The Optimum Radar Signal for Detection in Clutter

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

June 7, 2006

Abstract

It is shown that in order to maximize the detectability of a radar target in clutter whose Doppler is unknown and is uniformly distributed over the Doppler bandwidth a simple CW or narrowband signal is optimal. The optimality criterion is the average deflection coefficient, with the averaging being over target Doppler frequency. Most remarkably the result does not depend on the clutter spectrum but holds for any distribution of clutter energy with frequency.

1 Introduction

For radar detection in clutter it is normally assumed that low-Doppler targets are best detected with a wideband waveform while high-Doppler targets require a narrowband signal. Usually the Doppler frequency due to target motion is unknown. It is therefore customary to use an average figure of merit to assess detection performance. An example is the improvement factor for an MTI radar in which the output target-to-clutter ratio to the input target-to-clutter ratio is averaged uniformly over all Doppler frequencies to produce an overall measure of detectability [1]. We show in this paper that *somewhat surprisingly, the optimal transmit signal for detectability, using an averaged (over target Doppler) deflection coefficient as the performance measure, is the sinusoid. In addition, this result holds for all clutter spectra.* In the case of a pulsed Doppler radar a similar result is that for a sequence of pulses the optimal modulation is no modulation at all. In "slow-time" [4] the optimal sampled transmit waveform is therefore a discrete sinusoid. Note that this result assumes that the optimal detector as described in [2] is used in which the target Doppler shift is known and therefore, the optimal prewhitener/matched filter is matched to the known Doppler shift. In practice, for an unknown Doppler shift the performance will be degraded slightly due to a necessary search over Doppler bins [3]. However, it is expected that the same result will hold.

For the detection of a radar signal in clutter and noise the optimal detector is a prewhitener followed by a matched filter. The clutter, however, is a signal-dependent noise since the clutter spectrum depends on the transmit signal. It is well known that for a slowly fluctuating point target the optimal receiver for a radar is given by the filter with frequency response [2]

$$H(F) = \frac{S^*(F - F_d)}{E_s(F) \star P_d(F) + N_0} \qquad -W/2 \leq F \leq W/2 \tag{1}$$

where F is the frequency in Hz, $S(F)$ is the Fourier transform of the complex envelope of the transmitted signal, F_d is the target Doppler shift in Hz, $E_s(F) = |S(F)|^2$ is the energy spectral density (ESD) of the transmitted signal, $P_d(F)$ is the Doppler scattering function, $\star$ indicates convolution, N_0 is the noise power spectral density (PSD) of the received complex envelope noise, and W is the baseband bandwidth in Hz. This optimal receiver assumes a constant Doppler profile in range and an infinite range. The detection performance of this receiver is given by the probability of detection P_D with a probability of false alarm P_{FA} as

$$P_D = P_{FA}^{\frac{1}{1+\Delta_0(F_d)}} \tag{2}$$

where the *deflection coefficient* is defined as

$$\Delta_0(F_d) = \frac{\bar{E}_r}{\mathcal{E}} \int_{-W/2}^{W/2} \frac{E_s(F - F_d)}{E_s(F) \star P_d(F) + N_0} dF \tag{3}$$

and $\mathcal{E}$ is the energy of the transmitted signal while $\bar{E}_r$ is the average received signal energy ($\bar{E}_r = 2\sigma_b^2 \mathcal{E}$, where σ_b^2 models the effect of reflection from the slowly fluctuating point target). The detection performance as quantified by P_D is monotonically increasing with $\Delta_0(F_d)$. Note that the difficulty in detecting a signal in clutter is the degradation due to the term $E_s(F) \star P_d(F)$ in the denominator, which is the clutter PSD. The signal design problem is to choose $E_s(F)$ so as to maximize $\Delta_0(F_d)$ subject to an energy and bandwidth constraint. It is assumed that the bandwidth of the transmitted signal is B_s or that its ESD is concentrated in the frequency band $-B_s/2 \leq F \leq B_s/2$ and also that the maximum Doppler shift is $\pm F_{d_{\max}}$. Hence, the bandwidth W of the baseband processor is found from $W/2 = B_s/2 + F_{d_{\max}}$, which is chosen to accommodate the received signal with a maximum Doppler shift. For the purposes of this analysis we will assume that the signal may be infinite in length. Of course, in practice this is only approximate. (This assumption is theoretically required for (3) to hold although the approximation will be a good one if the time-bandwidth product exceeds 16 [7].) Also, note that we are not considering any subsidiary design criteria such as range or Doppler resolution but only detectability. In the next section we set up the problem and give its solution.

