New Methods for the Detection and Interception of Unknown, Frequency-Hopped Waveforms

by

William Edward Snelling

Dissertation submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1990

c.1

Advisory Committee:

Associate Professor Evaggelos A. Geraniotis, Chairman/Advisor

Associate Professor Benjamin Kedem

Professor Lee D. Davisson

Associate Professor Steven A. Tretter

Assistant Professor Adrianos Papamarcou

©Copyright by

William Edward Snelling

1990

Abstract

Title of Dissertation: **New Methods for the Detection and Interception of Unknown, Frequency-Hopped Waveforms**

William Edward Snelling, Doctor of Philosophy, 1990

Dissertation directed by: Dr. Evaggelos Geraniotis, Associate Professor, Department of Electrical Engineering, University of Maryland, College Park

Three new methods for the detection and interception of frequency-hopped waveforms are presented. The first method extends the optimal, fixed-block detection method based on the likelihood ratio to a sequential one based on the Sequential Probability Ratio Test (SPRT). The second method is structured around a compressive receiver and is highly efficient yet easily implemented. The third method is based on the new concept of Amplitude Distribution Function (ADF) and results in a detector that is an extension of the radiometer.

The first method presents a detector structured to make a decision sequentially, that is, as each data element is collected. Initially, a purely sequential test is derived and shown to require fewer data for a decision. A truncated sequential method is also derived and shown to reduce the data needed for a decision while operating under poor signal-to-noise ratios (SNRs). A detailed

performance analysis is presented along with numerical and Monte Carlo analyses of the detectors.

The second method assumes stationary, colored Gaussian interference and presents a detailed model of the compressive receiver. A locally optimal detector is developed via the likelihood ratio theory and yields a reference to which previous ad hoc schemes are compared. A simplified, suboptimal scheme is developed that trades off duty cycle for performance, and a technique for estimating hop frequency is developed. The performance of the optimal and suboptimal detectors is quantified. For the suboptimal scheme, the trade-off with duty cycle is studied. The reliability of the hop frequency estimator is bounded and traded off against duty cycle.

In the third method, a precise definition of the ADF is given, from which follows a convolutional relationship between the ADFs of signal and additive noise. A technique is given for deconvolving the ADF, with which signal and noise components can be separated. A detection statistic based directly on this deconvolution technique is defined and statistically characterized, yielding a framework on which to synthesize a detector. The detector's performance is analyzed and compared with the radiometer.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

1.1 Motivation

The goal of the intercept receiver is to detect, identify, and geolocate hostile electromagnetic (EM) sources and use this information to counter with Electronic Counter Measures (ECM) and Electronic Counter Counter Measures (ECCM). For military communications, the receiver is expected to intercept any one of a set of target communications. The interception task is hindered by a dense EM environment that consists of other communication signals, both friendly and hostile, possibly jamming or masking signals in addition to ever-present noise interference. With the advent of frequency-hopped and other spread-spectrum communications, the search bandwidth that assures a reasonable probability of intercept is greatly increased, thereby aggravating the problem of interference, because greater numbers and types of interfer-

1

ences obscure the target signal. The increased complexity of the interception problem motivates the search for new methods of detection and interception of frequency-hopped waveforms.

1.2 Target Signals

Military and other secure communications use spread-spectrum signaling involving some variety of modulation whose purpose is to add ambiguity or "randomness" to the waveform as a measure against unintended detection or interception. The usual procedure for randomizing the waveform is pseudo-random variation of transmission times (time hopping or TH), phases (direct sequence or DS), or frequencies (frequency hopping or FH). This work concentrates solely on the interception of FH waveforms that have form

$$s(t) = \sum_{i=1}^{N_h} x_i(t) \tag{1.1}$$

where

$x_i(t)$ equals $\sqrt{2S'}\sin(\omega_{k_i}t + \theta_i)$ for $iT_h \le t \le (i+1)T_h$;

$\{\omega_k\}_{k=1}^{K}$ is a family of known frequencies within the spread-spectrum bandwidth;

$\{k_i\}$ are integer-valued, independent, uniformly distributed, random variables ranging inclusively between 1 and K;

$\{\theta_i\}$ are continuous, independent, uniformly distributed, random variables ranging between 0 and 2π that represent carrier phase;

S' is a real constant denoting the average signal energy;

T_h is a real constant denoting the epoch, or time duration, of each hop;

N_h is a positive integer denoting the number of hops during message transmission.

