Optimal Signal Design for Detection of Gaussian Point Targets in Stationary Gaussian Clutter/Reverberation

Steven Kay*
Dept. of Electrical and Computer 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

March 5, 2007

Abstract

In this paper we address the design of an optimal transmit signal and its corresponding optimal detector for a radar or active sonar system. The focus is on the temporal aspects of the waveform with the spatial aspects to be described in a future paper. The assumptions involved in modeling the clutter/reverberation return are crucial to the development of the optimal detector and its consequent optimal signal design. In particular, the target is assumed to be a Gaussian point target and the clutter/reverberation a stationary Gaussian random process. In practice, therefore, the modeling will need to be assessed and possibly extended, and additionally a means of measuring the "in-situ" clutter/reverberation spectrum will be required. The advantages of our approach are that a simple analytical result is obtained which is guaranteed to be optimal, and also the extension to spatial-temporal signal design is immediate using ideas of frequency-wavenumber representations. Some examples are given to illustrate the signal design procedure as well as the calculation of the increase in processing gain. Finally, the results are shown to be an extension of the usual procedure which places the signal energy in the noise band having minimum power.

*This work was supported by the Air Force Research Lab, Rome, NY under subcontract FA8750–04–C–0230 to SAIC, Chantilly, VA.

1 Introduction

The problem of signal waveform 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 colored noise whose spectrum *does not* depend on the transmitted signal, the solution is well known. It says to place all the signal energy into the frequency band for which the noise power is minimum. Correspondingly, for discrete signal vector design one should choose the signal as the eigenvector of the noise covariance matrix whose eigenvalue is minimum [10].

To date there has been no analytical solution to the signal-dependent noise problem, although a slightly flawed analysis did appear in [11]. In this paper we describe an approach that yields a simple solution, subject to the assumptions of a particular scattering model. Some of these results were previously summarized in [12], but here we include the missing details as well as expand the discussion to include examples and an actual signal synthesis method. The scattering model assumes that the signal-dependent noise is the output of a random linear time invariant (LTI) filter, whose impulse response can be assumed to be a realization of a wide sense stationary (WSS) random process. The same model has been used before in [13] and more recently in [14]. It should be noted that this model does not allow for spectral spreading, as would be inherent in a moving platform and/or intrinsic clutter motion situation. Hence, it differs from the standard one usually assumed [8]. However, subject to this limitation the advantage of such a model is that

1. An analytical solution for the optimal waveform is obtained.

2. New insights into the signal design problem are evident.

3. The results can be extended to the design of a spatial-temporal transmit signal using concepts of frequency-wavenumber filtering. This will be addressed in a future paper. Hence, multidimensional techniques or even multichannel techniques as are now important for MIMO radar [15] are easily derived.

Previous results using the random LTI scattering model were of limited practical utility. For example in [6, 13] a Fredholm equation needs to be solved and in [14] an iterative solution is proposed, which is neither guaranteed to converge nor to produce the optimal signal.

For a practical implementation one can envision a probing signal that measures the channel characteristics needed for waveform design. Then, the optimal transmit signal may be designed "in-situ". Techniques such as reported in [16] for channel estimation then become immediately applicable.

In this paper we first consider the modeling assumptions in Section 2, followed by the optimal detector and its performance in Section 3. In Section 4 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 5. A method to realize the optimal signal is described in Section 6 and finally conclusions are given in Section 7.

2 Problem Statement and Modeling Assumptions

The model for the received waveform is shown in Figure 1. For the purposes of this paper a point target is assumed so that $g(t) = A\delta(t)$. However, a more general extended target, such as is considered in [17] is easily accommodated. To clarify the exposition, however, we will only discuss the point target.

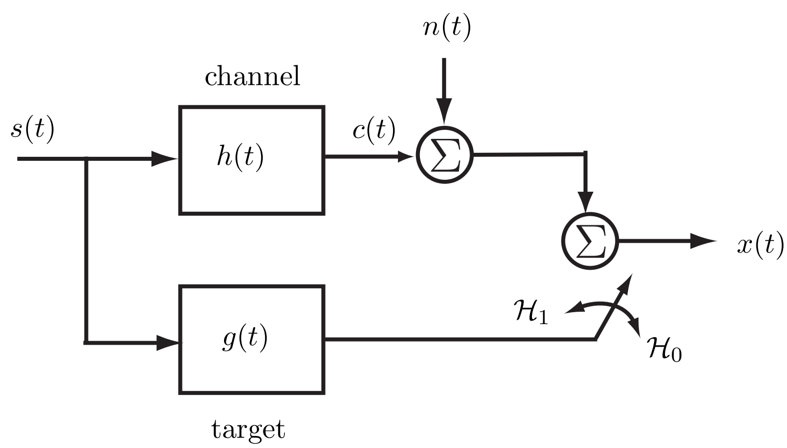

Figure 1: Modeling of received waveform. $s(t)$ is the transmitted signal, $h(t)$ is the impulse response of the random LTI channel filter, $g(t)$ is the impulse response of random LTI target filter, and $n(t)$ represents ambient noise and interference.

We assume that the received waveform is the complex envelope of the real bandpass data and is denoted by $x(t)$ for $|t| \leq T/2$. When no target return is present, i.e., under hypothesis $\mathcal{H}_0$, we have that $x(t) = c(t) + n(t)$, where $c(t)$ denotes clutter (henceforth, we will use radar terminology) and $n(t)$ is the sum of ambient noise and interference, i.e., jamming. Under the hypothesis $\mathcal{H}_1$, the target return is modeled as $As(t)$, where $s(t)$ is the complex envelope of the transmitted signal and A is a complex reflection factor with the probability density function (PDF) $A \sim \mathcal{CN}(0, \sigma_A^2)$. Here 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 detector is still applicable to the nonzero Doppler target but of course will require separate Doppler channels. Also, $n(t)$ is modeled as a complex WSS Gaussian random process with zero mean and power spectral density (PSD) $P_n(F)$. 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 $c(t)$ as the output of a *random LTI filter* with impulse response $h(t)$, whose input is the transmitted signal. This is the model used in [11, 14]. This type of modeling is appropriate for multipath [8] since the filtering will model the altered frequency spectrum of the return signal. (Note 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 $c(t) = s(t) \star h(t)$, where $\star$ denotes convolution, at the receiver. By reversing the convolution we can write this as $c(t) = h(t) \star s(t)$, where now the filter input is $h(t)$ and the filter impulse response is $s(t)$. If we now assume that $h(t)$ is a complex WSS Gaussian random process with zero mean and PSD $P_h(F)$, then $c(t)$ will also be a complex WSS Gaussian random process [18, 19] with zero mean and PSD $P_c(F) = T|S(F)|^2 P_h(F)$, where $S(F)$ is the normalized Fourier transform (the usual Fourier transform multiplied by $1/\sqrt{T}$) of $s(t)$. It should be noted that in modeling the channel impulse response by a random process it is implicitly assumed that the duration of the impulse response is time-limited. As a result, the model is capable of representing a stable and causal filter, which of course, the channel must be. In practice, this constraint is manifested in processing successive range windows over the range of interest.