2 Optimal Signal Design

The problem we address is the maximization of $\Delta_0(F_d)$, *when averaged uniformly over target Doppler* F_d, by choosing $E_s(F)$ subject to the energy constraint. Hence, dropping the $\bar{E}_r/\mathcal{E}$ in (3), which does not affect the solution, we wish to maximize

$$\begin{aligned} I(E_s) &= \int_{-F_{d_{\max}}}^{F_{d_{\max}}} \Delta_0(F_d) \frac{1}{2F_{d_{\max}}} dF_d \\ &= \int_{-F_{d_{\max}}}^{F_{d_{\max}}} \left[\int_{-(B_s/2+F_{d_{\max}})}^{B_s/2+F_{d_{\max}}} \frac{E_s(F-F_d)}{E_s(F) \star P_d(F) + N_0} dF \right] \frac{1}{2F_{d_{\max}}} dF_d \end{aligned} \tag{4}$$

subject to the constraints

$$\int_{-B_s/2}^{B_s/2} E_s(F) dF = \mathcal{E} \tag{5}$$

and $E_s(F) \geq 0$. Note that the phase of the signal can be chosen arbitrarily since it does not affect detectability. It is usually chosen for realizability and other design criteria such as range resolution.

Next interchanging integrals in (4) we have

$$\begin{aligned} I(E_s) &= \int_{-(B_s/2+F_{d_{\max}})}^{B_s/2+F_{d_{\max}}} \frac{1}{2F_{d_{\max}}} \int_{-F_{d_{\max}}}^{F_{d_{\max}}} E_s(F-F_d) dF_d \frac{1}{E_s(F) \star P_d(F) + N_0} dF && (6) \\ &\leq \int_{-(B_s/2+F_{d_{\max}})}^{B_s/2+F_{d_{\max}}} \frac{1}{2F_{d_{\max}}} \underbrace{\int_{-(B_s/2+F_{d_{\max}})}^{B_s/2+F_{d_{\max}}} E_s(F-F_d) dF_d}_{I} \frac{1}{E_s(F) \star P_d(F) + N_0} dF && (7) \end{aligned}$$

The integral denoted by I is equal to $\mathcal{E}$. We will show later that the optimal signal has a bandwidth of zero or $B_s = 0$ so that $W/2 = B_s/2$. As a result, we can maximize the upper bound on $I(E_s)$ to obtain the maximum of $I(E_s)$. Hence, the criterion to be maximized becomes the upper bound

$$J(E_s) = \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{E_s(F) \star P_d(F) + N_0} dF \tag{8}$$

subject to the energy and nonnegativity constraints. But it is shown in Appendix A that the functional $J(E_s)$ is convex so that for $0 < \alpha < 1$, we must have

$$J(\alpha E_{s_1} + (1-\alpha) E_{s_2}) \leq \alpha J(E_{s_1}) + (1-\alpha) J(E_{s_2}). \tag{9}$$

This says that any signal ESD that can be decomposed as a convex combination of ESDs will have a detectability index that is less than or equal to $\max(E_{s_1}, E_{s_2})$. It is well known that the maximum of a convex functional is at an extreme point [5] or at a point that *cannot* be decomposed as in (9).