This general model of frequency-hopped waveforms includes a large number of modulations such as frequency shift keying (FSK) and minimum shift keying (MSK). Some important modulations not included are those whose carrier phase is correlated from hop to hop, for example, continuous phase FSK (CPFSK). Even for these cases, these results apply but may not be optimal.

1.3 Intercept Receiver Functions

An intercept receiver extracts, for further processing, a small number of candidate signals from the plethora of signals in a communication band of interest. The initial processing steps that discard signals are called Pruning Functions. After pruning, secondary processing, known as feature extraction, yields information aiding in emitter identification and countering. After fea-

ture extraction, further processing could yield the actual information embedded in the communication signal, but this is peripheral to the primary function of the intercept receiver, namely ECM and weapons support, and consequently will not be explored here.

1.3.1   Pruning

Given the frenetic activity in most communication bands, the intercept receiver must, early in its processing, choose out of all candidate signals within the band a small number of potential target signals. Pruning Functions achieve this by eliminating all but the most promising prospects for processing. Pruning Functions fall into four categories: Initial Detection, Direction Finding, Frequency Estimation, and Time of Arrival Measurement. As is evident from their names, Initial Detection separates potential target signals from background noise, Direction Finding classifies and possibly eliminates signals by direction of origin, Frequency Estimation censors signals based on a measurement of carrier frequency, while Time of Arrival Measurement differentiates between pulsed signals based on their arrival times. It is useful conceptually to consider the Pruning Functions as independent processes applied separately, but in a practical system these functions are usually highly coupled in that a single processing step may accomplish two or more Pruning Functions.

Initial Detection separates the candidate signal from the background noise, usually in the form of a threshold operation applied on a test statistic derived from the received waveform. It may be a dedicated operation, meaning that the only information gleaned is the presence or nonpresence of a communication waveform, or may yield other information such as the time of arrival (time domain) or the dominant frequency (frequency domain), or, for feature detectors, the hop rate.

The signal direction can be found by using a scanning narrowbeam antenna whose scan position at the time of detection determines direction. Difference in signal phase from two separate antennas can yield the angle from which a candidate signal is emitted; however, this interferometry technique suffers from a vulnerability to coherent interference. In a similar fashion, the amplitude difference between the same signal received from different antennas or antenna-patterns can determine emission angle but the technique is even more vulnerable to interferences, coherent or noncoherent. The difference between the times of arrival of different receivers is an alternative way to determine angle. This method seems more directed to pulse signals, such as radar, in which arrival time is a relatively simple quantity to measure, but it could be applied to more complicated communication signals, by cross correlation for instance.

By estimating the dominant frequency of a candidate signal, narrowband

5

interferences such as other non-spread-spectrum communication signals can be identified and rejected. Additionally, the current hop frequency of the target signal can be determined and subsequently used to narrowband jam the current hop band. The estimated hop frequency also can be used for identification processing or information extraction.

Time of Arrival Measurement can not only determine emission angle, as previously mentioned, but also can prune. This use is primarily useful for pulse radar signals, in which times of arrival can determine pulse repetition rates and hence associate the intercepted pulse with the emitting radar. For frequency-hopped communications, the arrival times of individual hop intervals determine the emitters' hop rate and hence discriminate between target signal and interferers.

1.3.2   Feature Extraction

Feature Extraction is the measurement of characterizing features of the communication waveforms. Features such as hop rate, hop frequency, modulation type, and bit periods serve as examples. Feature Extraction overlaps the Pruning Functions in that operations such as center frequency estimation and time-of-arrival measurements yield useful features. The loose distinction is one of precision and purpose. Measurements made while pruning are coarse

and serve only to decimate what otherwise would be an unmanageable number of candidate signals, while extracted features are of sufficient accuracy to serve the intercept purposes of jamming and identification.