3 Optimal Detector and its Performance for a Given Transmit Signal

With the previous modeling assumptions and assuming that the time-bandwidth product WT satisfies $WT > 16$ [21], we can easily derive an optimal detector. This is done in Appendix A and is based on the Neyman-Pearson criterion [10]. The optimal detector, which is conveniently expressed in the frequency domain, decides a signal is present if

$$\left| \sum_{m=-M/2}^{M/2} \frac{X(F_m) S^*(F_m)}{P_h(F_m) T |S(F_m)|^2 + P_n(F_m)} \right|^2 > \gamma \tag{1}$$

where $F_m = m/T$, $M = WT$, and the Fourier transforms are defined as

$$X(F) = \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x(t) \exp(-j2\pi F t) dt. \tag{2}$$

In practice, FFTs would be used to approximate the continuous-time Fourier transforms used in (1). The detection performance of the optimal detector is shown in Appendix A to be monotonically increasing with the parameter

$$d^2 = \sigma_A^2 \int_{-W/2}^{W/2} \frac{T|S(F)|^2}{P_h(F)T|S(F)|^2 + P_n(F)} dF. \tag{3}$$

Hence, the optimal signal design problem reduces to the relatively simple problem of choosing a signal $s(t)$ that is constrained in energy, which is defined as

$$\mathcal{E} = \int_{-W/2}^{W/2} T|S(F)|^2 dF \tag{4}$$

and that maximizes d^2. Note that it is only the energy spectral density (ESD) or $\mathcal{E}_s(F) = T|S(F)|^2$ that affects performance. The phase of the Fourier transform can be chosen arbitrarily and in practice will be selected for ease of signal realizability.

4 Maximizing the Detection Performance by Transmit Signal Design

The key to maximizing (3) over all $|S(F)|^2$, subject to the constraints that $|S(F)|^2 \geq 0$ and the energy constraint of (4), lies in the property that d^2 is a *concave functional* of $|S(F)|^2$. This assures us that the solution found via differential means will produce a *global maximum*. In contrast to this, we note that in [14] the algorithm given is iterative, for which it is neither guaranteed to converge nor if it converges, to produce the global maximum. There are no such limitations with the approach described here. However, with the proposed approach, we will only find $|S(F)|^2$, so that a further necessary step is to synthesize a time-limited signal with the given ESD. Fortunately, this is possible and amounts to a filter design problem based on a given magnitude frequency response specification. Many techniques are available to effect the design [20]. In Section 6 we use Durbin's method for moving average parameter estimation to realize a signal with a given ESD.

In Appendix C we derive the ESD that maximizes d^2. It is given by

$$\mathcal{E}_s(F) = T|S(F)|^2 = \max\left(\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)}, 0\right) \tag{5}$$

where $\max(x, 0)$ means the maximum of x and 0. The parameter λ is found from the energy constraint of (4) so that we must solve

$$\int_{-W/2}^{W/2} \max\left(\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)}, 0\right) dF = \mathcal{E} \tag{6}$$

for λ, where λ is positive. A solution for λ is guaranteed since if

$$g(\lambda) = \int_{-W/2}^{W/2} \max\left(\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)}, 0\right) dF$$

then $g(0) = \infty$ and $g(\infty) = 0$ and g is a continuous function, which means that it takes on all values in between by the intermediate value theorem. We can narrow down the search region for λ, however, by noting that for $\mathcal{E} > 0$, we must have $\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)} > 0$ for at least some values of F. We can then exclude those values of λ for which $\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)} \leq 0$ for all F. These are the values $\lambda \geq 1/P_n(F)$. Thus, we need not search the values for which $\lambda \geq 1/\min(P_n(F))$ or we have that the search region is

$$0 < \lambda < \frac{1}{\min(P_n(F))}.$$

We can also compute the maximum value of d^2 to allow us to determine either improvements over other detectors, either in the case of a suboptimal detector, e.g., a matched filter, or in the case of the optimal detector that uses a suboptimal transmit signal. The maximum value of d^2 is given by (3) with the optimal signal ESD given by (5).

Before considering some examples, we note the following interesting result. Because of the max operation in (5) we see that the optimal ESD will be zero for those frequencies for which

$$\sqrt{P_n(F)/\lambda} - P_n(F) \leq 0$$

or equivalently for the frequencies

$$P_n(F) \geq \frac{1}{\lambda}.$$

Hence, if there is a large amount of ambient noise and jamming in a particular frequency band, the optimal signal will not put any of its energy into that band. An illustration of this phenomenon is shown in Figure 2. It is normally referred to as "water filling" and also appears as the solution for distributing energy to maximize channel capacity for parallel channels [22].

5 Some Examples

5.1 White Ambient Noise and No Jamming or $P_n(F)$ is Flat

A very simple result occurs if $P_n(F)$ is a constant. This will occur for no jamming and white ambient noise. In this case $P_n(F) = N_0$ and therefore from (5)

$$\mathcal{E}_s(F) = \max\left(\frac{\sqrt{N_0/\lambda} - N_0}{P_h(F)}, 0\right).$$

It is clear that we must have $\sqrt{N_0/\lambda} - N_0 > 0$ or else $\mathcal{E}_s(F) = 0$ for all F. When this condition holds, we have that

$$\mathcal{E}_s(F) = \frac{\sqrt{N_0/\lambda} - N_0}{P_h(F)} \tag{7}$$

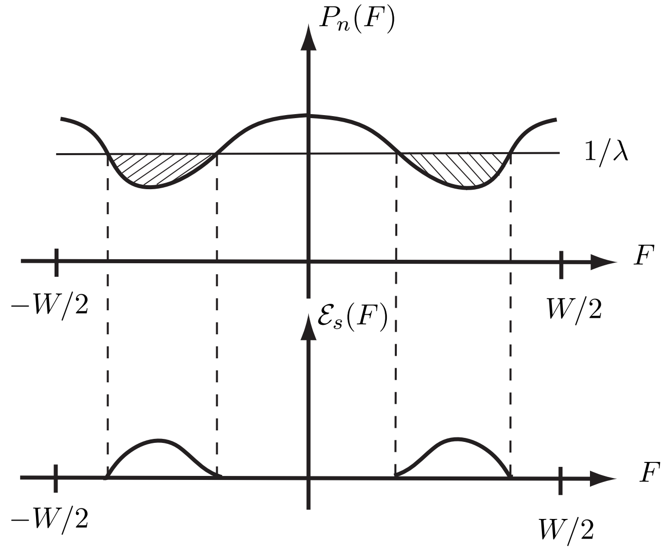

Figure 2: "Water-filling" interpretation of optimal signal design.

and for the energy constraint to hold

$$\int_{-W/2}^{W/2} \frac{\sqrt{N_0/\lambda} - N_0}{P_h(F)} dF = \mathcal{E}$$

or

$$\sqrt{N_0/\lambda} - N_0 = \frac{\mathcal{E}}{\int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF}$$

so that we have finally from (7)

$$\mathcal{E}_s(F) = \frac{\mathcal{E}/P_h(F)}{\int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF}. \tag{8}$$

Note that $\mathcal{E}_s(F) = c/P_h(F)$, where c is a constant, so that the clutter PSD $P_c(F) = P_h(F)\mathcal{E}_s(F) = c$ or the optimal signal is one that *whitens the clutter*. Since the ambient noise is also assumed to be white, this choice of signal results in the detection of a signal in *white noise*.

In this case the performance can be shown from (3) and (8) to be

$$d^2 = \sigma_A^2 \frac{\mathcal{E}}{N_0} \frac{N_0 \int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF}{\mathcal{E} + N_0 \int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF}$$

and noting that $\sigma_A^2 \mathcal{E}/N_0$ is an upper bound on performance that is attained in the case of no clutter (let $P_h(F) = 0$ and $P_n(F) = N_0$ in (3)), we see that the optimal detector with optimal signal design cancels clutter to within a factor of

$$\xi = 10 \log_{10} \frac{N_0 \int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF}{\mathcal{E} + N_0 \int_{-W/2}^{W/2} \frac{1}{P_h(F)} dF} \quad \text{dB}. \tag{9}$$

The intuitive result is that if $P_h(F)$ is small within a band of frequencies, we concentrate our energy in that band (see (8)) and also the detection performance is nearly the upper bound of $\sigma_A^2\mathcal{E}/N_0$ if $\int_{-W/2}^{W/2}\frac{1}{P_h(F)}dF$ is large.