The only ESD that cannot be decomposed as in (9) is shown in Appendix B to be one that is concentrated at a single frequency. Hence, the signal that maximizes $J(E_s)$ subject to the energy and nonnegativity constraints is

$$E_{s_{\text{opt}}}(F) = \mathcal{E} \delta(F - F_0)$$

for any F_0 such that $F_0 \leq B_s/2$, where $\delta(\cdot)$ is the Dirac delta function. This ESD thus corresponds to a sinusoidal signal of frequency F_0. To determine the optimal frequency we substitute $E_{s_{\text{opt}}}(F)$ into (8) to yield

$$\begin{aligned} J(E_s) &= \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{\mathcal{E}\delta(F-F_0) \star P_d(F) + N_0} dF \\ &= \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{\mathcal{E}P_d(F-F_0) + N_0} dF. \end{aligned} \tag{10}$$

The optimal signal is therefore sinusoidal with a frequency F_0 that maximizes

$$\int_{-W/2}^{W/2} \frac{1}{\mathcal{E}P_d(F-F_0) + N_0} dF. \tag{11}$$

We next make the reasonable assumption that the scattering function $P_D(F)$ is concentrated on the frequency interval $-F_{d_{\max}} \leq F \leq F_{d_{\max}}$. This says that the spreading due to Doppler of the clutter, which typically consists of ground clutter, bird clutter, and weather will be less the maximum Doppler shift of the target. As a result $P_D(F-F_0)$ will be concentrated in the baseband and hence, the choice of F_0 in (11) is arbitrary. We choose the transmit frequency of the complex envelope as $F_0 = 0$. As alluded to earlier, the optimal signal has zero bandwidth and hence for this signal the upper bound on $I(E_s)$ is achieved (let $B_s = 0$ in the inner integral in (7)).

In summary, *the optimal transmit signal is sinusoidal at the radar center frequency.* It is interesting to note that the optimal signal is the same *irregardless of the clutter scattering function.* Also, with $F_0 = 0$ the average deflection coefficient $\bar{\Delta}_0$ now becomes from (10) with $W/2 = F_{d_{\max}}$ and $\bar{\Delta}_0 = (\bar{E}_r/\mathcal{E})J(E_s)$

$$\bar{\Delta}_0 = \frac{\bar{E}_r}{2F_{d_{\max}}} \int_{-F_{d_{\max}}}^{F_{d_{\max}}} \frac{1}{\mathcal{E}P_d(F) + N_0} dF. \tag{12}$$

This result says that *on the average* the detectability is enhanced by using a signal that detects high-Doppler targets. This is because low-Doppler targets have poor detectability so that it is more important to ensure that high-Doppler targets are easily detected. In this way the overall detectability is maximized *on the average.* (We also point out that the use of the expected value of $\Delta_0(F_d)$ as our detectability criterion is not the same as using the expected value of the probability of detection as would be given by averaging (2). This is because the relationship between P_D and $\Delta_0(F_d)$ is nonlinear and the expectation does not commute over nonlinear operations [6]. However, we expect the same general type of result, that the transmit signal should be narrowband.) In the next section, we revisit the results in [2] to show for a specific example that a sinusoidal signal is optimal.

3 Van Trees' Example [2] – Revisited

In [2] it was assumed that the Doppler scattering function was Gaussian with the form

$$P_D(F) = \frac{N_r}{\sqrt{2\pi}\sigma_R} \exp\left(-\frac{1}{2}F^2/\sigma_R^2\right)$$

where σ_R is the root-mean-square (RMS) Doppler spread in Hz, and N_r is used to adjust the clutter power. The transmitted signal is one that has a Gaussian envelope and a linear FM sweep

$$s(t) = \sqrt{\mathcal{E}}\left(\frac{1}{\pi T^2}\right)^{1/4} \exp\left[-\left(\frac{1}{2T^2} - jb\right)t^2\right] \qquad -\infty < t < \infty.$$