1.4   Existing Methods

All aspects and functions of the intercept receiver were described. However, this work concentrates solely on initial detection and feature extraction and on feature extraction; only hop frequency estimation will be explored.

1.4.1   Wideband Energy Detector

The Wideband Energy Detector ([40] and [45]) is the simplest to implement of all existing detection schemes. Also called a radiometer, it is a device for measuring the energy of a signal over a prescribed time and bandwidth. A typical radiometer (Figure 1.1) consists of a bandpass filter followed by a square law device and an integrator. The bandpass filter, of bandwidth W, selects the frequency band over which the energy measurement is made. The square law device and integrator calculate the energy of this bandpassed signal over the interval of length T. The radiometer is used to detect spread-spectrum signals by setting the center frequency and bandwidth of the bandpass filter so that the filter will pass the expected spread-spectrum signal. A decision is

Comparator

BPF W — $(\ )^2$ — $\int_0^T (\ )\, dt$ — $+$ $\oplus$ $-$ — Decision $\lessgtr 0$
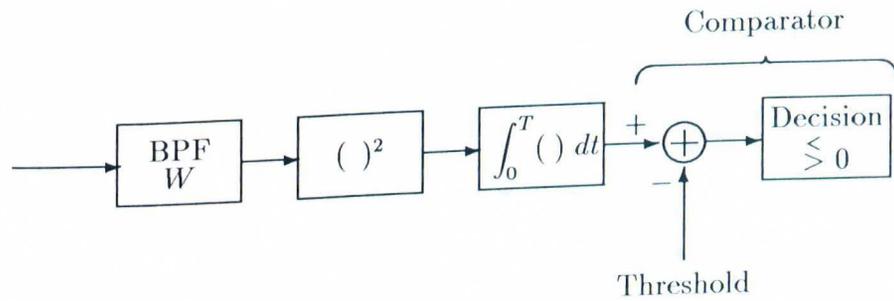
Threshold

Figure 1.1: Wideband Energy Detector

made by comparing the output of the radiometer to a threshold.

As can be guessed, the wideband radiometric detector is most efficient when the bandwidth exactly matches the spread-spectrum bandwidth K/T_h and the integration period matches the transmission time $N_h T_h$ of the spread-spectrum signal. Under these conditions and for large time bandwidth ($TW > 1000$) products, the performance of the wideband radiometer is described below for the case of white-noise interference with single-sided spectral density N_0 as

$$\frac{S'T_h}{N_0} = d\sqrt{\frac{K}{N_h}} \tag{1.2}$$

and

$$d = Q^{-1}(P_F) - Q^{-1}(P_D) \tag{1.3}$$

where $Q^{-1}(\cdot)$ is the inverse of the complementary Gaussian probability distribution and $S'T_h/N_0$ is the required signal-to-noise ratio for detection with probability P_D and false-alarm probability P_F.

8

In addition to being the easiest detector to implement, the wideband energy detector assumes the least known about the spread-spectrum signal. For optimal detection of a given spread-spectrum waveform, only the bandwidth and message duration need to be known. However, an approximate knowledge of these parameters degrades performance only slightly. Because the performance of the wideband energy detector is invariant to the details of the spread-spectrum waveform, it is equally effective in the detection of either FH, TH, or DS waveforms. It is also useful as a lower bound on the performance of other detectors designed around the particulars of a given spread-spectrum waveform.