5.2 Manasse's Result

As a special case of the previous result, assume we also have $P_h(F) = k$. This is the case of a clutter PSD that is a scaled replica of the transmitted signal. Then we essentially have the result of Manasse given in [26]. From (8) the optimal signal is just

$$\begin{aligned}\mathcal{E}_s(F) &= \frac{\mathcal{E}/P_h(F)}{\int_{-W/2}^{W/2}\frac{1}{P_h(F)}dF} \\ &= \frac{\mathcal{E}/k}{\int_{-W/2}^{W/2}\frac{1}{k}dF} \\ &= \frac{\mathcal{E}}{W}.\end{aligned}$$

The signal should have a flat ESD over the entire frequency band as reported by Manasse. Also, the degradation is from (9) with $P_h(F) = k$

$$\xi = 10\log_{10}\frac{N_0W/k}{\mathcal{E} + N_0W/k}$$

which approaches zero as $W \to \infty$.

5.3 A Numerical Example

We next illustrate the entire approach by using a numerical example. Consider the case where $P_h(F) = 1$ and $P_n(F) = \exp(-|F|)$ over the frequency band $-1 \le F \le 1$, with the allotted signal energy being $\mathcal{E} = 1/8$. Hence, this corresponds to a clutter return that is identical to the transmit signal and noise that is colored. This example has been chosen mainly for convenience since it leads to a simple analytical solution. However, we will now solve it numerically using MATLAB to illustrate the procedure followed for more complicated noise and clutter PSDs. From (5) the optimal ESD is

$$\mathcal{E}_s(F) = \max\left(\exp(-|F|/2)\frac{1}{\sqrt{\lambda}} - \exp(-|F|), 0\right) \qquad -1 \le F \le 1 \tag{10}$$

and it must integrate to $\mathcal{E} = 1/8$ over the entire frequency band so that

$$\mathcal{E}(\lambda) = \int_{-1}^{1}\max\left(\exp(-|F|/2)\frac{1}{\sqrt{\lambda}} - \exp(-|F|), 0\right)dF = \frac{1}{8}$$

must be solved for λ. As described previously, $0 < \lambda < 1/\min(P_n(F))$. Here, the minimum value of $P_n(F)$ is $P_n(1) = \exp(-1)$ so that $0 < \lambda < \exp(1)$. We can search for this noting that the integral or energy is

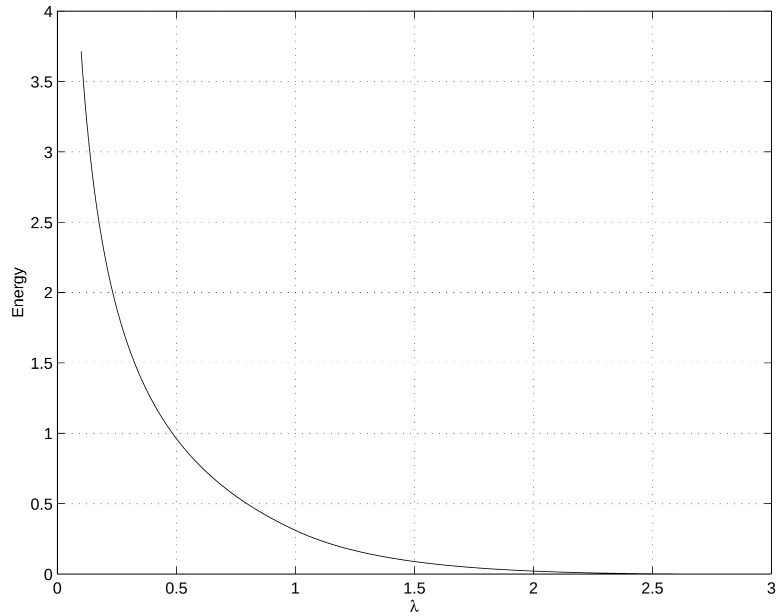

Figure 3: Energy of transmit signal as a function of λ.

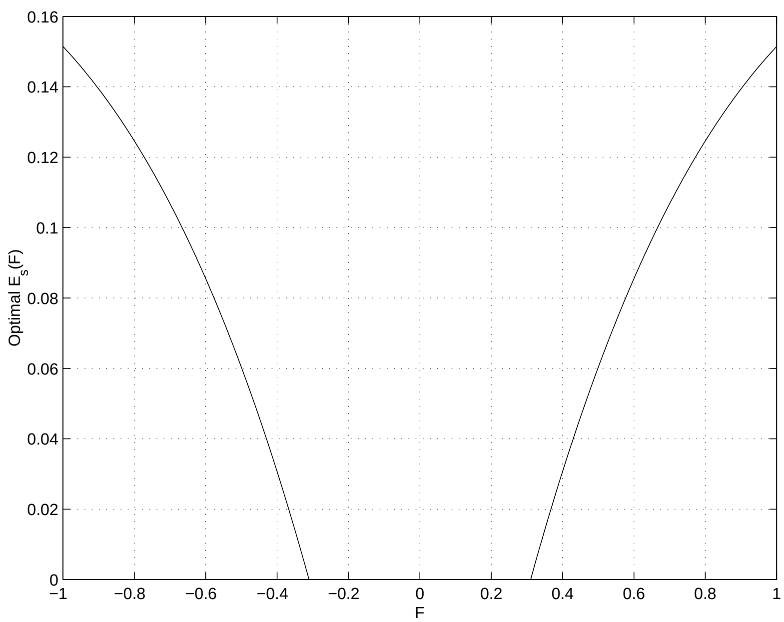

Figure 4: Energy spectral density of optimal transmit signal.

a monotonically decreasing function of λ. For this example, $\mathcal{E}(\lambda)$ is shown in Figure 3 from which it is found that $\lambda = 1.364$. Using this value in (10) produces the optimal signal ESD shown in Figure 4. It is interesting to note that for frequencies below the cutoff frequency of $F = \pm 0.31$, there is no signal energy. This is as expected since the optimal signal will not place energy where the noise PSD $P_n(F) = \exp(-|F|)$

is large. Finally, to verify that this is indeed optimal we compute d^2 for this ESD as well as for an ESD that is flat with frequency. We have from (3) that

$$\frac{d^2}{\sigma_A^2} = \int_{-1}^{1} \frac{\mathcal{E}_s(F)}{\mathcal{E}_s(F) + \exp(-|F|)} dF$$

which is found numerically to be $d^2/\sigma_A^2 = 0.2129$. A flat ESD of $\mathcal{E}_s(F) = \mathcal{E}/W = 1/16$ produces

$$\frac{d^2}{\sigma_A^2} = \int_{-1}^{1} \frac{1/16}{1/16 + \exp(-|F|)} dF$$

which numerically evaluates to $d^2/\sigma_A^2 = 0.1927$.

5.4 A Radar Example

In all cases we will assume that $P_h(F) = k$, which means that the clutter PSD, which is $P_c(F) = P_h(F)\mathcal{E}_s(F)$, is equal to a scaled version of the signal ESD. This is a worst case scenario in that there is no spreading in Doppler of the clutter return. Note that when $k = 0$ the noise background is devoid of clutter.

