Waveform Design for Multistatic Radar Detection

Steven Kay*
Dept. of Electrical, Computer, and Biomedical Engineering
University of Rhode Island
Kingston, RI 02881
401-874-5804 (voice) 401-782-6422 (fax)
kay@ele.uri.edu
Keywords: signal detection, radar, sonar

July 9, 2007

Abstract

We derive the optimal Neyman-Pearson detector and it performance, and then present a methodology for the design of the transmit signal for a multistatic radar receiver. The detector assumes a Swerling-I extended target model as well as signal-dependent noise, i.e., clutter. It is shown that the Neyman-Pearson detection performance does not immediately lead to an obvious signal design criterion so that as an alternative, a divergence criterion is proposed for signal design. A simple method for maximizing the divergence, termed the maximum marginal allocation algorithm, is presented and is guaranteed to find the global maximum. The overall approach is a generalization of previous work that determined the optimal detector and transmit signal for a monostatic radar.

1 Introduction

Signal design for optimal detection in signal-dependent noise has been a problem of long-standing interest. In particular, the fields of radar and sonar have seen much work in this area. Some of the salient references are listed in [1–9]. Signal-dependent noise is generally referred to as clutter in radar and reverberation in active sonar. In either case, the fact that the received noise characteristics are dependent on the transmitted signal greatly complicates the signal design. For the case of signal design in wide sense stationary (WSS) colored noise whose power spectral density (PSD) *does not* depend on the transmitted signal, the solution

*This work was supported by the Air Force Research Lab, Rome, NY and administered by Michigan Tech Research Institute under contract FA8750-05-C-0237.

is well known. It says to place all the signal energy into the frequency band for which the noise power is minimum. For the detection of a discrete signal vector in correlated noise with a given covariance matrix one should choose the signal as the eigenvector of the noise covariance matrix whose eigenvalue is minimum [10]. However, it was shown in [11] that this latter design approach is no longer viable when the noise covariance matrix depends on the transmitted signal, as for clutter. In that work the optimal transmit signal in the presence of signal-dependent noise was derived. In this paper we extend those results to the multistatic case.

The design of transmit waveforms for a multistatic radar is a topic of current interest. We henceforth assume a single transmitter and multiple receivers. The use of multiple transmitters is an obvious extension and falls under the category of MIMO radar. From a detectability standpoint, however, multiple transmitters may not be desirable unless the available frequency spectrum can accommodate transmit signals with nonoverlapping frequency bands. Otherwise, "crosstalk", a common problem in MIMO communication systems [12], will significantly raise the noise floor. For this paper we therefore address only the single transmitter scenario. The use of a Neyman-Pearson (NP) detector, although not viable for a practical implementation, can be viewed as providing an upper bound on the performance of any multistatic radar system since all statistical quantities are assumed known. It also lends insight into the signal design problem. It should be noted that when, for example, some parameters are unknown, then either a generalized likelihood ratio test (GLRT) or Rao test can be used to accommodate the unknown parameters [10]. This will lead to an actual system implementation. Some preliminary results to this effect rely on the complex linear model and can be found in [10,13].

Other attempts to design an optimal transmit signal in the multistatic case can be found in [14], where an iterative algorithm is applied. The difficulty with this approach is that the algorithm is not guaranteed to converge or may converge to only a local maximum. Other approaches can be found in [15,16] in which mutual information is used as the design criterion. It is not clear, however, how this criterion is related to detection performance and the reported results do not make this connection, if any, clear. Our approach attempts to tie the detection performance to the signal design.

It should be mentioned that the use of signal diversity for better system performance is well known in the communication literature and has been for some 40 years. The use of signal diversity and the optimal receiver structure for time diversity can be found in [17]. As must be the case, since the likelihood ratio forms the basis for the optimal receiver, the results there are similar to our NP formulation results.

In this paper we first define the problem and state the modeling assumptions in Section 2, followed by the optimal detector in Section 3 and its performance in Section 4. In Sections 5 and 6 the main results of the paper are given, which are the design of the signal with the optimal energy spectral density. Examples are next given in Section 7 and finally conclusions are given in Section 8.

2 Problem Statement and Modeling Assumptions

It is assumed that we have M receive sensors, physically separated so that all received clutter and noises are statistically independent from sensor to sensor. The model for the received waveforms is shown in Figure 1.

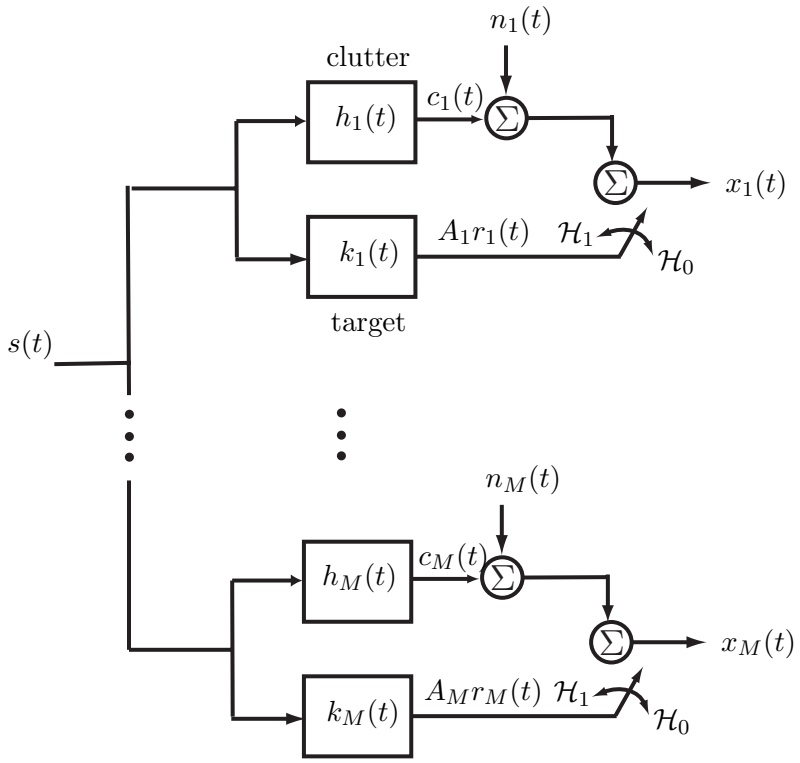

Figure 1: Modeling of received complex envelope waveforms. $s(t)$ is the transmitted signal, $h_i(t)$ is the impulse response of the random linear time invariant (LTI) clutter channel filter to the ith sensor, $k_i(t)$ is the impulse response of the random LTI target filter to the ith sensor, and $n_i(t)$ represents ambient noise and interference at the ith sensor.

We assume that the received waveforms are the complex envelopes of the real bandpass data, and are denoted by $x_i(t)$ for $|t| \leq T/2$ at the ith sensor. When no target return is present, i.e., under hypothesis $\mathcal{H}_0$, we have that $x_i(t) = c_i(t) + n_i(t)$, where $c_i(t)$ denotes clutter and $n_i(t)$ is the sum of ambient noise and interference, i.e., jamming. Under the hypothesis $\mathcal{H}_1$, the target return at the ith sensor is modeled as $s(t) \star k_i(t) = A_i s(t) \star g_i(t)$, where $\star$ denotes convolution. Hence, $s(t)$ is the complex envelope of the transmitted signal, A_i is a complex reflection factor with the probability density function (PDF) $A_i \sim \mathcal{CN}(0, \sigma^2_{A_i})$ (a Swerling-I target-type model), and $g_i(t)$ is the deterministic part of the target response, which models the impulse response of the extended target. The designation $\mathcal{CN}$ means *complex*

normal or Gaussian. We have assumed a *zero Doppler target*, which represents a worst case scenario. It is felt that if we can make progress on this signal design problem, then the nonzero Doppler target should yield improved performance as well. Note that it is only the optimality of the transmit signal that is in question for nonzero Doppler targets. The proposed approach can easily be modified to address the nonzero Doppler target but of course will require separate Doppler channels. Also, $n_i(t)$ is modeled as a complex WSS Gaussian random process with zero mean and PSD $P_{n_i}(F)$. The noises at the sensors as well as the target reflection amplitudes are assumed to be independent of each other. The baseband frequency band is assumed to be $-W/2 \leq F \leq W/2$ and hence all PSDs are defined over this band. Finally, we model the clutter return $c_i(t)$ as the output of a *random LTI filter* with impulse response $h_i(t)$, whose input is the transmitted signal. This is the model used in [18,19,20]. This type of modeling is appropriate for multipath [8] since the filtering will model the altered frequency spectrum of the return signal. (Note, however, that in [8] the statistical characteristics of the filter are different. There the *uncorrelated scattering model* is used, whereby each point of the impulse response is uncorrelated with any other point and the variance varies from point to point.) However, Doppler spreading due to clutter motion and/or platform motion is not accommodated. To model the latter the more usual model is a *convolution in frequency*, which yields frequency spreading, as opposed to a multiplication. We do not pursue this further.