1.4.2 Optimal Channelized Detector

The Optimal Channelized Detector ([45] and [50]) uses a more precise knowledge of a spread-spectrum waveform to achieve performance gains over the Wideband Energy Detector and other detector configurations. In the context of FH waveforms, the message duration and the period and phase of the pulse epochs are assumed to be known. Also assumed to be known are the exact frequencies of the pulsed sinusoids that constitute the individual "hops" of the spread-spectrum waveform. The signal amplitude relative to the background noise is also assumed to be known. Not known are the relative phase between the individual hops and, of course, the pseudo-random code that
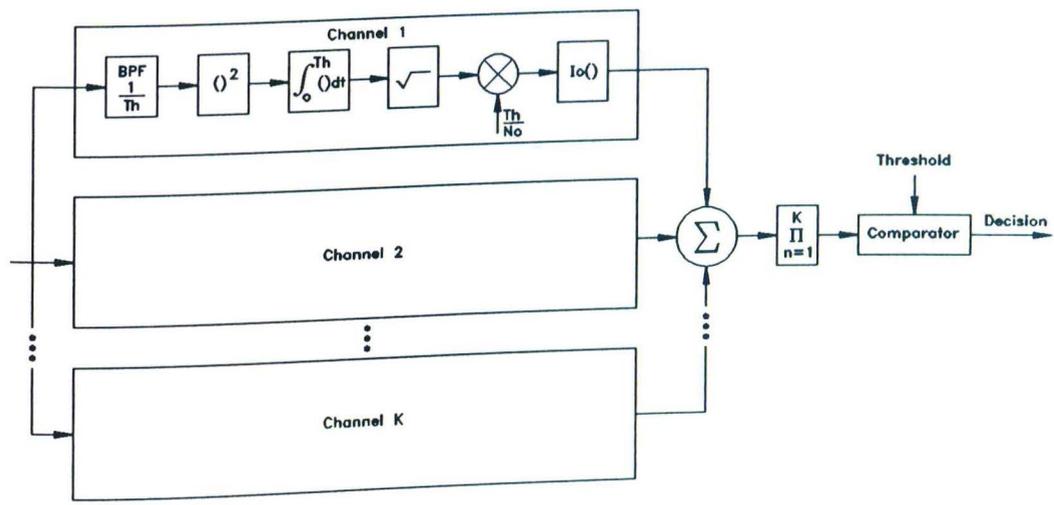
Figure 1.2: Optimal Channelized Detector

produces the hops.

With these assumptions and that of white-noise interference, detection the-
ory yields the detector shown in Figure 1.2. This detector consists of individual
filters matched in time and frequency to each of the possible pulsed sinusoids
component to the FH waveform. The envelopes of the matched filter outputs
are "emphasized" by normalizing by expected noise energy and applying the
zero-order modified Bessel function of the first kind I_0. The emphasized filter
outputs are summed to yield a likelihood function over a single epoch. These
individual likelihood functions for each epoch of the message are multiplied to
yield the overall likelihood function, from which a decision can be made via a
threshold comparison.

10

The generalized performance expression for the optimal multichannel detector cannot be obtained due to an inability to specify the output probability distribution functions. When the number of hops N_h is large (e.g., $N_h > 100$), it is possible to closely approximate the true answer by using Gaussian statistics. This analysis gives $S'T_h/N_0$ needed for a given P_{FA} and P_D as

$$\frac{S'T_h}{N_0} = \frac{1}{2}I_0^{-1}\left[1 - K + Ke^{\frac{d^2}{N_h}}\right] \tag{1.4}$$

where d is given before.

Unfortunately, however, the Optimal Channelized Detector is only of academic interest because of its implementation complexity and its sensitivity to the FH waveform parameters. It is useful primarily for establishing an upper bound to the performance of other more implementable and robust detectors.

1.4.3 Suboptimal Channelized Detectors

Because of the implementation complexity of the Optimal Channelized Detector, Suboptimal Channelized Detectors are considered. These are several classes of detectors that are simplifications in various ways of the Optimal Channelized Detector. These simplifications are listed below.

The combinations of a matched filter followed by an envelope detector are replaced by narrowband radiometers of bandwidth $1/T_h$. This simpli-

fication degrades the resulting performance over that of the predicted performance of the Optimal Channelized Detector. However, this performance difference would be less in practice, because the Optimal Channelized Detector would not achieve the optimum performance due to the inevitable mismatch between the actual frequencies received and the frequencies for which the filters are matched. Doppler shift and transmitter waveform diversity are the likely culprits of this mismatch.