The ESD of the transmit signal is

$$E_s(F) = \frac{\mathcal{E}}{\sqrt{2\pi}B} \exp\left(-\frac{1}{2}F^2/B^2\right)$$

where B, the RMS bandwidth in Hz, is given by

$$B = \frac{1}{2\pi}\left(\frac{1}{2T^2} + 2b^2T^2\right)^{1/2}.$$

Upon convolving the ESD with the Doppler scattering function we have

$$E_s(F) \star P_D(F) = \frac{N_r\mathcal{E}}{\sqrt{2\pi}\gamma} \exp\left(-\frac{1}{2}F^2/\gamma^2\right)$$

where $\gamma = \sqrt{\sigma_R^2 + B^2}$ is the RMS bandwidth of the clutter PSD. There are several parameters of interest in this example. One is the clutter power to noise power in the equivalent rectangular bandwidth of the clutter. Denoting this by D it is given as $\mathcal{E}N_r/(N_0\sqrt{2\pi}\sigma_R)$. The target Doppler to RMS Doppler spread is given by F_d/σ_R, and finally, the signal bandwidth to Doppler spread is B/σ_R. With these parameters the deflection coefficient is given by

$$\Delta_0(F_d) = \frac{\bar{E}_r}{N_0}\int_{-W/2}^{W/2} \frac{(\sqrt{2\pi}B)^{-1}\exp\left[-(1/(2B^2))(F - F_d)^2\right]}{1 + ((D\sigma_R)/\gamma)\exp[-F^2/(2\gamma^2)]}dF \tag{13}$$

and a normalized version is given by

$$\Delta_{0,n}(F_d) = \frac{\Delta_0(F_d)}{\frac{\bar{E}_r}{N_0}}.$$

In Figure 1 is plotted the normalized deflection coefficient $\Delta_{0,n}(F_d)$ versus B/σ_R, which is the normalized signal bandwidth, for $D = 100$. Note that each curve is for a different target Doppler, expressed in terms of target Doppler to RMS clutter bandwidth ratio or F_d/σ_R. This matches the results in [2]. Some extra curves have been added for clarity. Note that for zero Doppler the performance is poor but even for $F_d/\sigma_R = 4$ the performance is nearly the upper bound of 1 for small signal bandwidths. Equivalently, $\Delta_0(F_d) = \bar{E}_r/N_0$, which is the performance when no clutter is present. By averaging these curves in