Continuing with the clutter modeling, if $s(t)$ is the transmit signal, then the clutter return will be $c_i(t) = s(t) \star h_i(t)$ at the receiver. All clutter returns are assumed independent from sensor to sensor. By reversing the convolution we can write this as $c_i(t) = h_i(t) \star s(t)$, where now the filter input is $h_i(t)$ and the filter impulse response is $s(t)$. If we now assume that $h_i(t)$ is a complex WSS Gaussian random process with zero mean and PSD $P_{h_i}(F)$, then $c_i(t)$ will also be a complex WSS Gaussian random process [21,22] with zero mean and PSD $P_{c_i}(F) = T|S(F)|^2 P_{h_i}(F)$, where $S(F)$ is the normalized Fourier transform (the usual Fourier transform multiplied by $1/\sqrt{T}$) of $s(t)$.

3 The Neyman-Pearson Detector

Next we describe the Neyman-Pearson (NP) detector which assumes that all statistical quantities are known. In practice, these may need to be estimated and hence implementations based on the GLRT or Rao test will need to be derived. The detection problem can be summarized as

$$\begin{aligned} \mathcal{H}_0 &: x_i(t) = c_i(t) + n_i(t) \\ \mathcal{H}_1 &: x_i(t) = A_i r_i(t) + c_i(t) + n_i(t) \end{aligned} \tag{1}$$

for $i = 1, 2, \ldots, M$ and $|t| \leq T/2$, and $r_i(t) = g_i(t) \star s(t)$. The assumptions on the signal, clutter, and noises have already been set forth. Based on these assumptions it is shown in Appendix A that the NP

detector, which is implemented in the frequency domain, decides a signal is present if

$$T(\mathbf{X}) = 2\sum_{i=1}^{M} \frac{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \frac{\left|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i\right|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} > \gamma \tag{2}$$

where

$$\begin{aligned} \mathbf{X}_i &= \left[\begin{array}{ccc} X_i(F_{-N/2}) & \ldots & X_i(F_{N/2}) \end{array} \right]^T \\ \mathbf{R}_i &= \left[\begin{array}{ccc} R_i(F_{-N/2}) & \ldots & R_i(F_{N/2}) \end{array} \right]^T \end{aligned}$$

are vectors of dimension $(N+1) \times 1$ containing Fourier transform samples. The "two factor" multiplying the sum is included to produce a recognizable PDF for $T(\mathbf{X})$. Also we define the vector of all the sensor outputs in the frequency domain as

$$\mathbf{X} = \left[\begin{array}{ccc} \mathbf{X}_1^T & \ldots & \mathbf{X}_M^T \end{array} \right]^T.$$

The Fourier transforms are defined as

$$\begin{aligned} X_i(F) &= \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x_i(t) \exp(-j2\pi Ft) dt \\ R_i(F) &= \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} r_i(t) \exp(-j2\pi Ft) dt \end{aligned} \tag{3}$$

and $\mathbf{K}_i$ is the covariance matrix of $\mathbf{C}_i + \mathbf{N}_i$ where

$$\begin{aligned} \mathbf{C}_i &= \left[\begin{array}{ccc} C_i(F_{-N/2}) & \ldots & C_i(F_{N/2}) \end{array} \right]^T \\ \mathbf{N}_i &= \left[\begin{array}{ccc} N_i(F_{-N/2}) & \ldots & N_i(F_{N/2}) \end{array} \right]^T. \end{aligned}$$

In effect, the detection statistic is just the sum of the statistics that would be used for a single sensor. This is obviously due to the assumption of independence of the signal return, clutter, and noise between sensors. For a time-bandwidth product $WT > 16$ [24], the detection statistic of (2) can be approximately written as (see Appendix A)

$$T(\mathbf{X}) = 2\sum_{i=1}^{M} \alpha_i \frac{\left|\sum_{n=-N/2}^{N/2} \frac{X_i^*(F_n) R_i(F_n)}{P_{c_i}(F_n) + P_{n_i}(F_n)}\right|^2}{\sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n) + P_{n_i}(F_n)}} \tag{4}$$

where $R_i(F_n) = S(F_n) G_i(F_n) \sqrt{T}$, $P_{c_i}(F_n) = T|S(F_n)|^2 P_{h_i}(F_n) = \mathcal{E}_s(F_n) P_{h_i}(F_n)$, and

$$\alpha_i = \frac{\sigma_{A_i}^2 \sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n) + P_{n_i}(F_n)}}{1 + \sigma_{A_i}^2 \sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n) + P_{n_i}(F_n)}}.$$

We next examine the meaning of (2). Clearly, the weighting term given by

$$\frac{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}$$

serves to include the contributions of sensors that exhibit a large target return. If $\sigma_{A_k}^2 \mathbf{R}_k^H \mathbf{K}_k^{-1} \mathbf{R}_k << 1$, then the contribution of sensor k will not be included in the sum of (2) while if $\sigma_{A_k}^2 \mathbf{R}_k^H \mathbf{K}_k^{-1} \mathbf{R}_k >> 1$, the weighting is unity and the test statistic for sensor k is included without downweighting. Assuming for the sake of interpretation that $\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i >> 1$ for all i, we have from (2)

$$T(\mathbf{X}) = 2 \sum_{i=1}^{M} \frac{\left| \mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i \right|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \tag{5}$$

and if we let $\hat{A}_i = (\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)^{-1} \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{X}_i$, then this becomes

$$T(\mathbf{X}) = 2 \sum_{i=1}^{M} \frac{|\hat{A}_i|^2}{(\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)^{-1}}. \tag{6}$$

Since by Fourier transforming (1) we have the complex linear model form of $\mathbf{X}_i = A_i \mathbf{R}_i + \mathbf{W}_i$, where $\mathbf{W}_i = \mathbf{C}_i + \mathbf{N}_i$, and A_i is independent of $\mathbf{W}_i$, we see that the estimator $\hat{A}_i$ is just the *conditional* minimum variance unbiased (MVU) estimator of A_i, and is also efficient conditionally. By conditional we mean that we are given A_i so that equivalently it can be regarded as deterministic in a conditional sense. It follows that $\mathrm{var}(\hat{A}_i) = (\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)^{-1}$ is the conditional minimum variance, and satisfies the Cramer-Rao lower bound. Hence, the NP detector sums the normalized estimated powers of the signal received at the sensors.

Interestingly, the detector of (5) has also been derived *but under the assumption of deterministic and unknown A_i's* as a GLRT [10,13]. It should be noted, however, that if some of the sensors have little signal power, then the GLRT performance will be degraded. In this case, the test statistic of (5) will add in noise only. On the other hand, the NP detector of (2) recognizes this possibility by incorporating the weighting factor α_i, which will downweight any sensor contribution for which $\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i << 1$. Of course, to do so it requires a priori knowledge of the signal power, i.e., knowledge of the $\sigma_{A_i}^2$'s, which the GLRT does not have privy to. In radar it is quite often the case that the bistatic radar cross-section can vary by as much as 30 dB [23] so that this problem of adding in noise only is a concern. The GLRT can still be used but should be modified to alleviate this problem. Results to this effect will be reported on in a future paper. In the next section we determine the detection performance of (2).