The emphasizing function is linearized. This is mainly an implementation simplification. However, the optimality of the Optimal Channelized Detector depends on a priori knowledge of the amplitude of the FH waveform, which is a parameter of the emphasis. Thus the loss of optimality in practice is lessened because of the uncertainty of the FH waveform amplitude.

Decisions are made at the channel level and are then combined to form a statistic upon which the final decision is based. This type of detector is appealing if a frequency estimate of the detected is also desired.

Instead of having a filter for each FH frequency, the entire spread-spectrum bandwidth is subdivided into coarse subbands. The subbands most likely to contain the current hop are selected for application of any of the above channelized detector schemes.
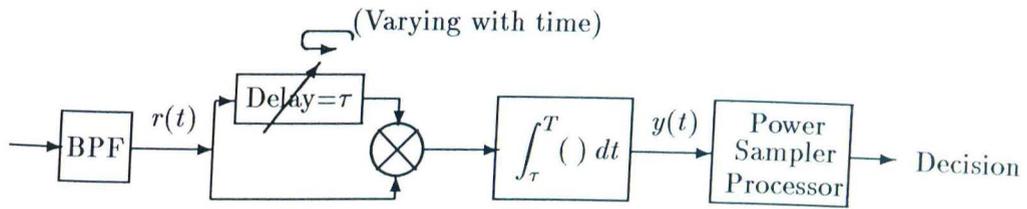
Figure 1.3: Autocorrelation Detector

1.4.4 Autocorrelation Detector

An Autocorrelation Detector ([29] and [31]) is composed of a bank of auto-correlators, each operating on a subband of the total spread-spectrum bandwidth (Figure 1.3). The autocorrelators estimate the autocorrelation $y(\tau)$ of their bandlimited input over the time period T. The power of each correlation is sampled yielding W_k which are weighted with a_k and summed to derive the decision statistic Y. This is summarized by the equation

$$Y = \sum_{k=1}^{G-1} a_k W_k \underset{H_0}{\overset{H_1}{\gtrless}} \text{Threshold.} \qquad (1.5)$$

There are three issues concerned with the design of the autocorrelation detector. The first is the coarseness of the individual subbands relative to the total spread-spectrum bandwidth. It seems intuitively appealing to assume that performance would improve by reducing the width of this subband up to the limits of the FH frequency spacing; however, no analytical or numerical

13

results confirm this conjecture. The second design issue is the time interval over which we estimate the autocorrelation function. The third issue involves the weights used in the computation of the decision statistic.