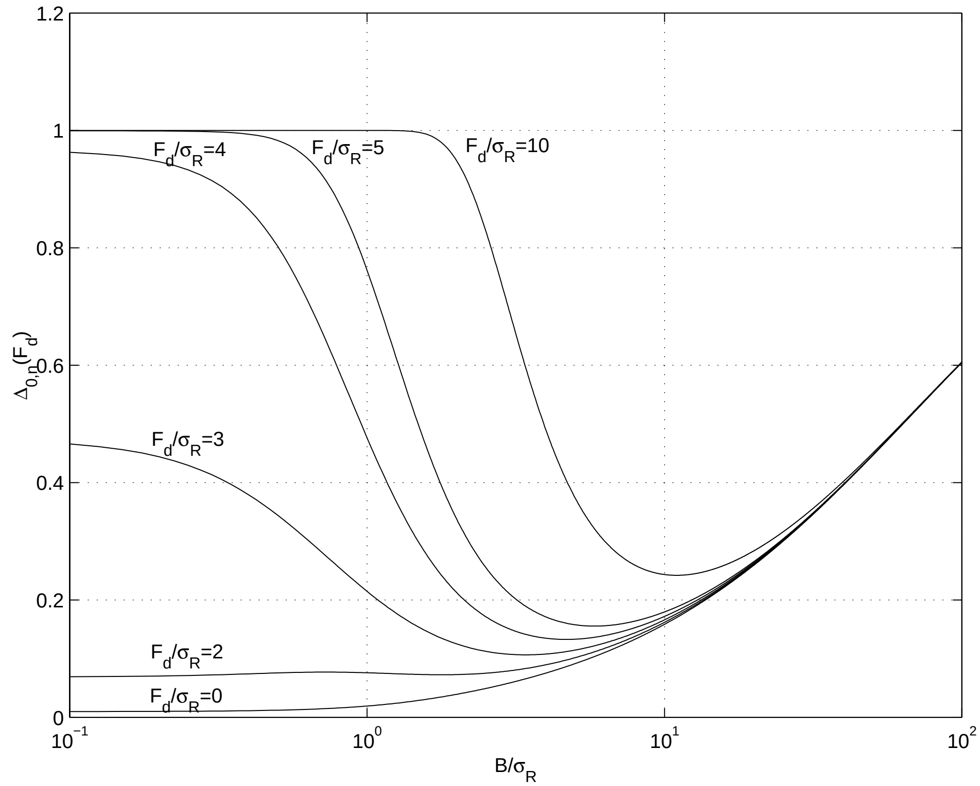

Figure 1: Normalized deflection coefficient of optimal receiver in the presence of clutter [2].

the vertical direction or uniformly over F_d we obtain from (13) the average deflection coefficient. As an example, if $F_{d_{\max}} = 100$ Hz, $0 \leq B \leq 100$ Hz, and $\sigma_R = 1$, then for each signal we require a bandwidth as determined by $W/2 = 3B + 100$. Hence, the average deflection coefficient is

$$\bar{\Delta}_0 = \frac{1}{2F_{d_{\max}}} \int_{-F_{d_{\max}}}^{F_{d_{\max}}} \Delta_0(F_d) dF_d$$

where $\Delta_0(F_d)$ is given by (13). The normalized average deflection coefficient $\bar{\Delta}_{0,n} = \bar{\Delta}_0/(\bar{E}_r/N_0)$ is shown in Figure 2. Note that as predicted from our analysis the performance is best for a signal bandwidth of zero. It also is seen to decrease monotonically as the bandwidth increases until a certain bandwidth, when it begins to increase. However, it is still poorer than for the zero bandwidth signal. If we had infinite bandwidth, then it can be shown that the deflection coefficient as well as the average deflection coefficient could be made to equal that for the case of no clutter. This is possible because the clutter spectrum would fall below the ambient noise level. This is seldom possible in practice. For this example of $D = 100$ the loss incurred is not that great. However, for larger clutter to noise ratios it can be substantial. For example, if $D = 100,000$, a clutter to noise ratio of 50 dB, the corresponding curve in dB is shown in Figure 3.