4 Performance of the NP Detector

It is shown in Appendix B that the detection statistic of (2) has the distribution

$$T(\mathbf{X}) \quad \sim \quad \sum_{i=1}^{M} \alpha_i^{(0)} \chi_2^2(i) \qquad \text{under } \mathcal{H}_0$$

$$\sim \sum_{i=1}^{M} \alpha_i^{(1)} \chi_2^2(i) \qquad \text{under } \mathcal{H}_1$$

where "$\sim$" means *is distributed according to*, and

$$\begin{aligned} \alpha_i^{(0)} &= \frac{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \\ \alpha_i^{(1)} &= \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i \end{aligned} \tag{7}$$

and $\{\chi_2^2(1), \chi_2^2(2), \ldots, \chi_2^2(M)\}$ are a set of independent and identically distributed chi-squared random variables with two degrees of freedom. Note that if $\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i >> 1$, then $\alpha_i^{(0)} = 1$, and $T(\mathbf{X}) \sim \chi_{2M}^2$ under $\mathcal{H}_0$ but otherwise not. It is well known that a weighted sum of independent χ_2^2 random variables has a mathematically tractable probability density function (PDF). Following the results in [10], we can determine the probability of false alarm P_{FA} and probability of detection P_D. To do so note that the characteristic function of $T = \sum_{i=1}^{M} \alpha_i \chi_2^2(i)$ for independent $\chi_2^2(i)$'s is given by

$$\begin{aligned} \phi_T(\omega) &= \prod_{i=1}^{M} \frac{1}{1 - 2j\alpha_i \omega} \\ &= \sum_{i=1}^{M} \frac{P_i}{1 - 2j\alpha_i \omega} \end{aligned}$$

where

$$P_i = \prod_{\substack{n=1 \\ n \neq i}}^{M} \frac{1}{1 - \alpha_n / \alpha_i}.$$

The second equality assumes that the α_i's are all distinct in employing a partial fraction expansion. Inverting the characteristic function produces the PDFs

$$\begin{aligned} p(T; \mathcal{H}_0) &= \sum_{i=1}^{M} \frac{P_i}{2\alpha_i^{(0)}} \exp[-T/(2\alpha_i^{(0)})] \\ p(T; \mathcal{H}_1) &= \sum_{i=1}^{M} \frac{Q_i}{2\alpha_i^{(1)}} \exp[-T/(2\alpha_i^{(1)})] \end{aligned}$$

with both PDFs being zero for $T < 0$ and

$$\begin{aligned} P_i &= \prod_{\substack{n=1 \\ n \neq i}}^{M} \frac{1}{1 - \alpha_n^{(0)} / \alpha_i^{(0)}} \\ Q_i &= \prod_{\substack{n=1 \\ n \neq i}}^{M} \frac{1}{1 - \alpha_n^{(1)} / \alpha_i^{(1)}}. \end{aligned} \tag{8}$$

Upon integrating the PDFs from γ to infinity we have the final results

$$P_{FA} = \sum_{i=1}^{M} P_i \exp[-\gamma/(2\alpha_i^{(0)})] \tag{9}$$

$$P_D = \sum_{i=1}^{M} Q_i \exp[-\gamma/(2\alpha_i^{(1)})]. \quad (10)$$

Although we have obtained a closed form expression for the detection performance, it unfortunately lends little insight into optimal signal design. Hence, in the next section we utilize an approximate measure of detection performance that is more amenable to signal optimization. Once the signal has been chosen, (9) and (10) can be used to determine the actual detection performance. An example of this is given in Section 7.

5 Alternative Signal Design Criterion - Divergence

The influence of the transmitted signal on the detection performance of the NP detector is embedded in the parameters $\alpha_i^{(0)}$ and $\alpha_i^{(1)}$ as seen by (7–10). If we let $\rho_i = \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i$, then we have from (7) that

$$\begin{aligned} \alpha_i^{(0)} &= \frac{\rho_i}{1+\rho_i} \\ \alpha_i^{(1)} &= \rho_i. \end{aligned}$$

Note that ρ_i is a signal-to-noise ratio (SNR) and in fact using the large time-bandwidth approximation it becomes

$$\rho_i = \sigma_{A_i}^2 \sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n) + P_{n_i}(F_n)} = \sum_{n=-N/2}^{N/2} \frac{\sigma_{A_i}^2 |G_i(F_n)|^2 T |S(F_n)|^2}{P_{c_i}(F_n) + P_{n_i}(F_n)}$$

which in the case of a single sensor has been shown to be the deflection coefficient of the NP detector [11]. This means that for $M = 1$ the detection probability of the NP detector is monotonically increasing with ρ_1. Hence, any reasonable detection measure for the multistatic case should be a monotonic function of the ρ_i's. Such a measure, as we now show, is the divergence or symmetrized Kullback-Liebler measure. It can be connected with performance of a hypothesis test asymptotically via the Stein lemma [25, 26] and as a bound on performance as shown in [27]. The divergence is defined as

$$D = D(p_1||p_0) + D(p_0||p_1)$$

where

$$D(p_i||p_j) = \int p_i(\mathbf{x}) \ln \frac{p_i(\mathbf{x})}{p_j(\mathbf{x})} d\mathbf{x}$$

and represents a type of "distance" between the two PDFs $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$. For our problem $p_0(\mathbf{x}) = p(\mathbf{X}; \mathcal{H}_0)$ and $p_1(\mathbf{x}) = p(\mathbf{X}; \mathcal{H}_1)$. Since each PDF is factorable due to the independence assumption, the divergence is additive and so can be shown to result in

$$D = \sum_{i=1}^{M} D_i$$

where

$$\begin{aligned} D_i &= \int p(\mathbf{X}_i;\mathcal{H}_0)\ln\frac{p(\mathbf{X}_i;\mathcal{H}_0)}{p(\mathbf{X}_i;\mathcal{H}_1)}d\mathbf{X}_i + \int p(\mathbf{X}_i;\mathcal{H}_1)\ln\frac{p(\mathbf{X}_i;\mathcal{H}_1)}{p(\mathbf{X}_i;\mathcal{H}_0)}d\mathbf{X}_i \\ &= E_0\left[\ln\frac{p(\mathbf{X}_i;\mathcal{H}_0)}{p(\mathbf{X}_i;\mathcal{H}_1)}\right] + E_1\left[\ln\frac{p(\mathbf{X}_i;\mathcal{H}_1)}{p(\mathbf{X}_i;\mathcal{H}_0)}\right] \end{aligned}$$

and the subscripts on the expected values indicate the averaging PDF. This can be evaluated by using results in Appendix A where it is shown that (see (22))

$$\begin{aligned} l_i(\mathbf{X}_i) &= \ln\frac{p(\mathbf{X}_i;\mathcal{H}_1)}{p(\mathbf{X}_i;\mathcal{H}_0)} \\ &= c_i + \alpha_i^{(0)}\frac{\left|\mathbf{X}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i\right|^2}{\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i} \end{aligned}$$

and therefore

$$\begin{aligned} D_i &= E_0\left[-c_i - \alpha_i^{(0)}\frac{\left|\mathbf{X}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i\right|^2}{\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}\right] + E_1\left[c_i + \alpha_i^{(0)}\frac{\left|\mathbf{X}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i\right|^2}{\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}\right] \\ &= \frac{\alpha_i^{(0)}}{\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}\left[-E_0\left[\left|\mathbf{X}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i\right|^2\right] + E_1\left[\left|\mathbf{X}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i\right|^2\right]\right] \\ &= \frac{\alpha_i^{(0)}}{\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}\left[-\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i + \mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i + \sigma_{A_i}^2(\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i)^2\right] \\ &= \alpha_i^{(0)}\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i. \end{aligned}$$