1.5 Hop Frequency Estimators

As previously described, feature detectors focus on a particular time-domain feature of the spread-spectrum waveform. In this work, we develop a detector based on the feature hop frequency. These devices can do the job of initial detection, but we focus on their estimation performance. Two noteworthy estimators in the literature fit this billing. The first, described in [3], is a maximum likelihood estimator with a structure similar to the optimal detector of Figure 1.2, except that, instead of summing the outputs of each channel, it selects the channel with the output of maximum magnitude. The hop frequency corresponding to that channel is declared the estimate (Figure 1.4). The second estimator of note, described in [39], is based on the first one but has reduced complexity. In this method, wideband radiometers cover the spread-spectrum bandwidth in order to select a small number of subbands that most likely contain the particular hop. These subbands are further processed into fine bands, enabling the ultimate selection of the band with the current hop.
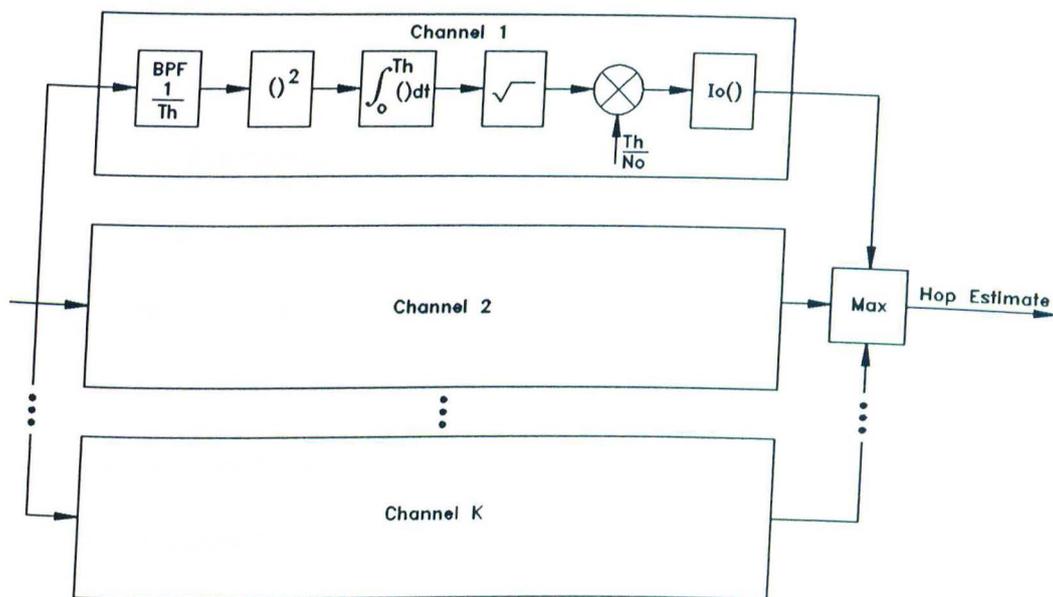
Figure 1.4: Maximum Likelihood Hop Frequency Estimator

1.6 New Methods for Detection and Hop Estimation

1.6.1 Sequential Detector

In this work, a new detector is developed based on the ideas of sequential detection. It is essentially like the optimal channelized detector but, instead of basing its decision on accumulated energy on a predetermined large number of hop dwells, it decides, after each hop dwell, on the presence or nonpresence of a frequency-hopped waveform. Because the detector is based on the Sequential Probability Ratio Test (SPRT), the test is optimal in the sense that no other sequential test will make a decision in less time on average than the SPRT. However, optimality occurs only if the signal is present at a predetermined SNR. For smaller SNRs, the SPRT actually can perform worse than a test based on a Fixed Sample Size (FSS). This shortcoming is averted by mixing the SPRT with the FSS test to create the Truncated Sequential Test (TST). From these results is designed an optimal test whose worst-case average decision time is minimal. Netted are three new detectors that exploit the advantages of sequential detection; the pure SPRT, the TST, and the optimal TST.

With each of these techniques, the number of samples needed for a reliable decision is dramatically reduced. One way this detector performance gain

16

can yield a performance gain in the interceptor itself is through decreased duty cycle. A typical interceptor might scan a particular direction in order to determine the presence or absence of communications. Because of the gains due to the sequential detector, the scan time is significantly reduced. Another way to take advantage of the performance gain is to add robustness to the signal parameters. It was already mentioned that the decision time of the sequential tests is dependent on the SNR. By hypothesizing that the signal rests in a band of SNRs, a sequential test can be designed that still outperforms tests based on a fixed sample time.

1.6.2 Compressive-Receiver-Based Detector and Hop Estimator

The compressive receiver, which simultaneously estimates frequency components over a wide, predetermined band, has promise as an interceptor with both the simplicity of a wideband device and the performance of the channelized device. The use of the compressive receiver for interception is a largely unexplored area with all previous results being superficial and ad hoc. By contrast, two different detectors and a hop frequency estimator are developed using an optimal likelihood function approach. The first, the locally optimal detector, is a detector with structure similar to the channelized detector but operating on the output of the compressive receiver. It is locally opti-

17

mal, meaning that for signals with low SNR it gives the greatest probability of detection for a given probability of false alarm. Because the locally optimal detector has an unwieldy structure, it defeats the motivation to use a compressive receiver: simplicity and high performance. Therefore, a time-multiplexed detector is used that, at the expense of duty cycle, can achieve performance as close to optimal as desired. Both the locally optimal and the time-multiplexed detector have hop frequency estimator versions. By choosing as the hop frequency estimate the hop frequency corresponding to the detector channel with maximum output, a hop frequency estimator is formed. In conclusion, two detectors and a hop frequency estimator are developed with performance comparable to the channelized devices but with the simplicity of the broadband devices.