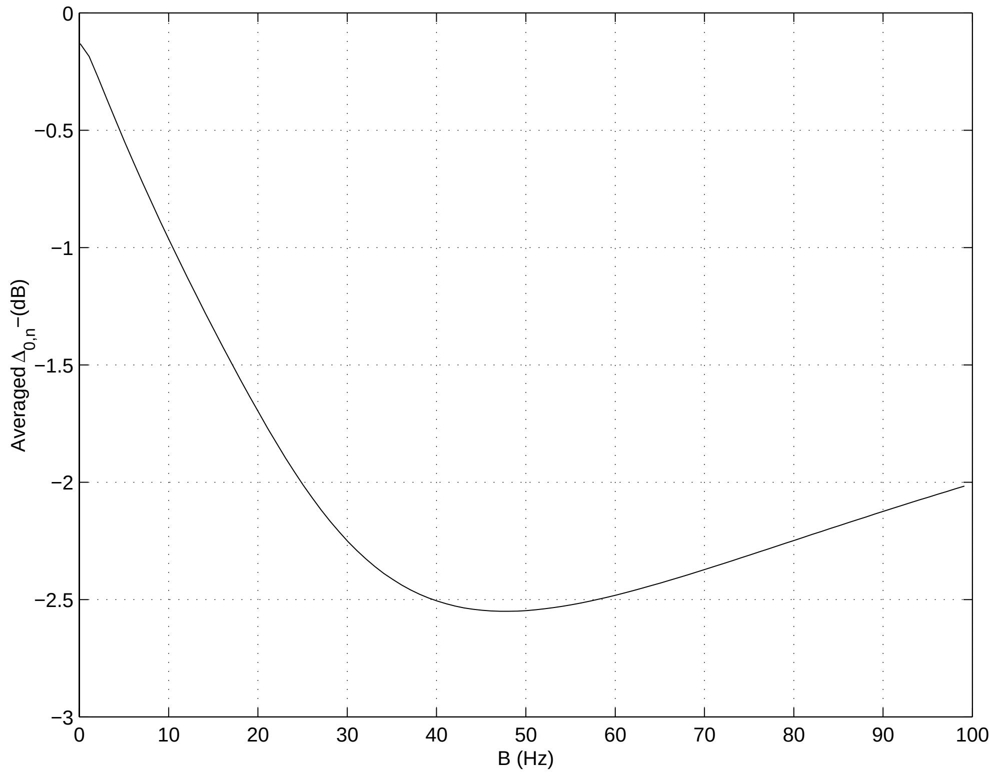

Figure 2: Normalized average deflection coefficient of optimal receiver in the presence of clutter for $D = 100$.