We have that

$$D = \sum_{i=1}^{M}\frac{\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}{1+\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i}\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i$$

and for mathematical expediency as well as a practical implementation, we assume that $\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i >> 1$. Finally, we have that

$$\begin{aligned} D &\approx \sum_{i=1}^{M}\sigma_{A_i}^2\mathbf{R}_i^H\mathbf{K}_i^{-1}\mathbf{R}_i \\ &\approx \sum_{i=1}^{M}\sigma_{A_i}^2\sum_{n=-N/2}^{N/2}\frac{|R_i(F_n)|^2}{P_{c_i}(F_n)+P_{n_i}(F_n)}. \end{aligned} \tag{11}$$

It is seen that the divergence is the sum of the SNRs of the sensors. Observe that the weighting coefficients are still present but are now given by $\sigma_{A_i}^2$ and also that the detectability measure is monotonic with the SNR at each sensor. Although not directly tied to the performance of the NP detector as it was for a single receive sensor, it nonetheless appears to be a reasonable measure. And in fact, it appears to accurately predict performance gains as will be described in Section 7. As mentioned previously, under some large

data record conditions the probability of detection can be written as a monotonic function of the divergence but not in general for finite data records. In the next section we will see how to maximize (11) by a suitable choice of the transmit signal.

6 Signal Design

The signal design problem is to maximize D as given by (11) subject to a constraint on the signal energy. This generalizes the problem that was solved in [11]. Since $F_n = n/T = n\Delta F$, we have from (11) that

$$\begin{aligned} D &= \sum_{n=-N/2}^{N/2} \sum_{i=1}^{M} \sigma_{A_i}^2 \frac{|R_i(F_n)|^2 T}{P_{c_i}(F_n) + P_{n_i}(F_n)} \Delta F \\ &\approx \int_{-W/2}^{W/2} \sum_{i=1}^{M} \frac{\sigma_{A_i}^2 T |R_i(F)|^2}{P_{c_i}(F) + P_{n_i}(F)} dF \end{aligned}$$

and since $r_i(t) = s(t) \star g_i(t)$, we have that $\sqrt{T}R_i(F) = \sqrt{T}S(F)\sqrt{T}G_i(F)$, and also $P_{c_i}(F) = P_{h_i}(F)T|S(F)|^2$. Letting the signal energy spectral density (ESD) be $\mathcal{E}_s(F) = T|S(F)|^2$, we have

$$D = \int_{-W/2}^{W/2} \sum_{i=1}^{M} \frac{\sigma_{A_i}^2 T |G_i(F)|^2 \mathcal{E}_s(F)}{\mathcal{E}_s(F) P_{h_i}(F) + P_{n_i}(F)} dF. \tag{12}$$

Note that for $M = 1$ and $T|G_i(F)|^2 = 1$ we have the same measure of detectability as given in [11]. The energy constraint can now be expressed as

$$\int_{-W/2}^{W/2} \mathcal{E}_s(F) dF \leq \mathcal{E}$$

and as has been shown in [11], the maximum of D will be attained when

$$\int_{-W/2}^{W/2} \mathcal{E}_s(F) dF = \mathcal{E}. \tag{13}$$

Maximization of D can by accomplished by two different numerical methods. Both ensure that the *global maximum* will be found. The methods are dynamic programming (DP) [28] and the maximum marginal allocation (MMA) approach [29]. To see how to apply these approaches we first revert back to the discretized version of D since ultimately a numerical solution will be sought. From (12) this is

$$D = \sum_{n=-N/2}^{N/2} \sum_{i=1}^{M} \frac{\sigma_{A_i}^2 \mathcal{E}_s(F_n) T |G_i(F_n)|^2}{\mathcal{E}_s(F_n) P_{h_i}(F_n) + P_{n_i}(F_n)} \Delta F \tag{14}$$

and next relabel the frequencies as $F_k = -W/2 + k\Delta F$ for $k = 0, 1, \ldots, N$ and $\Delta F = W/N$ so that

$$D = \sum_{k=0}^{N} \sum_{i=1}^{M} \frac{\sigma_{A_i}^2 \mathcal{E}_s(F_k) T |G_i(F_k)|^2 \Delta F}{\mathcal{E}_s(F_k) P_{h_i}(F_k) + P_{n_i}(F_k)}.$$

The constraint becomes

$$\sum_{k=0}^{N} \mathcal{E}_s(F_k)\Delta F = \mathcal{E}.$$

To simplify the notation we let $u(k) = \mathcal{E}_s(F_k)$ and

$$\begin{aligned} a_i^{(k)} &= \frac{P_{h_i}(F_k)}{\sigma_{A_i}^2 |G_i(F_k)|^2} \\ b_i^{(k)} &= \frac{P_{n_i}(F_k)}{\sigma_{A_i}^2 |G_i(F_k)|^2} \end{aligned}$$

so that

$$D = \sum_{k=0}^{N} \sum_{i=1}^{M} \frac{u(k)}{a_i^{(k)} u(k) + b_i^{(k)}}$$

and the constraint becomes $\sum_{k=0}^{N} u(k) = \mathcal{E}/\Delta F = u_{\max}$ and of course $u(k) \geq 0$. Hence we seek to maximize

$$D = \sum_{k=0}^{N} L(u(k), k) \tag{15}$$

where

$$L(u(k), k) = \sum_{i=1}^{M} L_i(u(k), k) \tag{16}$$

for

$$L_i(u(k), k) = \frac{u(k)}{a_i^{(k)} u(k) + b_i^{(k)}}. \tag{17}$$

As shown in [11] $L_i(u(k), k)$ is a concave function and therefore $L(u(k), k)$, being a sum of concave functions, is also concave. As a result D is a concave function. Also, it is easy to see that each $L_i(u(k), k)$ is monotonically increasing and so $L(u(k), k)$ is also monotonically increasing. This type of problem has been well studied and is termed the *resource allocation problem* [30]. Because of the concavity the use of a Lagrangian multiplier approach will produce the global maximum. Unfortunately, the equations obtained by differentiation cannot easily be solved, as was the case when $M = 1$. Alternative numerical means, however, are possible and include the DP and the MMA approaches. DP can be used even if the $L(u(k), k)$'s are not concave and even if additional constraints on the $u(k)$'s are necessary [30]. However, for our current problem we can take advantage of the form of the $L(u(k), k)$'s to utilize the very simple and computationally efficient approach of MMA, which we now describe.

We first discretize the range of $u(k)$. Since $\sum_{k=0}^{N} u(k) = u_{\max}$ and $0 \leq u(k) \leq u_{\max}$ for all k, we let $u(k)$ take on values in the set $\{0, \Delta, 2\Delta, \ldots, P\Delta\}$, where $P\Delta = u_{\max}$. We then allocate the $P\Delta$ units of "energy" by allocating Δ units at each step of the algorithm. In the first step, we set $u(j) = \Delta$ if $L(u(j), j) > L(u(k), k)$ for all $k \neq j$. We then repeat the same procedure except we choose to allocate the next Δ to the value of k that produces the maximum value of

$$\{L(\Delta, 0), L(\Delta, 1), \ldots, L(\Delta, j-1), L(2\Delta, j) - L(\Delta, j), L(\Delta, j+1), \ldots, L(\Delta, N)\}$$

or to the k that results in the maximum *marginal* increase. Note that since we have already allocated Δ to $k = j$ in step 1, the marginal increase for $k = j$ is $L(2\Delta, j) - L(\Delta, j)$ in going from an allocation of Δ to 2Δ. For the other values of k the marginal increases are the same as before. We continue this process until $\sum_{k=0}^{N} u(k) = P\Delta$ and so only P steps are needed for completion. A simple numerical example follows.