From (5) the optimal ESD is given by

$$\mathcal{E}_s(F) = T|S(F)|^2 = \frac{1}{k} \max\left(\sqrt{P_n(F)/\lambda} - P_n(F), 0\right)$$

and the maximum detectability index is from (3)

$$d^2_{\text{opt}} = \sigma_A^2 \int_{-W/2}^{W/2} \frac{\mathcal{E}_s(F)}{k\mathcal{E}_s(F) + P_n(F)} dF.$$

As a benchmark to performance consider also a linear FM (LFM), whose ESD is given by

$$\mathcal{E}_s(F) = \frac{\mathcal{E}}{W} \quad |F| \leq W/2.$$

With this signal the detectability index is

$$d^2_{\text{LFM}} = \sigma_A^2 \int_{-W/2}^{W/2} \frac{\mathcal{E}/W}{k\mathcal{E}/W + P_n(F)} dF.$$

We now consider the following scenario. The bandwidth is $W = 5$ Mhz, the signal pulse width is $T = 1\ \mu$ sec, the signal energy is $\mathcal{E} = 10^6$ joules, the clutter constant is $k = 1$, and the noise PSD is $P_n(F) = P_i(F) + N_0$, where $P_i(F)$ is the interference PSD, and the ambient noise PSD is $N_0 = 1$ watts/Hz. (Note that $WT = 5$ was chosen for this example to illustrate typical radar parameters. Although it does not satisfy $WT > 16$, the comparative results will still be correct. Many wideband radar systems do indeed satisfy this requirement.) The interference is composed of three sinusoidal jammers whose PSD in watts/Hz is

$$P_i(F) = P_1 \left(\frac{\sin(\pi(F - F_1)T)}{\pi(F - F_1)T}\right)^4 + P_2 \left(\frac{\sin(\pi(F - F_2)T)}{\pi(F - F_2)T}\right)^4 + P_3 \left(\frac{\sin(\pi(F - F_3)T)}{\pi(F - F_3)T}\right)^4$$

where $P_1 = P_2 = 10^4$ and $P_3 = 10^5$, and the center frequencies are $F_1 = 1$ Mhz, $F_2 = 0.5$ Mhz, and $F_3 = -0.25$ Mhz. The value obtained for the energy constraint parameter is $\lambda = 0.1098$. The optimal ESD is shown in Figure 5 as the solid curve along with the PSD of the interference and ambient noise shown as the dashed curve. It can be seen that as expected the signal energy is placed in the frequency bands where

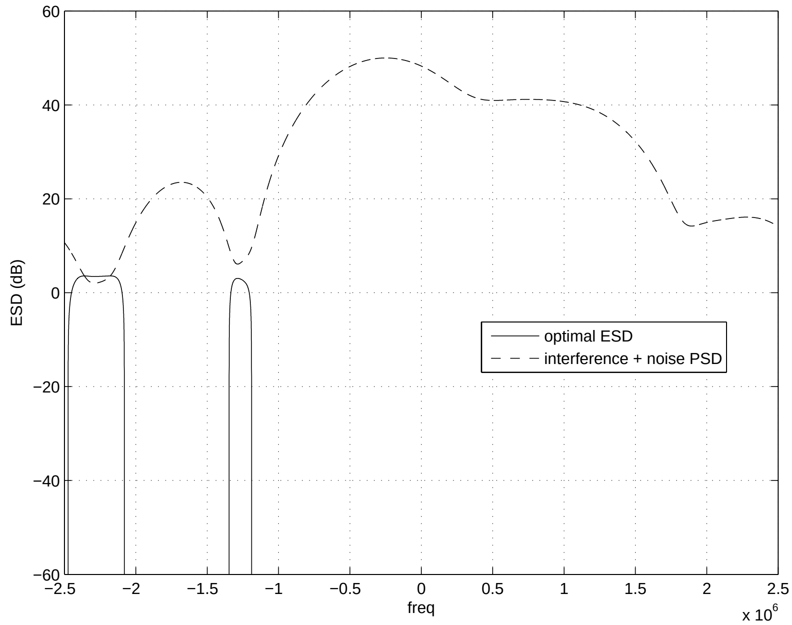

Figure 5: Optimal transmit ESD and interference plus noise PSD.

the interference power is least. Also, note that over much of the band the signal energy is zero. However, all the energy in the band is not placed where the interference PSD is minimum. This would only be the case if the colored noise did not depend on the transmit signal. This is not the case here unless $k = 0$. In fact, it is shown in Appendix D that if $k = 0$, the optimal transmit signal does indeed place all its energy in the frequency band where the colored noise PSD is minimum. To verify this we let $k \to 0$. In this case the results are shown in Figure 6 and we recover the usual result for $k = 0$. Hence, the solution presented here can be viewed as an extension of the usual result [10].

Finally, the detectability indices for $k = 1$ (corresponding to Figure 5) are

$$\begin{aligned} d^2_{\text{opt}} &= 53.0 \quad \text{dB} \\ d^2_{\text{LFM}} &= 46.4 \quad \text{dB} \end{aligned}$$

an improvement over the LFM signal of 6.6 dB. For the case of no clutter or $k = 0$ (corresponding to Figure 6) the detection improvement over the LFM is 11.4 dB since

$$d^2_{\text{opt}} = 58.0 \quad \text{dB}$$

$$d^2_{\text{LFM}} = 46.6 \quad \text{dB}.$$

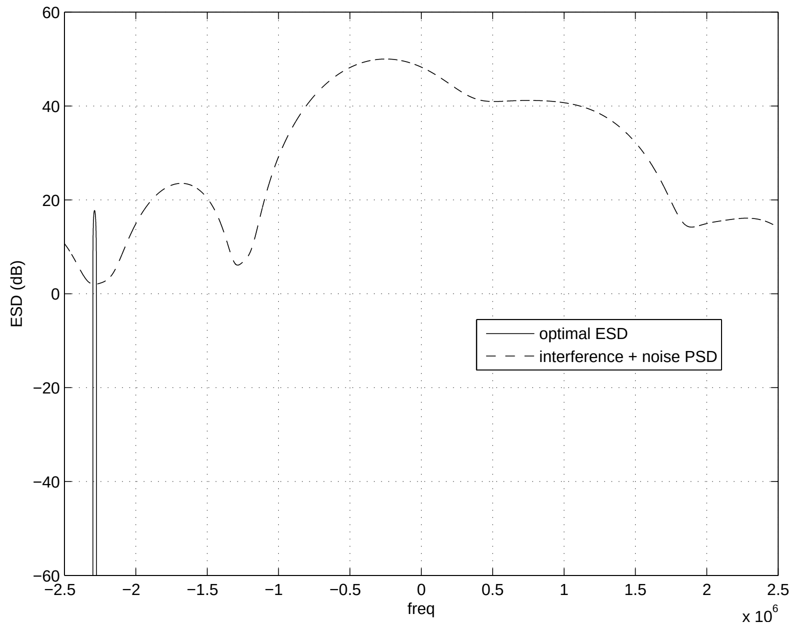

Figure 6: Optimal transmit ESD and interference plus noise PSD in the absence of clutter.

