Can Detectability Be Improved by Adding Noise?

Steven Kay, *Fellow, IEEE*

***Abstract*—It is shown that under certain conditions the performance of a suboptimal detector may be improved by adding noise to the received data. The reasons for this counterintuitive result are explained and a computer simulation example given.**

***Index Terms*—Decision making, Gaussian noise, signal detection.**

I. Introduction

IT IS A TENET of detection theory that the performance of a detector increases as the signal-to-noise ratio (SNR) increases. Somewhat paradoxically, however, under certain conditions detectability can be improved by adding *independent* noise to the received data, in effect decreasing the SNR! Before describing a situation in which this phenomenon occurs we first discuss in general the use of additional noise samples as a means for improving detectability.

To simplify the discussion we consider the illustrative example of detection of a dc level embedded in independent and identically distributed (i.i.d.) noise. Formally, we wish to choose between the two hypotheses

$$\begin{aligned}\mathcal{H}_0&: x[n] = w[n]\\ \mathcal{H}_1&: x[n] = A + w[n]\end{aligned} \tag{1}$$

for $n = 0, 1, \cdots, N-1$, where the dc level A is known and $A > 0$, and $w[n]$ is i.i.d. noise. Now assume that we have available the additional noise samples $\{u[0], u[1], \cdots, u[N-1]\}$. It is well known that if the additional noise samples are statistically dependent on $w[n]$ and/or the probability density function (pdf) of $u[n]$ depends upon whether $\mathcal{H}_0$ or $\mathcal{H}_1$ is true, then knowledge of $u[n]$ will improve signal detectability. In fact, if either of these conditions hold, the optimal likelihood ratio test (LRT) will depend upon $u[n]$. However, if the conditional PDF of $\boldsymbol{u} = [u[0]\, u[1] \cdots u[N-1]]^T$ given the data $\boldsymbol{x} = [x[0]\, x[1] \cdots x[N-1]]^T$ does not depend upon which hypothesis is true or

$$p(\boldsymbol{u}|\boldsymbol{x}; \mathcal{H}_0) = p(\boldsymbol{u}|\boldsymbol{x}; \mathcal{H}_1) \tag{2}$$

then by the theorem of irrelevance [4], we can ignore $\boldsymbol{u}$. This is because it is irrelevant to any decision between $\mathcal{H}_0$ and $\mathcal{H}_1$. It is easily shown that the LRT in this case will not depend upon $\boldsymbol{u}$ [2]. Alternatively, for the detection problem of (1) $\boldsymbol{x}$ is a sufficient statistic. Therefore, we can restrict attention to the class of detectors that depend on $\boldsymbol{x}$, without any loss in performance [3]. That $\boldsymbol{x}$ is a sufficient statistic follows from the observation that (2) can be rewritten as

$$p(\boldsymbol{u}|\boldsymbol{x}; A = 0) = p(\boldsymbol{u}|\boldsymbol{x}; A > 0)$$

or once $\boldsymbol{x}$ is known, the PDF of $\boldsymbol{u}$ does not depend on A. The key point here is that we can ignore $\boldsymbol{u}$ *in determining the optimal detector in the Neyman–Pearson sense*. In the case of a *suboptimal* detector, however, the use of $\boldsymbol{u}$ may still improve detectability, as we show in the next section.

II. An Example

For the problem of (1) consider the detector that decides $\mathcal{H}_1$ if

$$T(\boldsymbol{x}) = \sum_{n=0}^{N-1} \operatorname{sgn}(x[n]) > \gamma_x.$$

This detector is suboptimal (unless $w[n]$ is Laplacian and the signal A is close to zero [2]). Assume that $w[n]$ is i.i.d. with the Gaussian mixture pdf

$$p(w) = \tfrac{1}{2}\,\phi(w; \mu, 1) + \tfrac{1}{2}\phi(w; -\mu, 1)$$

where

$$\phi(w; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(w-\mu)^2\right].$$