As an illustration let $N = 2$, $M = 1$, and $u(0) + u(1) + u(2) = u_{\max} = 4$, and say we wish to maximize

$$D = L(u(0), 0) + L(u(1), 1) + L(u(2), 2) = \frac{u(0)}{2u(0) + 4} + \frac{u(1)}{2u(1) + 1} + \frac{u(2)}{3u(2) + 2}.$$

We let $\Delta = 1$ so that we can allocate either 0, 1, 2, 3, or 4 units to $u(0), u(1), u(2)$ subject to the total energy constraint. The possible values of $L(u(k), k)$ are shown in Table 1, and correspond to the values of $L(u(k), k)$ for $u(k) = 0, 1, 2, 3, 4$. Note that the values of the $L(u(k, k)$'s for an initial allocation of $\Delta = 1$ units are 0.167, 0.333, and 0.200 for $k = 0, 1, 2$, respectively so that at step 1 Δ units are allocated to $k = 1$. After doing so the new *marginal* energies are 0.167, 0.067, and 0.200 as shown in Table 2 so that $k = 2$ is chosen in step 2. For step 3 the new marginal energies are 0.167, 0.067, and 0.050 as shown in Table 3 so that $k = 1$ is chosen, etc. Similarly, the last allocation is determined by examination of Table 4. The final allocations are summarized in Table 5. The final allocation is 2Δ for $k = 0$, Δ for $k = 1$, and Δ for $k = 2$ resulting in the maximizing values of $u(0) = 2\Delta = 2$, $u(1) = \Delta = 1$, and $u(2) = \Delta = 1$. The maximum value of D is 0.783. Note that if Δ is decreased the maximizing values can be found more accurately. In fact, by letting $u(0) = 4 - u(1) - u(2)$ and tabulating D versus $u(1)$ and $u(2)$ over the allowable range of $0 \leq u(1) + u(2) \leq 4$, it is found that the maximum value of D is 0.791 by using $\Delta = 0.01$. The corresponding values of the $u(k)$'s are $u(0) = 1.63$, $u(1) = 1.32$, and $u(2) = 1.05$.

$u(k) = 4$	0.333	0.444	0.286
$u(k) = 3$	0.300	0.429	0.273
$u(k) = 2$	0.250	0.400	0.250
$u(k) = 1$	0.167	<u>0.333</u>	0.200
$u(k) = 0$	0	0	0
	$k = 0$	$k = 1$	$k = 2$

Table 1: Values of $L(u(k), k)$ for various values of $u(k)$. The underlined entry is the maximum marginal increase obtained by letting $u(k) = \Delta = 1$ at step one.

7 Signal Design Example

In this section we give an example of a signal design for $M = 3$ receive sensors. The signal bandwidth is $W = 5$ Mhz, the signal time duration is $T = 1\mu$s , and the energy constraint is $\mathcal{E} = 10^6$ joules. The noise

$u(k)=4$	0.333		0.286
$u(k)=3$	0.300	0.111	0.273
$u(k)=2$	0.250	0.096	0.250
$u(k)=1$	0.167	0.067	<u>0.200</u>
$u(k)=0$	0	0	0
	$k=0$	$k=1$	$k=2$

Table 2: Marginal values of $L(u(k),k)$ for various values of $u(k)$ after first allocation. The underlined entry is the maximum marginal increase obtained by letting $u(k)=\Delta=1$.

$u(k)=4$	0.333		
$u(k)=3$	0.300	0.111	0.086
$u(k)=2$	0.250	0.096	0.073
$u(k)=1$	<u>0.167</u>	0.067	0.050
$u(k)=0$	0	0	0
	$k=0$	$k=1$	$k=2$

Table 3: Marginal values of $L(u(k),k)$ for various values of $u(k)$ after second allocation. The underlined entry is the maximum marginal increase obtained by letting $u(k)=\Delta=1$.

$u(k)=4$			
$u(k)=3$	0.183	0.111	0.086
$u(k)=2$	0.133	0.096	0.073
$u(k)=1$	<u>0.083</u>	0.067	0.050
$u(k)=0$	0	0	0
	$k=0$	$k=1$	$k=2$

Table 4: Marginal values of $L(u(k),k)$ for various values of $u(k)$ after third allocation. The underlined entry is the maximum marginal increase obtained by letting $u(k)=\Delta=1$.

characteristics are given by

$$P_{n_i}(F)=\sum_{k=1}^{3}P_k^{(i)}\left(\frac{\sin[\pi(F-F_k)T]}{\pi(F-F_k)T}\right)^4+N_0$$

where $F_1=1$ Mhz, $F_2=0.5$ Mhz, $F_3=-2$ Mhz. The interference peak powers $P_k^{(i)}$ for the three sensors are shown in Table 6, and $N_0=1$. For the clutter we assume a worst case of $P_{h_i}(F)=1$ for all sensors.

Step	$k=0$	$k=1$	$k=2$	D
1	–	Δ	–	0.333
2	–	–	Δ	0.533
3	Δ	–	–	0.700
4	Δ	–	–	0.783
	2Δ	Δ	Δ	

Table 5: Final allocation of energies.

Sensor	$\sigma^2_{A_i}$	$P_1^{(i)}$	$P_2^{(i)}$	$P_3^{(i)}$
$i=1$	10	1×10^4	10×10^4	100×10^4
$i=2$	1	10×10^4	1×10^4	100×10^4
$i=3$	1	100×10^4	10×10^4	1×10^4

Table 6: Signal powers and interference peak powers for the three sensors.

This says that the clutter return has the same spectral shape as the transmit signal. The target frequency response is also assumed flat so that $T|G_i(F)|^2 = 1$ for all sensors and the target reflection powers $\sigma^2_{A_i}$ are given in Table 6. The interference PSD (which is $P_{n_i}(F) - N_0$) for the three sensors is plotted in Figure 2. Note that sensor 1 has the largest target return, although its minimum interference, which occurs at at

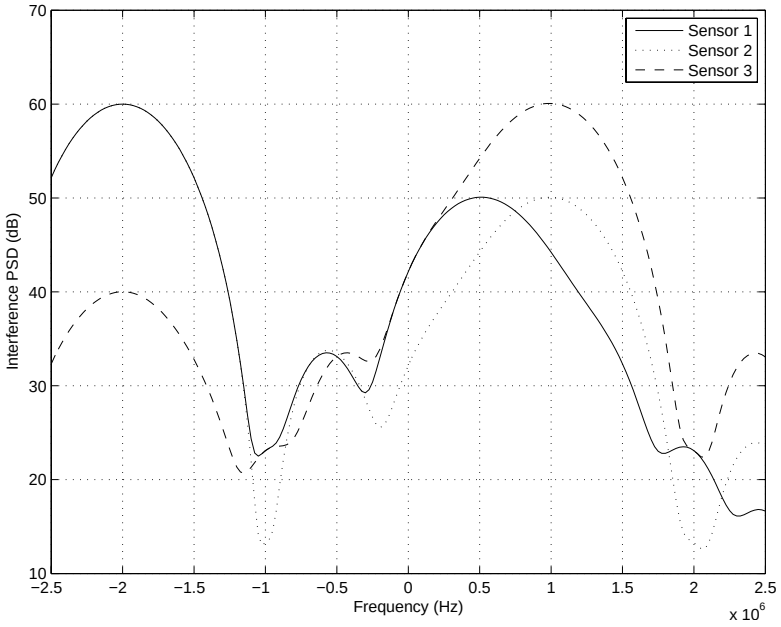

Figure 2: Interference PSD for the three sensors.