6 A Signal Synthesis Example

We will assume that $P_h(F) = k$, which means that the clutter PSD, which is $P_c(F) = P_h(F)\mathcal{E}_s(F)$, is equal to a scaled version of the signal ESD. From (5) the optimal ESD is given by

$$\mathcal{E}_s(F) = T|S(F)|^2 = \frac{1}{k}\max\left(\sqrt{P_n(F)/\lambda} - P_n(F), 0\right) \tag{11}$$

and the maximum detectability index is from (3)

$$d^2_{\text{opt}} = \sigma_A^2 \int_{-W/2}^{W/2} \frac{\mathcal{E}_s(F)}{k\mathcal{E}_s(F) + P_n(F)} dF.$$

In synthesizing a signal with the optimal ESD we will assume that the radar transmits a pulse train with the mth pulse multiplied by an amplitude $A_m > 0$ and phase shifted by a phase ϕ_m so that the *real* transmitted signal is [23]

$$s_R(t) = \sum_{m=0}^{M-1} A_m \cos(2\pi F_0(t - mT_p) + \phi_m)u(t - mT_p)$$

where F_0 is the radar center frequency, $1/T_p$ is the PRF and $u(t)$ is a square pulse of width τ defined as

$$u(t) = \begin{cases} 1 & 0 \leq t \leq \tau \\ 0 & \text{otherwise} \end{cases}$$

and $\tau \ll T_p$. Alternatively, we can write this as

$$\begin{aligned} s_R(t) &= \text{Re}\left(\sum_{m=0}^{M-1} A_m \exp(j(2\pi F_0(t - mT_p) + \phi_m))u(t - mT_p) \right) \\ &= \text{Re}\left(\sum_{m=0}^{M-1} A_m \exp(j\phi_m) \exp(-j2\pi F_0 mT_p)u(t - mT_p) \exp(j2\pi F_0 t) \right) \end{aligned}$$

and letting $s[m] = A_m \exp(j\phi_m)$ be the "slow-time" signal sequence and assuming F_0T_p is an integer, we have that

$$s(t) = \sum_{m=0}^{M-1} s[m]u(t - mT_p)$$

is the complex envelope of the transmit signal. If it is assumed that the received waveform is sampled at the PRF rate (after matched filtering of the pulse), then the received discrete signal will be just $s[m]$. Hence, we wish to choose $s[m]$ so that the ESD of $s(t)$ is given by (11) over the normalized frequency range $-1/2 \leq F/(1/T_p) \leq 1/2$. It is assumed that $W = 1/T_p$.

We now consider the following scenario. The bandwidth is $W = 5000$ Hz so that the PRF is $1/T_p = 5000$ pulses per second, the signal pulse width is $\tau = 1$ μ sec, the signal energy is $\mathcal{E} = 10^4$ joules, the clutter constant is $k = 1$, and the noise PSD is $P_n(F) = P_i(F) + N_0$, where $P_i(F)$ is the interference PSD and the ambient noise PSD is $N_0 = 1$ watts/Hz. The interference is composed of three Gaussian jammers whose PSD is

$$P_i(F) = \sum_{i=1}^{3} P_i \exp\left[-(1/(2B))(F - F_i)^2\right]$$

where $P_1 = P_2 = 100$ and $P_3 = 1000$, and the center frequencies are $F_1 = 1000$ Hz, $F_2 = 500$ Hz, and $F_3 = -250$ Hz, and $B = 10^4$. The PSD of the jammers and ambient noise is shown in Figure 7 as the dashed curve. The value of the energy constraint parameter is found to be $\lambda = 0.0699$. The ESD of the optimal signal is shown in Figure 7 as the solid curve. It can be seen that as expected the signal energy is placed in the frequency bands where the interference power is least. Also, note that over much of the band the signal energy is zero. However, it does not place all the energy in the band where the interference PSD is minimum.

To synthesize $s[m]$ we use Durbin's method as described in [20, 24]. Although it is normally used to estimate the parameters of a moving average random process based on a set of time series data, it can also be used for signal synthesis. An advantage of Durbin's method is that it produces a minimum phase signal. Hence, most of the energy is concentrated up front, making truncation of the sequence less problematic.

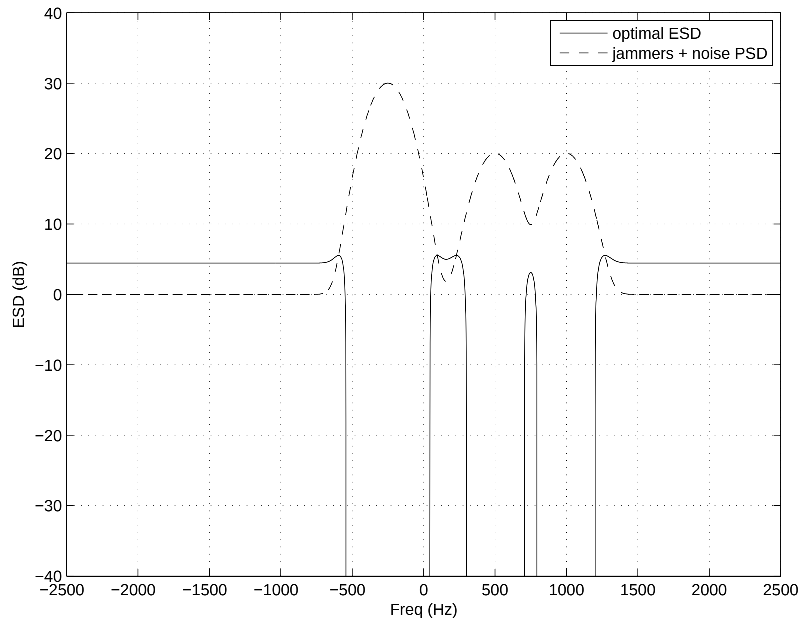

Figure 7: Optimal transmit ESD and jammers plus noise PSD.

To implement Durbin's method we first take the inverse Fourier transform of the desired ESD to obtain an autocorrelation function. The autocorrelation samples are used in the Levinson algorithm to solve for a set of autoregressive parameters using a large autoregressive model order. Then the autoregressive samples are used as data, the autocorrelation function estimated, and finally the solution of the Yule-Walker equations provides the moving average parameters. These moving average parameters are our $s[m]$ samples. The approximation using Durbin's method for $M = 200$ yields $s[m]$ as shown in Figure 8. Note that since $s[m]$ is complex we plot only the magnitude in the figure. Its ESD is shown in Figure 9 and is seen to closely match the optimal one of Figure 7.

7 Conclusions

A new method has been proposed to design signals for optimal detection performance in signal-dependent noise. If the channel power spectral density is known, then coupled with knowledge of the noise power spectral density, it has been shown how to synthesis the transmit waveform. The actual gain realized by invoking this procedure will be dependent upon the environmental conditions. Extensions to the case of an extended target and/or spatial processing will be discussed in a future paper.

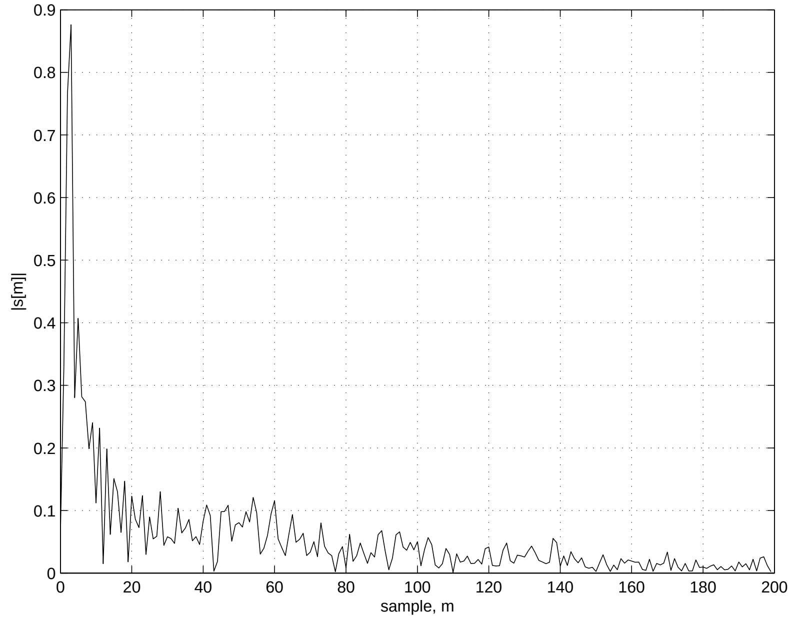

Figure 8: Magnitude of complex weighting for transmit signal pulses.