The optimal LRT for this problem is easily shown to decide $\mathcal{H}_1$ if

$$L(\boldsymbol{x}) = \sum_{n=0}^{N-1} \ln\left[\exp(-A^2/2)\exp(Ax[n]) \frac{\cosh(\mu(x[n]-A))}{\cosh(\mu x[n])}\right] > \gamma \tag{3}$$

and for a given probability of false alarm P_{FA} and hence threshold γ, it will produce the maximum probability of detection P_D. Now assume that we have a realization of white Gaussian noise (WGN) with variance σ^2, which is *independent* of $w[n]$. Calling this $u[n]$, we form $y[n] = x[n] + u[n]$ and define the vector of data-plus-noise samples as $\mathbf{y} = [y[0]\; y[1] \cdots y[N-1]]^T$. We now show that the detector that decides $\mathcal{H}_1$ if

$$T(\boldsymbol{y}) = \sum_{n=0}^{N-1} \operatorname{sgn}(y[n]) > \gamma_y$$

can have a higher P_D (for a fixed P_{FA}) than $T(\boldsymbol{x})$. Hence, *adding WGN to the data improves detectability*. Note that similar investigations have been carried out under the name of *nondynamical stochastic resonance* [1]. However, in these studies the figure of merit is SNR, which may or may not imply higher

Manuscript received May 20, 1999. This work was supported by Code 321SI of the Office of Naval Research and administered by D. Johnson. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. A. Nehorai.

The author is with the Department of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI 02881 USA (e-mail: kay@ele.uri.edu).

Publisher Item Identifier S 1070-9908(00)00273-X.

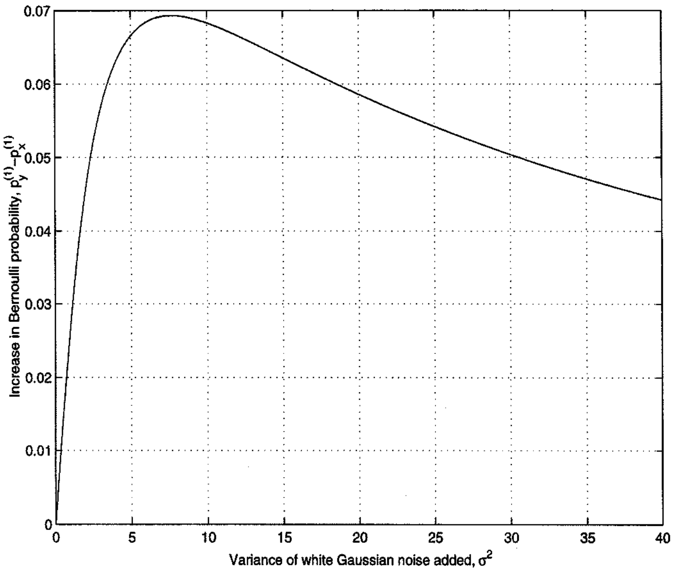

Fig. 1. Difference in Bernoulli probabilities.

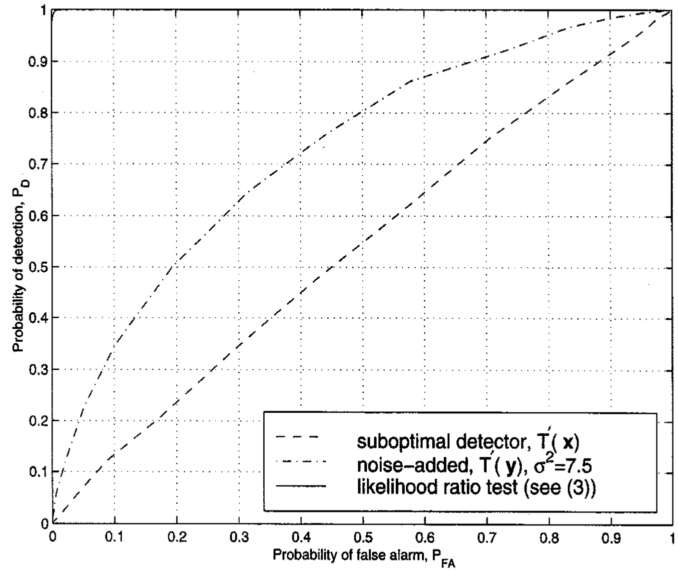

Fig. 2. Receiver operating characteristics.

detectability for a higher SNR. In our example, the detectability actually *decreases* as the SNR *increases*.