References

1. Shrader, W.W., V. Gregers-Hansen, "MTI Radar", in *Radar Handbook*, M. Skolnik, Ed., McGraw-Hill, New York, 1990.

2. Van Trees, H.L., *Detection, Estimation and Modulation Theory, Part III*, pp. 477–479, J. Wiley, New York, 1971.

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

4. Richards, M.A., *Fundamentals of Radar Signal Processing*, McGraw-Hill, New York, 2005.

5. Roberts, A.W., D.E. Varberg, *Convex Functions*, Academic Press, New York, 1973.

6. Kay, S., *Intuitive Probability and Random Processes Using MATLAB*, Springer, New York, 2006.

7. Van Trees, H.L., *Optimum Array Processing*, pg. 339, J. Wiley, New York, 2002.

8. Taylor, A.E., *General Theory of Functions and Integration*, Dover, New York, 1985.

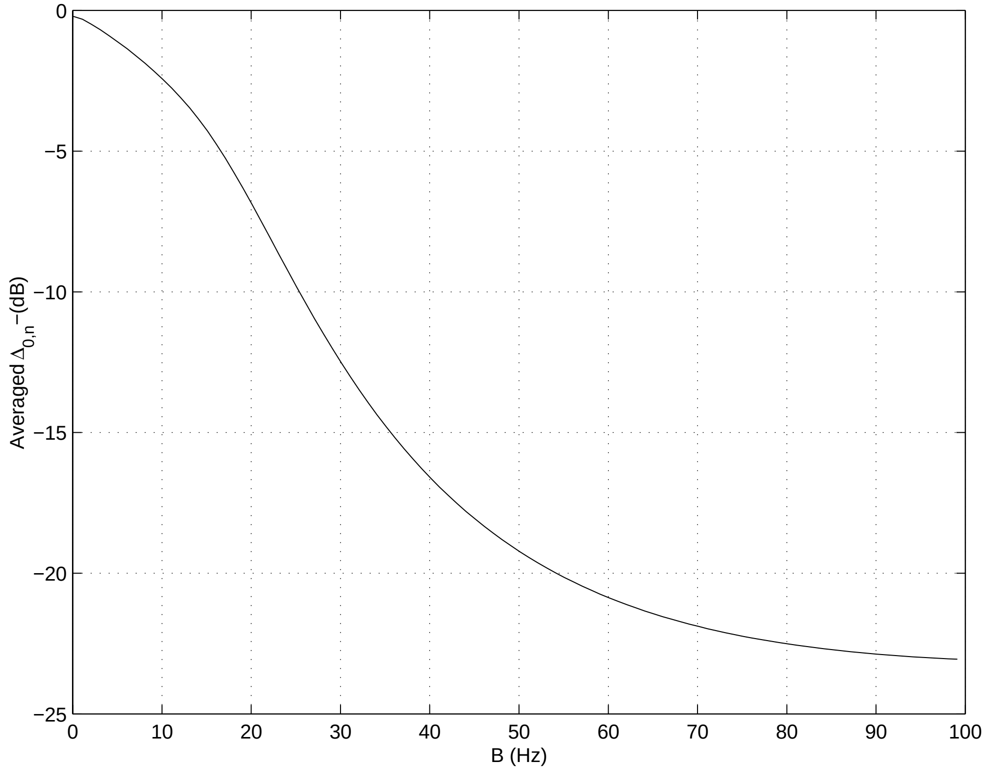

Figure 3: Normalized average deflection coefficient of optimal receiver in the presence of clutter for $D = 100,000$.

A Appendix – Convexity of Deflection Coefficient Functional

From (8) we have that

$$J(E_s) = \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{E_s(F) \star P_d(F) + N_0} dF. \tag{14}$$

First note that the function $g(x) = 1/(x+c)$ for $x \geq 0$ and $c > 0$ is convex. This is easily verified by showing that the second derivative is $g''(x) = 2/(x+c)^3 > 0$. Hence, it follows that for $0 < \alpha < 1$

$$g(\alpha x_1 + (1-\alpha)x_2) \leq \alpha g(x_1) + (1-\alpha)g(x_2)$$

and as a result

$$\frac{1}{\alpha x_1 + (1-\alpha)x_2 + c} \leq \frac{\alpha}{x_1 + c} + \frac{1-\alpha}{x_2 + c}.$$

Thus, (14) becomes

$$\begin{aligned} J(\alpha E_{s_1} + (1-\alpha)E_{s_2}) &= \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{(\alpha E_{s_1}(F) + (1-\alpha)E_{s_2}(F)) \star P_d(F) + N_0} dF \\ &\leq \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{\alpha}{E_{s_1}(F) \star P_d(F) + N_0} + \frac{1-\alpha}{E_{s_2}(F) \star P_d(F) + N_0} dF \end{aligned}$$

$$
\begin{aligned}
&= \alpha \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{E_{s_1}(F) \star P_d(F) + N_0} dF \\
&\quad +(1-\alpha) \frac{\mathcal{E}}{2F_{d_{\max}}} \int_{-W/2}^{W/2} \frac{1}{E_{s_2}(F) \star P_d(F) + N_0} dF \\
&= \alpha J(E_{s_1}) + (1-\alpha) J(E_{s_2})
\end{aligned}
$$

proving the convexity of $J(E_s)$.

B Appendix – Decomposition of Energy Spectral Density

It was shown in Appendix A that $J(E_s)$ is a convex functional on the space of nonnegative, energy constrained ESDs. To find the maximum we need to determine the form of the extreme points of the space, i.e., the ones that *cannot be written as in (9)*. We next show that these extreme points take the form of ESDs that are concentrated on a single point in frequency. To do so we will use the theory of spectral measures, which may also be thought of as probability measures, although the total mass is $\mathcal{E}$. We first prove the following lemma.

Lemma B.1 (Measure Decomposition) *A nonnegative measure on a measure space $\chi = [-a, a]$ with $\mu(\chi) = \mathcal{E}$ can be decomposed as $\mu(A) = \alpha\mu_1(A) + (1-\alpha)\mu_2(A)$ for $0 < \alpha < 1$ and for all measurable sets A with $\mu_1 \neq \mu_2$ if and only if there exists a set $S \subset \chi$ such that $\mu(S) > 0$ and $\mu(S^c) > 0$.*