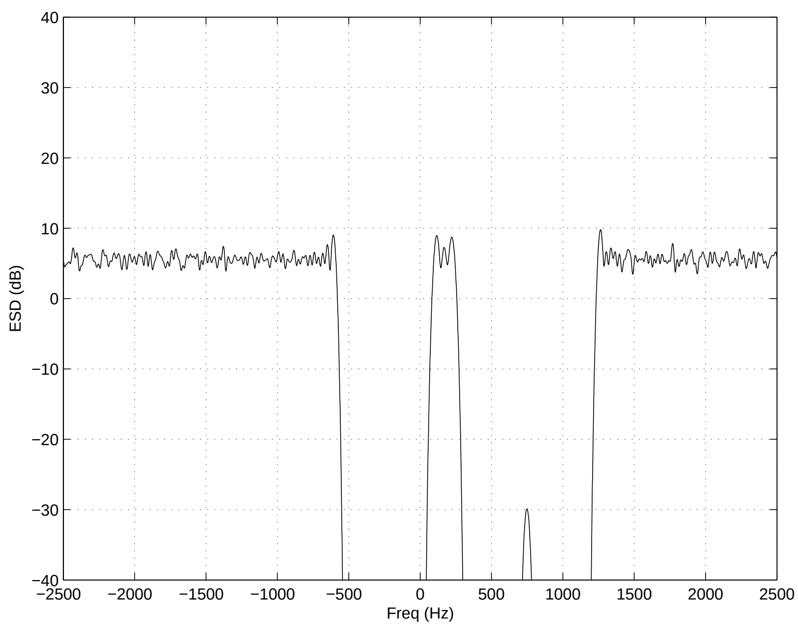

Figure 9: Energy spectral density of synthesized transmit signal.

Acknowledgments

Thanks are due to Muralidhar Rangaswamy of the Air Force Research Laboratory, Hanscom AFB, MA for useful discussions of this work.

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. 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.

12. Kay, S., J.H. Thanos,"Optimal Transmit Signal Design for Active Sonar/Radar", 2002 Int. Conf. on Acoustics, Speech, and Signal Processing, Orlando, FL.

13. Bell, M.R., "Information Theory and Radar Waveform Design", *IEEE Trans. on Info. Theory*, pp. 1578–1597, Sept. 1993.

14. 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.

15. Fishler, E., A. Haimovich, R.S. Blum, L.J. Cimini, Jr., D. Chizhik, R.A. Valenzuela, "Spatial Diversity in Radars–Models and Detection Performance", *IEEE Trans. on Signal Processing*, pp. 823–838, March 2006.

16. Kay, S., S.B. Doyle, "Rapid Estimation of the Range-Doppler Scattering Function", *IEEE Trans. on Signal Processing*, pp. 255-268, Jan. 2003.

17. 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.

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

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

20. Jackson, L.B., *Digital Filters and Signal Processing*, Springer, NY, 1995.

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

22. Gallagher, R.B., *Information Theory and Reliable Communication*, pp. 383–389, J. Wiley, NY, 1968.

23. Richards, M., *Fundamentals of Radar Signal Processing*, McGraw-Hill, NY, 2005.

24. Kay, S., *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

25. Luenberger, D.E., *Optimization by Vector Space Methods*, J. Wiley, NY, 1969.

26. Manasse, R., "The Use of Pulse Coding to Discriminate Against Clutter", MIT Lincoln Lab, Lexington, MA, Rept. 312–12, August 1957.

A Derivation of Neyman-Pearson Detector and its Performance

As per the assumptions described in Section 2 we consider the following detection problem.

$$\begin{aligned} \mathcal{H}_0 &: x(t) = c(t) + n(t) \\ \mathcal{H}_1 &: x(t) = As(t) + c(t) + n(t) \end{aligned}$$

for $-T/2 \leq t \leq T/2$. The assumptions are that $s(t)$ is a known complex signal, A is a complex random variable with $A \sim \mathcal{CN}(0, \sigma_A^2)$, $c(t)$ is a complex WSS Gaussian random process with zero mean and PSD $P_c(F) = P_h(F)T|S(F)|^2 = P_h(F)\mathcal{E}_s(F)$, and $n(t)$ is a complex WSS Gaussian random process with zero mean and PSD $P_n(F)$. The random variable A, the random processes $c(t)$ and $n(t)$ are all independent of each other. We convert the received data into the frequency domain using a Fourier transform as defined by (2) and using the *frequency snapshot* model [21]. For $WT > 16$ we can assert that the frequency samples $X(F_m)$ for $F_m = m/T$ and $m = -(M/2), \ldots, M/2$ are all independent. They are also complex Gaussian random variables with zero mean and variance equal to the PSD value, which is $P_x(F_m)$. Hence, after Fourier transforming we obtain the $(M+1) \times 1$ complex vector

$$\mathbf{X} = [X(F_{-M/2}) \ldots X(F_{M/2})]^T$$

so that the equivalent detection problem is

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X} = \mathbf{C} + \mathbf{N} \\ \mathcal{H}_1 &: \mathbf{X} = A\mathbf{S} + \mathbf{C} + \mathbf{N} \end{aligned}$$

where all complex vectors are $(M+1) \times 1$ vectors of the Fourier transform samples. Furthermore, because of the statistical properties of the vectors, being all complex Gaussian random vectors and all independent of each other, we have

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_0) \\ \mathcal{H}_1 &: \mathbf{X} \sim \mathcal{CN}(\mathbf{0}, \sigma_A^2 \mathbf{S}\mathbf{S}^H + \mathbf{C}_0) \end{aligned}$$

where $\mathbf{C}_0$ is the covariance matrix of $\mathbf{C} + \mathbf{N}$. We will give this explicitly later. We now derive the Neyman-Pearson detector for this problem. A similar one can be found in [10] on page 481, where the noise is assumed to be white. Hence, this is a slight extension of that result, which is termed the *rank one signal covariance matrix*. The PDFs under either hypothesis are complex multivariate Gaussian so that

$$\begin{aligned} p(\mathbf{X}; \mathcal{H}_0) &= \frac{1}{\pi^{M+1}|\det(\mathbf{C}_0)|} \exp(-\mathbf{X}^H \mathbf{C}_0^{-1} \mathbf{X}) \\ p(\mathbf{X}; \mathcal{H}_1) &= \frac{1}{\pi^{M+1}|\det(\mathbf{C}_1)|} \exp(-\mathbf{X}^H \mathbf{C}_1^{-1} \mathbf{X}) \end{aligned}$$

where $\mathbf{C}_1 = \sigma_A^2 \mathbf{S}\mathbf{S}^H + \mathbf{C}_0$. The log-likelihood ratio is

$$\begin{aligned} l(\mathbf{X}) &= \ln \frac{p(\mathbf{X}; \mathcal{H}_1)}{p(\mathbf{X}; \mathcal{H}_0)} \\ &= \underbrace{\ln|\det(\mathbf{C}_0)| - \ln|\det(\mathbf{C}_1)|}_{c} + \mathbf{X}^H(\mathbf{C}_0^{-1} - \mathbf{C}_1^{-1})\mathbf{X} \end{aligned}$$

and ignoring the constant c, we need to simplify the difference of the inverse covariance matrices. But using Woodbury's identity we have

$$\begin{aligned} (\mathbf{C}_0^{-1} - \mathbf{C}_1^{-1}) &= \mathbf{C}_0^{-1} - \left(\mathbf{C}_0 + \sigma_A^2 \mathbf{S}\mathbf{S}^H\right)^{-1} \\ &= \mathbf{C}_0^{-1} - \mathbf{C}_0^{-1} + \frac{\mathbf{C}_0^{-1}\sigma_A^2 \mathbf{S}\mathbf{S}^H \mathbf{C}_0^{-1}}{1 + \sigma_A^2 \mathbf{S}^H \mathbf{C}_0^{-1}\mathbf{S}} \end{aligned}$$

and hence we have

$$\begin{aligned} l(\mathbf{X}) &= \mathbf{X}^H \frac{\mathbf{C}_0^{-1}\sigma_A^2 \mathbf{S}\mathbf{S}^H \mathbf{C}_0^{-1}}{1 + \sigma_A^2 \mathbf{S}^H \mathbf{C}_0^{-1}\mathbf{S}} \mathbf{X} \\ &= \sigma_A^2 \frac{|\mathbf{X}^H \mathbf{C}_0^{-1}\mathbf{S}|^2}{1 + \sigma_A^2 \mathbf{S}^H \mathbf{C}_0^{-1}\mathbf{S}}. \end{aligned}$$