The analysis of detection performance is simplified if we consider the equivalent detector

$$T'(\boldsymbol{x}) = \sum_{n=0}^{N-1} \left(\tfrac{1}{2} + \tfrac{1}{2}\mathrm{sgn}(x[n])\right) = \sum_{n=0}^{N-1} \zeta_x[n] > \gamma_x'$$

and similarly for $T(\boldsymbol{y})$. Note that $\zeta_x[n]$ is a Bernoulli random variable, taking on values zero and one with Bernoulli probabilities, and therefore $T'(\boldsymbol{x})$ is binomially distributed. The Bernoulli probabilities depend on the hypothesis $\mathcal{H}_i$, either $\mathcal{H}_0$ or $\mathcal{H}_1$, and so are denoted as $p_x^{(i)} = \Pr\{\zeta_x[n] = 1;\, \mathcal{H}_i\}$. Likewise, we have

$$T'(\boldsymbol{y}) = \sum_{n=0}^{N-1} \left(\tfrac{1}{2} + \tfrac{1}{2}\mathrm{sgn}(y[n])\right) = \sum_{n=0}^{N-1} \zeta_x[n] > \gamma_y'$$

and $T'(\boldsymbol{y})$ is binomially distributed with $p_y^{(i)} = \Pr\{\zeta_x[n] = 1;\, \mathcal{H}_i\}$. The detection performance is easily shown to be monotonically increasing with $p_x^{(1)}$, $p_y^{(1)}$, respectively, so that we restrict our attention to showing that $p_y^{(1)} > p_x^{(1)}$. To constrain P_{FA} for the two detectors we choose $\gamma_x' = \gamma_y' = 0$. Then, under $\mathcal{H}_0$ since the PDF of $x[n] = w[n]$ is even, we have that $p_x^{(0)} = \Pr\{\zeta_x[n] = 1;\ \mathcal{H}_0\} = 1/2$. Similarly, since the PDF of $y[n] = w[n] + u[n]$ is also even ($w[n]$ and $u[n]$ have even PDF's which after convolution produces an even PDF), $p_y^{(0)} = \Pr\{\zeta_y[n] = 1;\ \mathcal{H}_0\} = 1/2$. Thus, if we choose $\gamma_x' = \gamma_y' = 0$, the PDF's for $T'(\boldsymbol{x})$ and $T'(\boldsymbol{y})$ will be the same under $\mathcal{H}_0$ and the P_{FA}'s will be identical. To show that $p_y^{(1)} > p_x^{(1)}$ we note that

$$\begin{aligned} p_x^{(1)} &= \Pr\{\zeta_x[n] = 1;\, \mathcal{H}_1\} = \Pr\{A + w[n] > 0\} \\ &= \Pr\{w[n] > -A\} = \tfrac{1}{2}\,Q(-A-\mu) + \tfrac{1}{2}\,Q(-A+\mu) \end{aligned}$$

where $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) \exp(-t^2/2)\,dt$. Also, $p_y^{(1)} = \Pr\{w[n]+u[n] > -A\}$. But $z[n] = w[n]+u[n]$ has a Gaussian