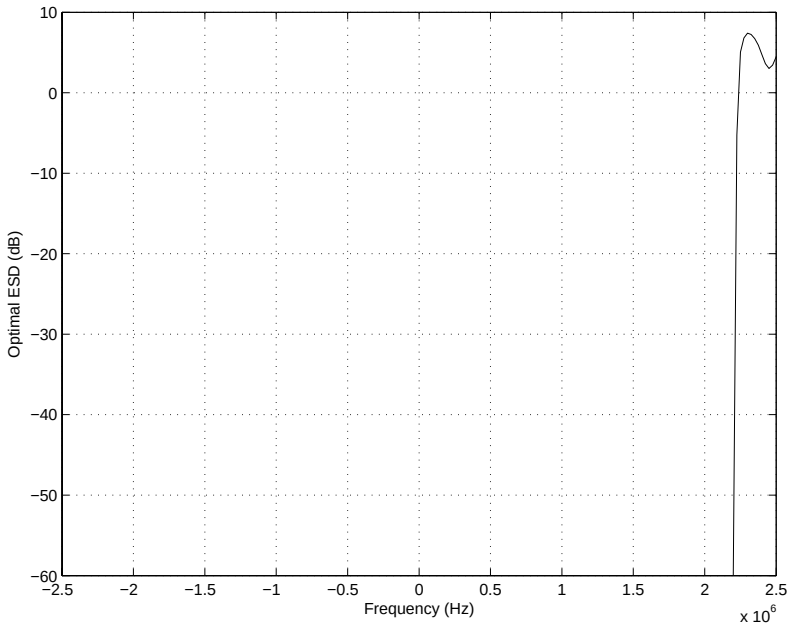

Figure 3: Optimal transmit signal when sensor 1 has largest target power.

$F = 2.5$ Mhz, is slightly higher than the minimum interference of sensor 2, which occurs at 2.1 Mhz. As a result, as shown in Figure 3 the optimal transmit signal places most of its energy at about 2.5 Mhz. If, however, we reduce the target power for sensor 1 from $\sigma_{A_1}^2 = 10$ to $\sigma_{A_1}^2 = 1$ so that all three sensors have the same target reflection powers, then the signal design is quite different as shown in Figure 4. Now the design is determined by the disparity of interference power with frequency. Since sensor 2 exhibits the least amount of power versus frequency in bands around -1 Mhz and 2 Mhz, this is where the signal energy is concentrated. To determine the improvement in the sum of SNRs of the sensors for this example we can use (14) with $T|G_i(F_n)|^2 = 1$ to yield

$$D = \sum_{n=-N/2}^{N/2} \sum_{i=1}^{M} \frac{\sigma_{A_i}^2 \mathcal{E}_s(F_n)}{\mathcal{E}_s(F_n) P_{h_i}(F_n) + P_{n_i}(F_n)} \Delta F. \tag{18}$$

A typical transmit signal is a linear FM whose ESD is $\mathcal{E}_s(F) = \mathcal{E}/W$ for $|F| \leq W/2$. Using this in (18) with $N = 200$ so that $\Delta F = W/N = 2.5 \times 10^4$ and for the conditions shown in Table 6 produces $D_{\text{lfm}} = 44.8$ dB. For the optimal transmit signal we have $D_{\text{opt}} = 53.3$ dB, a gain of 8.5 dB. For the second condition in which all three sensors have the same target reflection powers the results are $D_{\text{lfm}} = 39.1$ dB and $D_{\text{opt}} = 47.1$ dB, a gain of 8 dB.

To actually synthesize a transmit signal with the ESD given in Figures 3 or 4 the procedure described in [11] can be used. There it was shown how to weight a pulse train using complex weights in order to achieve the desired ESD.

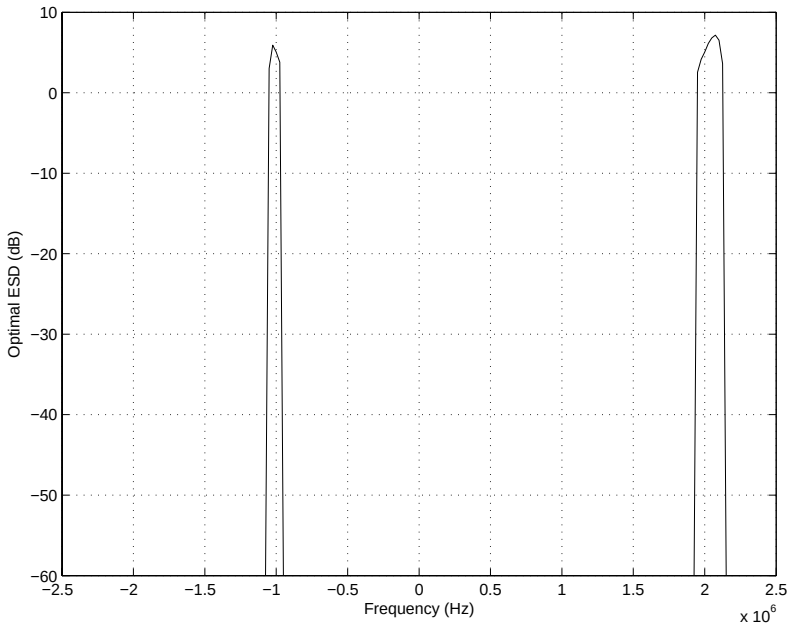

Figure 4: Optimal transmit signal when target reflection powers are the same for all sensors.

Finally, we give an example of the detection performance via the use of (9) and (10). We use the same conditions as used to generate the optimal transmit signal shown in Figure 4 except that we vary the target reflection powers from $\sigma_{A_i}^2 = 10^{-4}$ to 10^{-2} and assume that the $\sigma_{A_i}^2$'s are the same for each sensor. When the $\sigma_{A_i}^2$'s are equal the optimal signal will be the same as in Figure 4 as may be seen by examining D as given in (18) (for Figure 4 we had set $\sigma_{A_i}^2 = 1$ for $i = 1, 2, 3$). For $\sigma_{A_i}^2 = 0.001$ for all i we have that $D_{\text{lfm}} = 9.1$ dB and $D_{\text{opt}} = 17.1$ dB, a gain again of 8 dB. Note that by reducing $\sigma_A^2 = 1$ to $\sigma_A^2 = 0.001$, the divergence criterion is reduced by 30 dB as expected. We next plot P_d for the optimal signal and an LFM signal. A probability of false alarm of $P_{FA} = 10^{-5}$ is chosen. Note that to determine the threshold from (9) we need to solve this for each set of $\alpha_i^{(0)}$'s since the latter will depend upon $\sigma_{A_i}^2$ as given in (7). After finding the threshold γ for the given P_{FA} and also for each set of $\sigma_{A_i}^2$'s, we use (10) to determine the P_D for the corresponding set of $\alpha_i^{(1)}$'s. The detection performance versus σ_A^2 in dB is shown in Figure 5. Note that for $P_D = 0.5$ the improvement in detection performance is about 8 dB as predicted by the improvement in the divergence.

8 Conclusions

We have derived the optimal Neyman-Pearson detector for multiple independent receive sensors. Although the probability of detection is not monotonically related to a single parameter that depends on the signal, we have seen that the use of divergence for signal design is a reasonable alternative. It allows prediction

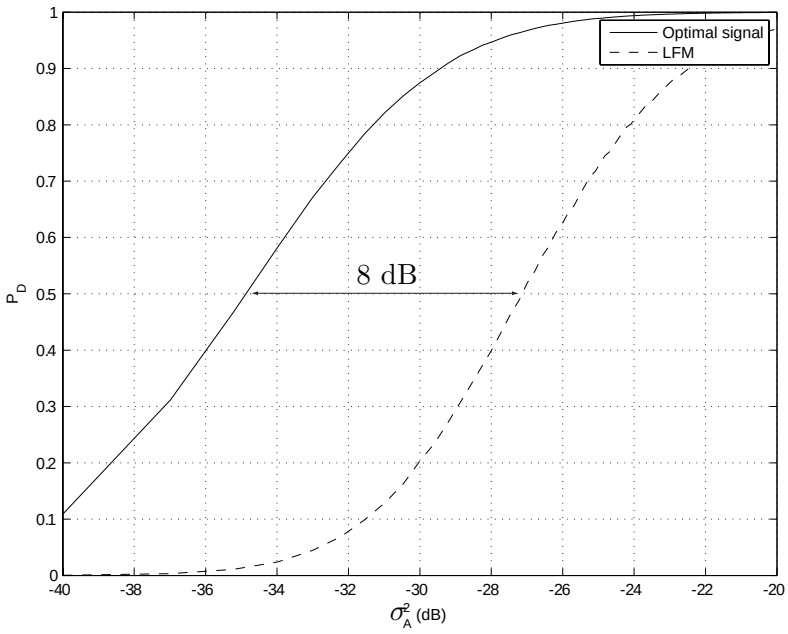

Figure 5: Probability of detection performance when target reflective power is the same for all sensors. The improvement in detection performance is accurately predicted by the divergence.

of the gain in detection performance. Furthermore, we have presented a numerical method for signal design that is guaranteed to maximize the divergence and is simple to implement. It is believed that if the assumed statistical quantities are known or can be accurately estimated in real-time, then the signal design approach presented can be used "in-situ" to optimize radar performance. Finally, it should be noted that the analysis and solution has included the effects of clutter as signal-dependent noise which complicates the problem but is more realistic in practice.