Again the constants may be ignored so that finally we have the test statistic

$$T(\mathbf{X}) = |\mathbf{X}^H \mathbf{C}_0^{-1}\mathbf{S}|^2$$

which is recognized as a prewhitener/matched filter, although in the frequency domain. It can be simplified further if we recall that the elements of $\mathbf{C}_0$ are

$$\begin{aligned} [\mathbf{C}_0]_{mn} &= E[X(F_m)X^*(F_n)] \\ &\approx \begin{cases} 0 & m \neq n \\ P_x(F_m) & m = n. \end{cases} \end{aligned}$$

Hence, $\mathbf{C}_0$ is diagonal with diagonal elements $[\mathbf{C}_0]_{mm} = P_x(F_m) = P_c(F_m) + P_n(F_m)$, and we have that

$$\begin{aligned} T(\mathbf{X}) &= |\mathbf{X}^H \mathbf{C}_0^{-1}\mathbf{S}|^2 \\ &= |\mathbf{S}^H \mathbf{C}_0^{-1}\mathbf{X}|^2 \\ &= \left| \sum_{m=-(M/2)}^{M/2} \frac{S^*(F_m)X(F_m)}{P_c(F_m) + P_n(F_m)} \right|^2 \\ &= \left| \sum_{m=-(M/2)}^{M/2} \frac{X(F_m)S^*(F_m)}{P_h(F_m)T|S(F_m)|^2 + P_n(F_m)} \right|^2 \end{aligned}$$

which is (1). The performance of this detector is next derived. We first state a general result, which is derived in Appendix B. It states that if X is a complex random variable with $X \sim \mathcal{CN}(0, \sigma^2)$, then

$P[|X|^2 > \gamma] = \exp(-\gamma/\sigma^2)$. Furthermore, if only the variance changes under the two hypotheses with σ_i^2 denoting the variance under $\mathcal{H}_i$, it is also shown in Appendix B that the probability of detection P_D is related to the probability of false alarm P_{FA} by

$$P_D = P_{FA}^{\sigma_0^2/\sigma_1^2}. \tag{12}$$

We apply this result by determining the PDF of $X = \mathbf{X}^H \mathbf{C}_0^{-1} \mathbf{S}$ under both hypotheses. Since $\mathbf{X}$ is a complex Gaussian random vector, any linear transformation of it produces another complex Gaussian random variable [18]. The mean of X under either hypothesis is zero since $\mathbf{X}$ is zero mean. Hence, we need only determine its variance or equivalently its second moment. Considering first X under $\mathcal{H}_0$ we have

$$\begin{aligned}
\sigma_0^2 &= E[|X|^2] \\
&= E[|\mathbf{X}^H \mathbf{C}_0^{-1} \mathbf{S}|^2] \\
&= E[\mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{X} \mathbf{X}^H \mathbf{C}_0^{-1} \mathbf{S}] \\
&= \mathbf{S}^H \mathbf{C}_0^{-1} E[\mathbf{X} \mathbf{X}^H] \mathbf{C}_0^{-1} \mathbf{S} \\
&= \mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{C}_0 \mathbf{C}_0^{-1} \mathbf{S} \\
&= \mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}.
\end{aligned}$$

Using the diagonal nature of $\mathbf{C}_0$, we have that

$$\begin{aligned}
\sigma_0^2 &= \mathbf{S}^H \text{diag}^{-1}(P_c(F_m) + P_n(F_m)) \mathbf{S} \\
&= \sum_{m=-M/2}^{M/2} \frac{|S(F_m)|^2}{P_c(F_m) + P_n(F_m)}.
\end{aligned}$$

Similarly under $\mathcal{H}_1$ we have that

$$\begin{aligned}
\sigma_1^2 &= \mathbf{S}^H (\mathbf{C}_0^{-1} (\mathbf{C}_0 + \sigma_A^2 \mathbf{S} \mathbf{S}^H) \mathbf{C}_0^{-1} \mathbf{S} \\
&= \sigma_0^2 + \underbrace{\sigma_A^2 |\mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}|^2}_{\Delta}.
\end{aligned}$$

Now

$$\begin{aligned}
\frac{\sigma_0^2}{\sigma_1^2} &= \frac{\sigma_0^2}{\sigma_0^2 + \Delta} \\
&= \frac{1}{1 + \Delta/\sigma_0^2}
\end{aligned}$$

and from (12)

$$P_D = P_{FA}^{\frac{1}{1+\Delta/\sigma_0^2}}.$$

But P_D is monotonically increasing with Δ/σ_0^2 for a fixed P_{FA} so that we can quantify our detection performance by defining

$$d^2 = \frac{\Delta}{\sigma_0^2} = \frac{\sigma_A^2 |\mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}|^2}{\mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}}.$$

Since $\mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}$ is a real positive quantity (just take the hermitian transpose) we have finally that

$$d^2 = \sigma_A^2 \mathbf{S}^H \mathbf{C}_0^{-1} \mathbf{S}$$

as our figure of merit. This is explicitly evaluated as

$$d^2 = \sigma_A^2 \sum_{m=-M/2}^{M/2} \frac{|S(F_m)|^2}{P_c(F_m) + P_n(F_m)}$$

and can be approximated as follows

$$\begin{aligned} d^2 &= \sigma_A^2 \frac{1}{\Delta F} \sum_{m=-M/2}^{M/2} \frac{|S(F_m)|^2}{P_c(F_m) + P_n(F_m)} \Delta F \\ &\approx \sigma_A^2 \frac{1}{\Delta F} \int_{-W/2}^{W/2} \frac{|S(F)|^2}{P_c(F) + P_n(F)} dF \\ &= \sigma_A^2 \frac{1}{\Delta F} \int_{-W/2}^{W/2} \frac{|S(F)|^2}{P_h(F) T |S(F)|^2 + P_n(F)} dF \\ &= \sigma_A^2 \int_{-W/2}^{W/2} \frac{T|S(F)|^2}{P_h(F) T |S(F)|^2 + P_n(F)} dF \end{aligned}$$

where $\Delta F = 1/T$ and T is assumed large, and hence producing (3).