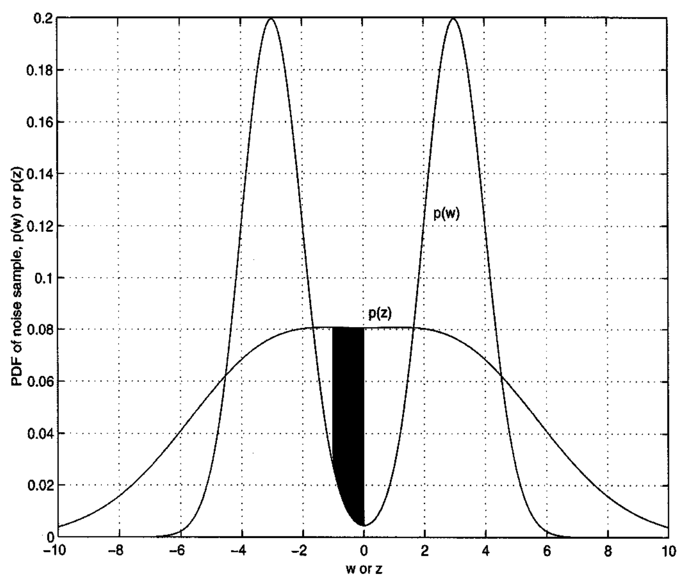

Fig. 3. Probability density functions for original $(p(w))$ and transformed noise $(p(z))$.

mixture PDF of $1/2\phi(z;\, \mu,\, 1+\sigma^2) + 1/2\phi(z;\, -\mu,\, 1+\sigma^2)$ so that

$$p_y^{(1)} = \frac{1}{2}\,Q\left(\frac{-A-\mu}{\sqrt{1+\sigma^2}}\right) + \frac{1}{2}\,Q\left(\frac{-A+\mu}{\sqrt{1+\sigma^2}}\right).$$

As an example, for $A = 1$, $\mu = 3$ we plot $p_y^{(1)} - p_x^{(1)}$ versus σ^2, the variance of the added noise, in Fig. 1. Note that $p_y^{(1)} > p_x^{(1)}$ and the difference appears to be maximized at $\sigma^2 = 7.5$. Hence, P_D *can be improved by adding WGN to* $x[n]$. The actual detection performance can be obtained using these values of $p^{(1)}$ in the binomial distribution. However, to verify the conclusions we performed a Monte Carlo computer simulation for $N = 30$. The receiver operating characteristics (ROC's) are shown in Fig. 2. As claimed the addition of noise improves detection. Of course, the *optimal* LRT of (3) outperforms both suboptimal detectors, its performance nearly perfect ($P_D \approx 1$ for all P_{FA}'s as seen in Fig. 2). Note that the optimal detector does not utilize the $u[n]$ noise samples in its statistic. This is because $u[n]$ is independent of $w[n]$ and the PDF of $u[n]$ is the same under either hypothesis. Hence, the conditional pdf of $\boldsymbol{u}$ satisfies (2) and the additional noise samples provide no discrimination.

Finally, the question arises as to where the gain in performance came from. The answer is that by adding WGN we have effectively changed the noise pdf from one with very little mass at zero to one with more mass as shown in Fig. 3. The increase in $p^{(1)}$ is just the area shown. This is because

$$\begin{aligned} p^{(1)} &= \Pr\{\zeta_y[n] = 1; \mathcal{H}_1\} \quad \text{or} \quad \Pr\{\zeta_x[n] = 1; \mathcal{H}_1\} \\ &= \Pr\{A + z[n] > 0\} \quad \text{or} \quad \Pr\{A + w[n] > 0\} \\ &= \Pr\{z[n] > -A\} \quad \text{or} \quad \Pr\{w[n] > -A\}. \end{aligned}$$

Since the pdf's of $z[n]$ and $w[n]$ are even and thus satisfy $\Pr\{z[n] > 0\} = \Pr\{w[n] > 0\} = 1/2$, the increase in $p^{(1)}$ is $\Pr\{-A < z[n] < 0\} - \Pr\{-A < w[n] < 0\}$ as shown by the shaded area in Fig. 3. This is about 0.07 which is consistent with Fig. 1. *In effect, the loss in detectability incurred by reducing the SNR is more than offset by the increased sensitivity of the new noise pdf near the origin.* This is because as the pdf is shifted to the right due to the presence of a signal, the probability of a threshold crossing increases more dramatically for $p(z)$ than for $p(w)$.

REFERENCES

[1] Z. Gingl, L. B. Kiss, and F. Moss, "Non-dynamical stochastic resonance: Theory and experiments with white and arbitrarily colored noise," *Europhys. Lett.*, pp. 191–196, Jan. 1995.

[2] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[3] M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Vol. II*, New York: Macmillan, 1977.

[4] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, New York: Wiley, 1965.