Proof: If

We assume the existence of S with $\mu(S) > 0$ and $\mu(S^c) > 0$. Hence, by construction we arrive at the required convex combination. To do so we have for an arbitrary set A

$$
\begin{aligned}
\mu(A) &= \mu(A \cap \chi) \\
&= \mu(A \cap (S \cup S^c)) \\
&= \mu(A \cap S) + \mu(A \cap S^c) \\
&= \mu(S) \frac{\mu(A \cap S)}{\mu(S)} + \mu(S^c) \frac{\mu(A \cap S^c)}{\mu(S^c)} \\
&= \underbrace{\frac{\mu(S)}{\mathcal{E}}}_{\alpha} \underbrace{\frac{\mathcal{E}\mu(A \cap S)}{\mu(S)}}_{\mu_1(A)} + \underbrace{\frac{\mu(S^c)}{\mathcal{E}}}_{1-\alpha} \underbrace{\frac{\mathcal{E}\mu(A \cap S^c)}{\mu(S^c)}}_{\mu_2(A)}.
\end{aligned}
$$

Note that since by assumption $\mu(S^c) > 0$ and $\mu(S) + \mu(S^c) = \mathcal{E}$, we must have $\mu(S) < \mathcal{E}$ and hence $0 < \alpha < 1$. Also, μ_1 and μ_2 are valid measures within the same space as μ since $\mu_1(\chi) = \mu_2(\chi) = \mathcal{E}$, and $\mu_1 \neq \mu_2$ since $\mu_1(S) = \mathcal{E}$ while $\mu_2(S) = 0$. The convex combination is just a restatement of the law of total probability [6].

Proof: Only if

Next assume that $\mu(A) = \alpha\mu_1(A) + (1-\alpha)\mu_2(A)$. Since $\mu_1 \neq \mu_2$ there must exist a set S such that $\mu_1(S) \neq \mu_2(S)$. On this set we must have $\mu(S) > 0$ since otherwise $\mu(S) = 0$ and hence $\mu_1(S) = \mu_2(S) = 0$, violating the assumption $\mu_1(S) \neq \mu_2(S)$. Also, it follows that $\mu(S) < \mathcal{E}$ since otherwise $\mu(S) = \mathcal{E}$ and therefore, $\mu_1(S) = \mu_2(S) = \mathcal{E}$, violating the assumption $\mu_1(S) \neq \mu_2(S)$. Thus, $\mu(S) > 0$ and $\mu(S^c) > 0$, completing the proof of the lemma. With this lemma we next prove the main theorem.

Theorem B.1 (Point measure as the extreme point) *If μ cannot be decomposed as $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, then μ is a point measure or $\mu = \delta_{x_0}$ for some $x_0 \in \chi$. Here δ_{x_0} is a measure with $\delta_{x_0}(A) = 1$ if $x_0 \in A$ and zero otherwise.*

Proof:

Since μ cannot be decomposed as given in the theorem, we have from the preceding lemma that there cannot exist a set S such that $\mu(S) > 0$ and $\mu(S^c) > 0$. Alternatively, for every set S if $\mu(S) > 0$, then $\mu(S^c) = 0$. Consider the half-interval sets $[-a, 0)$ and $[0, a]$. Let S be the set for which the half-interval set measure is positive, say $S = [0, a]$, for example. (At least one of the sets must have a positive measure since $\chi = S \cup S^c$ and $\mu(\chi) = \mathcal{E}$.) Then, since $\mu(S^c) = 0$ if $\mu(S) > 0$, we have $\mu([-a, 0)) = 0$. Note that if $S^c = [-a, 0)$ had positive measure, then we could augment S^c with the point set $\{0\}$ to form $[-a, 0]$. The latter set is closed and has the same measure as S^c since $\{0\} \subset S$, which has measure zero. Next partition $[0, a]$ into two half-intervals again. As before, since one of the half-intervals must have a positive measure, the other has zero measure. Choose the half-interval with positive measure, augmenting it with an end point if necessary to form a closed interval, and note that its measure is $\mathcal{E}$ since the discarded half-interval must again have measure zero since $\mu(S^c) = 0$. Continuing to bisect the intervals we generate a sequence of nested closed intervals (each closed set is a subset of the preceding one) whose length decreases to zero. By the nested interval theorem, the intersection of these sets, i.e., the limit set is a point [8]. Also, by construction the limit set has measure $\mathcal{E}$, proving the theorem.