B Derivation of Detection Performance

Assume that we have $X \sim \mathcal{CN}(0, \sigma_0^2)$ under $\mathcal{H}_0$ and $X \sim \mathcal{CN}(0, \sigma_1^2)$ under $\mathcal{H}_1$. The detector decides $\mathcal{H}_1$ if $|X|^2 > \gamma$. We derive the performance of this detector. Note that for a complex Gaussian random variable with variance σ^2 the real and imaginary parts each have a variance of $\sigma^2/2$ and are independent. Hence, we have with $X = U + jV$

$$\begin{aligned} P_{FA} &= P[|X|^2 > \gamma; \mathcal{H}_0] \\ &= P[U^2 + V^2 > \gamma; \mathcal{H}_0] \\ &= P\left[\left(\frac{U}{\sigma_0/\sqrt{2}}\right)^2 + \left(\frac{V}{\sigma_0/\sqrt{2}}\right)^2 > 2\gamma/\sigma_0^2; \mathcal{H}_0\right] \\ &= P[N_1^2 + N_2^2 > 2\gamma/\sigma_0^2; \mathcal{H}_0] \end{aligned}$$

where N_1 and N_2 are independent $N(0,1)$ random variables. Hence, the sum of their squares is a χ_2^2 random variable for which

$$P_{FA} = \int_{2\gamma/\sigma_0^2}^{\infty} \frac{1}{2}\exp(-x/2)dx = \exp(-\gamma/\sigma_0^2).$$

Similarly, we have that

$$P_D = \exp(-\gamma/\sigma_1^2)$$

and eliminating the threshold γ produces the desired result of

$$P_D = P_{FA}^{\frac{\sigma_0^2}{\sigma_1^2}}.$$

C Derivation of Optimal Signal Energy Spectral Density

The optimal signal ESD is given by the frequency function $|S(F)|^2$ that maximizes (3). As before we let $\mathcal{E}_s(F) = T|S(F)|^2$ and maximize

$$I = \int_{-W/2}^{W/2} \frac{\mathcal{E}_s(F)}{P_h(F)\mathcal{E}_s(F) + P_n(F)} dF$$

subject to the energy constraint

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

and the nonnegativity constraint $\mathcal{E}_s(F) \geq 0$ for $-W/2 \leq F \leq W/2$. Because I is a concave functional a local maximum as determined by a differentiation will also be a global maximum. Note that the constraint of energy is a linear one (and hence the region is convex) and the constraint of a nonnegative function also produces a convex region so that this is a convex progamming problem. It is well known that the maximum can be found by maximizing the Lagrangian [25]. Hence, we form

$$L(\mathcal{E}_s(F)) = \int_{-W/2}^{W/2} \frac{\mathcal{E}_s(F)}{P_h(F)\mathcal{E}_s(F) + P_n(F)} dF - \lambda \left(\int_{-W/2}^{W/2} \mathcal{E}_s(F)dF - \mathcal{E} \right)$$

or

$$L(\mathcal{E}_s(F)) = \int_{-W/2}^{W/2} \left(\frac{\mathcal{E}_s(F)}{P_h(F)\mathcal{E}_s(F) + P_n(F)} - \lambda\mathcal{E}_s(F) \right) dF + \lambda\mathcal{E}. \qquad (13)$$

We note in passing that the actual energy constraint is such that

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

However, it can be shown that I is monotonically increasing with $\mathcal{E}$ and hence we can replace the inequality by equality. This also says that for the maximum value of I

$$\frac{\partial I_{\max}}{\partial \mathcal{E}} > 0$$

and since it is well known that

$$\lambda = \frac{\partial I_{\max}}{\partial \mathcal{E}}$$

this also asserts that the Lagrangian multiplier must be positive. In the course of the derivation we will establish that I is a concave functional. Since we are free to choose any value of $\mathcal{E}_s(F)$ for each F in (13), as long as it is nonnegative, we can equivalently maximize

$$G(\mathcal{E}_s(F)) = \frac{\mathcal{E}_s(F)}{P_h(F)\mathcal{E}_s(F) + P_n(F)} - \lambda \mathcal{E}_s(F)$$

for each F over $\mathcal{E}_s(F) \geq 0$. We let $x = \mathcal{E}_s(F)$, $\alpha = P_h(F) > 0$, and $\beta = P_n(F) > 0$, so that we wish to maximize

$$g(x) = \frac{x}{\alpha x + \beta} - \lambda x.$$

This function is strictly concave on the interval $[0, \infty)$ as we now show by differentiation. The first derivative is

$$g'(x) = \frac{\beta}{(\alpha x + \beta)^2} - \lambda \tag{14}$$

and the second derivative is

$$g''(x) = -\frac{2\alpha\beta}{(\alpha x + \beta)^3} < 0.$$

Since the second derivative is negative for $x \geq 0$, $g(x)$ is concave. This also shows that I is a concave functional since it is a "sum" of concave functions. To determine the maximizing value of $g(x)$, we set the first derivative equal to zero to yield

$$\frac{\beta}{(\alpha x + \beta)^2} - \lambda = 0$$

and solving for x produces

$$x_0 = \frac{\sqrt{\beta/\lambda} - \beta}{\alpha}$$

as long as x_0 is nonnegative. If x_0 as given above is nonnegative, then the positivity constraint is not binding and hence this is the solution. When x_0 is negative, however, the constraint is binding and therefore, g is maximized over $[0, \infty)$ by $x = 0$. This is because if x_0 is negative, then

$$\sqrt{\beta/\lambda} - \beta < 0$$

which is equivalent to $\lambda > 1/\beta$. But from (14) we have for $\lambda > 1/\beta$ and $x > 0$

$$\begin{aligned} g'(x) &= \frac{\beta}{(\alpha x + \beta)^2} - \lambda \\ &< \frac{\beta}{\beta^2} - \lambda = \frac{1}{\beta} - \lambda \\ &< 0 \end{aligned}$$

and therefore $g(x)$ is monotonically decreasing for $x > 0$. It is clear then that the maximum over the interval $[0, \infty)$ is at $x = 0$. We have then that

$$x = \max\left(\frac{\sqrt{\beta/\lambda} - \beta}{\alpha}, 0\right)$$

and finally

$$\mathcal{E}_s(F) = \max\left(\frac{\sqrt{P_n(F)/\lambda} - P_n(F)}{P_h(F)}, 0\right)$$

by substituting in the definitions for α and β.

D Recovery of Optimal Signal for the Case of No Clutter

In order to recover the non-signal-dependent noise case we will let $P_h(F) = k$ in (5) and (6) and then let $k \to 0$. First multiplying both sides of (6) by $k > 0$, we have for the energy constraint

$$\int_{-W/2}^{W/2} \max\left(\sqrt{P_n(F)/\lambda} - P_n(F), 0\right) dF = k\mathcal{E}.$$

Next note that the integrand is nonnegative and monotonically decreasing with λ. It becomes zero when $1/\lambda = \min(P_n(F))$. To simplify the discussion we assume that the minimum of $P_n(F)$ occurs at a single frequency. If we let $1/\lambda = \min(P_n(F)) + \epsilon$ for $\epsilon > 0$ (the necessary value of λ as $k \to 0$), then the integrand as $k \to 0$ is

$$\max\left(\sqrt{P_n(F)(\min(P_n(F)) + \epsilon)} - P_n(F), 0\right).$$

Letting $\alpha = \min(P_n(F)) + \epsilon$ and $x = P_n(F)$, then

$$\sqrt{P_n(F)(\min(P_n(F)) + \epsilon)} - P_n(F) = \sqrt{\alpha x} - x.$$

It is clear that for

$$\sqrt{\alpha x} - x \leq 0$$

we must have $x \geq \alpha$. Thus, the integrand will be zero if $P_n(F) \geq \min(P_n(F)) + \epsilon$ and since ϵ was arbitrary, the integrand will be zero if $P_n(F) > \min(P_n(F))$. This says that as $k \to 0$, the only part of the frequency band that will contribute to the energy is the frequency band for which $P_n(F) \leq \min(P_n(F))$ and hence all the signal energy will be concentrated at the frequency for which $P_n(F)$ is minimum. This is the classical result.