1.6.3   Detector Based on the Amplitude Distribution Function

A new idea for detection is developed based on the Amplitude Distribution Function (ADF). The ADF is precisely defined as a function from which, through a sequence of lemmas and theorems, two results are established. One result is that the ADF is roughly the average distribution of a stochastic process, and the second is that, for signal plus noise, the resulting ADF is the convolution of the ADFs of signal and noise individually. The convolutional

relationship for signal plus noise motivates the construction of statistical transform, called the deconvolution statistic, that converges to something that is arbitrarily close to the signal ADF and hence has potential for separating signal from noise even for low signal levels. How close the deconvolved ADF matches the signal ADF depends on the proper choice of the kernel of the deconvolution statistic. An optimal detector is presented that directly observes samples of the deconvolution statistic, yielding a test statistic of quadratic form. The ADF-based detector is a robust device that is a generalization of the radiometer and is quite immune to the details of spread-spectrum modulation.

1.7 Document Organization

This work is partitioned into five chapters. The first chapter presents the problem, precisely defines the type of frequency-hopped waveforms under consideration, and describes the functions of the intercept receiver from a system viewpoint. It also briefly describes existing interception methods in the categories of initial detection and hop frequency estimation and then contrasts them with the new methods developed.

The second chapter describes in detail the new sequential detection methods. Within this chapter, the likelihood function for a single epoch is developed

and asymptotically analyzed for a large number of hop frequencies. Based on this analysis, the synthesis of the FSS test, SPRT, and TST are developed. Performance equations are presented along with the results of numerical and Monte Carlo analyses. The optimal TST is described and the asymptotic efficiencies, which capture the low-SNR test behavior, are presented. Finally, conclusions are drawn.

The third chapter gives a detailed description of the detector and hop frequency estimator based on a compressive receiver. It does this first by precisely defining the signal and compressive-receiver models and then using them to develop equations for the output signal component and to characterize statistically the noise at the compressive-receiver output. With the detection problem translated to the output of the compressive receiver, likelihood ratio theory is applied for the low-SNR case to create the locally optimal detector. A simplified detector, the time-multiplexed detector, is also presented along with a hop frequency estimator. All detectors are performance analyzed and numerical results given. Finally, conclusions are drawn.

The fourth chapter introduces the ADF-based detector and proceeds with an exposition of the mathematical tools developed for the ADF, which consist of a sequence of theorems and lemmas culminating in a convolutional relationship between the ADFs of signal and noise. The deconvolution statistic is introduced along with family of kernels to be used in the statistic. The

large-time statistical character of the deconvolution statistic is shown to be the basis of the ADF-based detector. Synthesis and performance analyses of the detector are presented and conclusions are drawn.

The fifth chapter summarizes the previous chapters and highlights the important points. It then suggests possible extensions to be investigated in the future and finally concludes the document.

Chapter 2

Sequential Detection of Unknown, Fast Frequency-Hopped Waveforms

2.1 Background and Introduction

The first task in the interception of spread-spectrum communications is the detection of the waveform. This is a prelude to other interception processes, such as feature detection, channel tracking, and message extraction. As a new development toward the detection problem, this chapter applies and extends previously published results in sequential detection to the problem of the optimal detection of noncoherent frequency-hopped (FH) waveforms. By using likelihood function methods, the problem was solved in [10] for an FH waveform with a known signal-to-noise ratio (SNR) and epochs with known starting times and durations. However, in that approach, the decision was based on a data segment of fixed size. Here a sequential approach is taken, meaning that whenever a new data element is collected, a decision about the

22

presence or nonpresence of an FH waveform is attempted. If no decision is reached, another data element is collected.