References

1. Rihaczek, A.W., "Optimal Filters for Signal Detection in Clutter", *IEEE Trans. on Aero. and Electr. Systems*,pp.297–299, Dec. 1965.

2. Rummler, W.D., "Clutter Supression by Complex Weighting of Coherent Pulse Trains", *IEEE Trans. on Aero. and Electr. Systems*,pp. 689–699, 1966.

3. Rummler, W.D., "A Technique for Improving the Clutter Performance of Coherent Pulse Trains", *IEEE Trans. on Aero. and Electr. Systems*, pp. 898–906, 1967.

4. Thompson, J.S., E.L. Titlebaum, "The Design of Optimal Radar Waveforms for Clutter Rejection using the Maximum Principle", *Supplement to IEEE Trans. on Aero. and Electr. Systems*, pp. 581–589, 1967.

5. Spafford, L.J., "Optimal Radar Signal Processing in Clutter", *IEEE Trans. on Info. Theory*, pp. 734–743, 1968.

6. Balakrishnan, A.V., "Signal Design for a Class of Clutter Channels", *IEEE Trans. on Info. Theory*, pp. 170–173, Jan. 1968.

7. Delong, D.F., Jr., E.M. Hofstetter, "The Design of Clutter-Resistant Radar Waveforms with Limited Dynamic Range", *IEEE Trans. Info. Theory*, pp. 376–385, 1969.

8. H.L. Van Trees, *Detection, Estimation, and Modulation Theory, Part III*, J. Wiley, NY, 1971.

9. Castella, F.R., J.R. Moore, "Improved Target Detection in a Littoral Environment", in *Radar'97*, pp. 429–433, Oct, 1997.

10. Kay, S., *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall, Upper Saddle River, NJ, 1998.

11. Kay, S., "Optimal Signal Design for Detection of Gaussian Point Targets in Stationary Gaussian Clutter/Reverberation", *IEEE Trans. on Selected Topics in Signal Processing*, pp. 31–41, June 2007.

12. Wang, X., H.V. Poor, *Wireless Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 2004.

13. Jaffer,A., B. Evans, P. Zulch and M. Rangaswamy, "Combined GLRT Processing Method for Multi-static Radar Systems", Proceedings 53rd MSS Tri-Service Radar Symposium, Orlando, Florida, 21 - 25 May, 2007.

14. Subotic, N.,B. Bachelor, K. Cooper, P. Zulch, "Joint Radar Waveform Design for Networked ISR Systems", 2006 International Conference on Waveform Diversity and Design, Hawaii

15. Leshem, A., O. Naparstek, A. Nehorai, "Information Theoretic Adaptive Radar Waveform Design for Multiple Extended Targets", *IEEE Trans. on Selected Topics in Signal Processing*, pp. 42–55, June 2007.

16. Yang, Y., R. Blum, "Minimax Robust MIMO Radar Waveform Design", *IEEE Trans. on Selected Topics in Signal Processing*, pp. 147–155, June 2007.

17. Wozencraft, J.M., I.M. Jacobs, *Principles of Communication Engineering*, pp. 534–539, J. Wiley and Sons, New York, 1965.

18. Kooij, T., "Optimum Signals in Noise and Reverberation", in *Proceedings of NATO Advanced Study Institute on Signal Processing with emphasis on underwater acoustics*, Vol. I, Enschede, the Netherlands, 1968.

19. Pillai, S.U., H.S. Oh, D.C. Youla, J.R. Guerci, "Optimal Transmit-Receiver Design in the Presence of Signal-Dependent Interference and Channel", *IEEE Trans. on Info. Theory*, Vol. 46, No. 2, pp. 577-584, March 2000.

20. Bergin, J.S., P.M. Techau, J.E. Don Carlos, J.R. Guerci, "Radar Waveform Optimization for Colored Noise Mitigation", 2005 IEEE Radar Conf., Alexandria, VA, May 2005.

21. Kay, S., *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Upper Saddle River, NJ, 1993.

22. Kay, S., *Intuitive Probability and Random Process using MATLAB*, Springer, NY, 2006.

23. Willis, N.J, *Bistatic Radar*, SciTech Pub., NC, 2007.

24. H.L. Van Trees, H.L., *Optimal Array Processing*, pp. 334–343, J. Wiley, NY, 2002.

25. Kullback, H., *Information Theory and Statistics*, Dover Pub., New York, 1959.

26. Blahut, R.E., *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, 1987.

27. Kailath, T., "The Divergence and Bhattacharyya Distance Measures in Signal Selection", *IEEE Trans. Comm.*, pp. 52–60, Feb. 1967.

28. Larson, R.E., J.L. Casti, *Principles of Dynamic Programming, Part I, II*, Marcel Dekker, New York, 1978.

29. Ibaraki, T., N. Katoh, *Resource Allocation Problems, Algorithmic Approaches*, MIT Press, 1988.

30. Bellman, R., *Dynamic Programming*, Dover Pub., New York, 2003.

A Derivation of Neyman-Pearson Detector and its Performance

The modeling assumptions have been given in Section 2. We first transform $x_i(t), c_i(t), n_i(t)$ into the frequency domain using the normalized Fourier transform of (3) and then take frequency samples at $F_n = n/T$ for $n = -N/2, \ldots, N/2$. Assuming $WT > 16$ [24], the frequency samples are approximately independent and are also complex Gaussian distributed. Hence, the detection problem of (1) can be recast as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X}_i = \mathbf{C}_i + \mathbf{N}_i \\ \mathcal{H}_1 &: \mathbf{X}_i = A_i \mathbf{R}_i + \mathbf{C}_i + \mathbf{N}_i \end{aligned}$$

for $i = 1, 2, \ldots, M$ and where for example $\mathbf{X}_i = [X(F_{-N/2}) \ldots X(F_{N/2})]^T$ is an $(N+1) \times 1$ complex random vector. Since $A_i \sim \mathcal{CN}(0, \sigma^2_{A_i})$, where "$\mathcal{CN}$" denotes a complex Gaussian PDF, the clutter random process $h_i(t)$ is Gaussian, the noise random process $n_i(t)$ is Gaussian, and $r_i(t) = s(t) \star g_i(t)$ is a deterministic signal, it follows that $A_i \mathbf{R}_i$, $\mathbf{C}_i$, and $\mathbf{N}_i$ are all complex multivariate Gaussian random vectors with a zero mean vector. Also, $A_i, h_i(t), n_i(t)$ are independent of each other so that the PDF under $\mathcal{H}_1$ can be written as

$$p(\mathbf{X}_i; \mathcal{H}_1) = \frac{1}{\pi^{N+1} \det(\sigma^2_{A_i} \mathbf{R}_i \mathbf{R}_i^H + \mathbf{K}_i)} \exp\left[-\mathbf{X}_i^H (\sigma^2_{A_i} \mathbf{R}_i \mathbf{R}_i^H + \mathbf{K}_i)^{-1} \mathbf{X}_i\right] \tag{19}$$

where $\mathbf{K}_i$ is the covariance matrix of $\mathbf{C}_i + \mathbf{N}_i$ and the PDF under $\mathcal{H}_0$ is given by

$$p(\mathbf{X}_i; \mathcal{H}_0) = \frac{1}{\pi^{N+1} \det(\mathbf{K}_i)} \exp\left[-\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{X}_i\right]. \tag{20}$$

