\left[-\frac{\alpha r^2 + \beta}{2(1+r^2)}\right] I_o[\sqrt{\alpha\beta}\, r/(1+r^2)]\, d\beta$$

$$- \frac{1}{(1+r^2)} \exp\left[-\frac{\alpha r^2 + \beta}{2(1+r^2)}\right] I_o[\sqrt{\alpha\beta}\, r/(1+r^2)]\, d\beta$$

$$= \int_{b/\sqrt{1+r^2}}^\infty (b/\sqrt{1+r^2}) \exp\left\{-\frac{(ar/\sqrt{1+r^2})^2 + (b/\sqrt{1+r^2})^2}{2}\right\}$$

$$I_o\{(ar/\sqrt{1+r^2})\, (b/\sqrt{1+r^2})\}\, d(b/\sqrt{1+r^2})$$

$$- \frac{1}{(1+r^2)} \exp\left[-\frac{a^2 r^2 + b^2}{2(1+r^2)}\right] I_o\left[\frac{abr}{1+r^2}\right] \quad .$$

Therefore,

$$P_o(a, b, r) = Q\left(\frac{ar}{\sqrt{1 + r^2}}, \frac{b}{\sqrt{1 + r^2}}\right) - \frac{1}{(1 + r^2)} \exp-\left[\frac{a^2r^2 + b^2}{2(1 + r^2)}\right] I_o\left[\frac{abr}{1 + r^2}\right] \tag{C-50}$$

where we have returned to the parameters $a = \sqrt{\alpha}$, $b = \sqrt{\beta}$ and have referred to Eq. (C-4). Maximon's result [Eqs. (C-2) and (C-3)] for $\nu = 0$ is checked by substituting the second series of Eq. (C-4) in Eq. (C-50).

REFERENCES

. . .

14. C.W. Helstrom, Statistical Theory of Signal Detection (Pergamon Press, New York, 1960), op. cit., p. 174, with f = 1.

. . .

23. C.W. Helstrom, "The Resolution of Signals in White, Gaussian Noise," Proc. IRE, 43, 1111-1118 (September 1955).

. . .

29. L.C. Maximon, "On the Representation of Indefinite Integrals Containing Bessel Functions by Simple Neumann Series," Proc. Amer. Math. Soc. 7, 1054-1062 (1956).
30. S.O. Rice, "Mathematical Analysis of Random Noise," Bell System Tech. J. 23, 282-332 (July 1944) and 24, 46-156 (January 1945), Sec. 3.10. Reprinted in N. Wax, Selected Papers on Noise and Stochastic Processes (Dover Publications, New York, 1954).
31. G.N. Watson, A Treatise on the Theory of Bessel Functions (MacMillan, New York, 1948), Sec. 16.5.
32. J.I. Marcum, "A Statistical Theory of Target Detection by Pulsed Radar—Mathematical Appendix," Report RM-753, The RAND Corporation, Santa Monica, California (1 July 1948), and Trans. IRE, PGIT IT-6, 159 (April 1960).
33. J.J. Jones, "The Probability of Error for a Four-Tone FSK Demodulator in the Presence of Crosstalk or Interference," Engineering Note 290(R), Sylvania Applied Research Laboratory, Waltham, Massachusetts (7 February 1962).
34. G.N. Watson, op. cit., p.79, Eq. (5).
35. ———, p.79, Eq. (6).
36. ———, p.79.
37. A. Erdélyi, et al., Tables of Integral Transforms, Vol. I, Bateman Manuscript Project, California Institute of Technology (McGraw-Hill, New York, 1954), p.197, Eq. (14).

For convenience in ordering copies of Lincoln Laboratory reports cited in this document, each reference is followed by its ASTIA number. In addition, Unclassified (released) reports have also been assigned Hayden serials (designated H-), indicating that they are obtainable, at cost, as microfilm or photoprint copies from the Micro-Reproduction Service, Hayden Memorial Library, M.I.T., Cambridge 39, Massachusetts.

Appendix D

COMMENTS AND ERRATA FOR APPENDIX C

Appendix C has been reproduced in its entirety from Price [7]. It contains the necessary derivation to complete the analysis begun in the present report's text of probability of error with no smoothing of the data.

The mathematics in Appendix C was checked and found correct with the following exceptions:

- In the equation which follows Eq. (C-13) the $\exp[-(a^2 + b^2)/2]$ factor in the second line should be multiplying the entire right-hand side of the equation rather than just the second and third lines as shown.

- In Eq. (C-21) the $\exp[-(\alpha + \beta)]$ factor in the first line should be $\exp[-(\alpha + \beta)/2]$.

In neither of these cases does the error propagate to the equations which follow.