Since it was furthermore assumed that the sensor outputs are all independent, it follows that the $\mathbf{X}_i$'s are independent as well. Thus,

$$p(\mathbf{X}; \mathcal{H}_k) = \prod_{i=1}^{M} p(\mathbf{X}_i; \mathcal{H}_k) \qquad k = 0, 1 \tag{21}$$

where the vector of all the Fourier transform samples from all the sensor outputs is $\mathbf{X} = [\mathbf{X}_1^T \ldots \mathbf{X}_M^T]^T$. Twice the log-likelihood ratio becomes

$$\begin{aligned} 2l(\mathbf{X}) &= 2 \ln \frac{p(\mathbf{X}; \mathcal{H}_1)}{p(\mathbf{X}; \mathcal{H}_0)} \\ &= 2 \ln \frac{\prod_{i=1}^{M} p(\mathbf{X}_i; \mathcal{H}_1)}{\prod_{i=1}^{M} p(\mathbf{X}_i; \mathcal{H}_0)} \\ &= 2 \sum_{i=1}^{M} \ln \frac{p(\mathbf{X}_i; \mathcal{H}_1)}{p(\mathbf{X}_i; \mathcal{H}_0)} \\ &= 2 \sum_{i=1}^{M} l_i(\mathbf{X}_i) \end{aligned}$$

where $l_i(\mathbf{X}_i)$ denotes the log-likelihood ratio resulting from the ith sensor output. The use of (19) and (20) produces the log-likelihood ratio for the ith sensor output of

$$l_i(\mathbf{X}_i) = c_i + \mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{X}_i - \mathbf{X}_i^H (\sigma_{A_i}^2 \mathbf{R}_i \mathbf{R}_i^H + \mathbf{K}_i)^{-1} \mathbf{X}_i$$

where c_i is a constant term not dependent on $\mathbf{X}_i$ that will ultimately be dropped since it does not affect the resultant NP test. We now have using Woodbury's identity

$$\begin{aligned} l_i(\mathbf{X}_i) &= c_i + \mathbf{X}_i^H \left[\mathbf{K}_i^{-1} - (\sigma_{A_i}^2 \mathbf{R}_i \mathbf{R}_i^H + \mathbf{K}_i)^{-1} \right] \mathbf{X}_i \\ &= c_i + \mathbf{X}_i^H \left[\mathbf{K}_i^{-1} - \mathbf{K}_i^{-1} + \frac{\sigma_{A_i}^2 \mathbf{K}_i^{-1} \mathbf{R}_i \mathbf{R}_i^H \mathbf{K}_i^{-1}}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \right] \mathbf{X}_i \\ &= c_i + \sigma_{A_i}^2 \frac{\left| \mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i \right|^2}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \\ &= c_i + \frac{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \frac{\left| \mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i \right|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} \end{aligned} \tag{22}$$

Dropping the constant term c_i, summing over i, and multiplying by two produces (2).

To express the log-likelihood ratio in simpler form we first note that since $\mathbf{C}_i$ and $\mathbf{N}_i$ are independent

$$\mathbf{K}_i = E[(\mathbf{C}_i + \mathbf{N}_i)(\mathbf{C}_i + \mathbf{N}_i)^H] = E[\mathbf{C}_i \mathbf{C}_i^H] + E[\mathbf{N}_i \mathbf{N}_i^H].$$

For a large WT each covariance matrix is diagonal with diagonal elements equal to the PSD values [24]. Hence,

$$\mathbf{K}_i = \mathrm{diag}(P_{c_i}(F_{-N}) + P_{n_i}(F_{-N}), \ldots, P_{c_i}(F_N) + P_{n_i}(F_N))$$

and as a result (22) becomes upon dropping c_i

$$l_i(\mathbf{X}_i) = \frac{\sigma_{A_i}^2 \sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n)+P_{n_i}(F_n)}}{1 + \sigma_{A_i}^2 \sum_{n=-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n)+P_{n_i}(F_n)}} \frac{\left| \sum_{n=-N/2}^{N/2} \frac{X_i^*(F_n) R_i(F_n)}{P_{c_i}(F_n)+P_{n_i}(F_n)} \right|^2}{\sum_{n-N/2}^{N/2} \frac{|R_i(F_n)|^2}{P_{c_i}(F_n)+P_{n_i}(F_n)}}$$

and hence we have (4).

B Detection Performance of NP Detector

The NP detector decides $\mathcal{H}_1$ if

$$2l(\mathbf{X}) = 2 \sum_{i=1}^{M} \alpha_i \frac{|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}$$

exceeds a threshold. Under $\mathcal{H}_0$ we have that $\mathbf{X}_i = \mathbf{C}_i + \mathbf{N}_i = \mathbf{W}_i$ and therefore $\mathbf{X}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}_i)$. Under $\mathcal{H}_1$ we have that $\mathbf{X}_i = A_i \mathbf{R}_i + \mathbf{W}_i$ and therefore $\mathbf{X}_i \sim \mathcal{CN}(\mathbf{0}, \sigma_{A_i}^2 \mathbf{R}_i \mathbf{R}_i^H + \mathbf{K}_i)$. First considering the PDF of $l(\mathbf{X})$ under $\mathcal{H}_0$, we have

$$\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i = \mathbf{W}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i \sim \mathcal{CN}(0, \sigma_0^2)$$

and under $\mathcal{H}_1$

$$\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i = (A_i \mathbf{R}_i + \mathbf{W}_i)^H \mathbf{K}_i^{-1} \mathbf{R}_i \sim \mathcal{CN}(0, \sigma_1^2)$$

since the random vectors are complex Gaussian and therefore a linear transformation produces a scalar complex Gaussian random variable [21]. The variances are found as

$$\sigma_0^2 = E[|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2] = E[\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{W}_i \mathbf{W}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i] = \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i$$

and similarly

$$\sigma_1^2 = \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i + \sigma_{A_i}^2 (\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)^2.$$

Thus, under $\mathcal{H}_0$ we have that

$$\frac{2|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i} = 2 \left| \underbrace{\frac{\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{\sigma_0}}_{\mathcal{CN}(0,1)} \right|^2 \sim \chi_2^2$$

and under $\mathcal{H}_1$

$$\frac{2|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i (1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)} = 2 \left| \underbrace{\frac{\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{\sigma_1}}_{\mathcal{CN}(0,1)} \right|^2 \sim \chi_2^2.$$

Now under $\mathcal{H}_0$

$$2l_i(\mathbf{X}_i) = \underbrace{\frac{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}{1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}}_{\alpha_i^{(0)}} \underbrace{\frac{2|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}}_{\sim \chi_2^2(i)}$$

and under $\mathcal{H}_1$

$$2l_i(\mathbf{X}_i) = \underbrace{\sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i}_{\alpha_i^{(1)}} \underbrace{\frac{2|\mathbf{X}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i|^2}{\mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i (1 + \sigma_{A_i}^2 \mathbf{R}_i^H \mathbf{K}_i^{-1} \mathbf{R}_i)}}_{\sim \chi_2^2(i)}$$

and where all $\chi_2^2(i)$ random variables are independent. Finally, we have

$$2l(\mathbf{X}) = \begin{cases} \sum_{i=1}^M \alpha_i^{(0)} \chi_2^2(i) & \text{under } \mathcal{H}_0 \\ \sum_{i=1}^M \alpha_i^{(1)} \chi_2^2(i) & \text{under } \mathcal{H}_1. \end{cases}$$