The sequential approach to detection has a rich history. For the binary hypothesis problem with discrete-time independent identically distributed (i.i.d.) data, Wald [48] has derived the optimal sequential test. This test is optimal in the sense that no other test can reach a decision of the same Neyman-Pearson reliability within a shorter average time. This result has been extended to continuous time data in references [38] and [6]. Others have suggested tests that must make a decision within a prescribed time. These are the "truncated" tests given in [1], [43], and [44]. Truncation is desirable not only for implementation reasons, but also for improving the performance of a sequential test when the input statistics differ from those assumed in designing the test. In particular, Tantaratana and Poor in [43] derive a truncated sequential test for i.i.d. Gaussian data with an unknown mean, which is the foundation of the results in this chapter.

Development of the sequential test is begun by defining the observations model for a composite hypothesis problem. Specifically, given the observation $y(t)$, the problem is one of choosing between H_0, which is the hypothesis that an FH waveform is not present, and $H_{\gamma'}$, which is the hypothesis that an FH

waveform is present with an SNR γ' where $0 < \gamma'$. Exactly, the model is

$$\text{versus} \quad \begin{aligned} H_0: \quad y(t) &= n(t) \\ H_{\gamma'}: \quad y(t) &= s(t) + n(t) \qquad 0 < \gamma' \end{aligned} \qquad (2.1)$$

where $s(t)$ is given by (1.1) and $n(t)$ is white Gaussian noise with two-sided spectral density $\frac{N_0}{2}$. The hypothesized SNR γ' is related to the other model parameters by $\gamma' = S'T_h/N_0$.

Because a reliable test cannot be devised for an FH waveform with an arbitrarily small SNR, the preceding composite hypothesis problem is simplified to a binary hypothesis problem: H_0 versus H_γ, where γ is specified as the smallest SNR that is to be accurately detected. The quantity $\gamma = ST_h/N_0$ with S being the corresponding signal energy. The relative SNR $r = \sqrt{\gamma'/\gamma}$ is also used.

Using the above observations model, the design of a sequential test for the detection of FH signals is approached as follows. An asymptotically optimal test is derived by applying the likelihood function theory to the simplified binary hypothesis problem H_0 versus H_γ. The parameters of this test are specified to ensure a maximum probability of detection for a given probability of false alarm. This binary hypothesis test is then applied to the more general composite hypothesis problem with a resulting degradation in detection time that is shown to be controllable by properly truncating the test procedure.

The derivation of the asymptotically optimal test begins with the derivation

of the likelihood function for a single-epoch observation, which is appropriately called the Single-Epoch Likelihood Function (SELF). By invoking the central limit theorem, Gaussian densities are found that are asymptotic to the actual SELF densities as the number of frequencies becomes large. In determining these densities, the SELF's means and variances will be explicitly computed under each hypothesis. By next considering individual SELFs as the observations, the problem will be reduced to a binary hypothesis problem with Gaussian i.i.d. observations. This simplification is justified, because each epoch of FH waveform has independent statistics and because the SELF's statistics do not depend on the particular hop frequency. Using these equivalent observations and their asymptotic densities, the Asymptotic Log-Likelihood Function (ALLF) is derived. The ALLF is then used to synthesize tests for the binary hypothesis problem. This procedure requires extending the previously published sequential tests to the cases of data with variances that depend on the hypothesis. Applying these results, a Fixed-Sample Size (FSS) test, a Sequential Probability Ratio Test (SPRT), and a Truncated Sequential Test (TST) are designed.

Each of the three tests is analyzed by approximating the test statistic by a Wiener process and then employing the classical theory of diffusion, as outlined in [6] and [1]. This analysis is more general because it yields the performance of each test to the composite hypothesis problem rather than just the binary

25

hypothesis problem on which the tests are based. This analysis yields the average decision time of each test as a function of the input SNR, as well as the operating characteristic of each test. From these results comes an optimal test, whose worst-case average decision time is minimal. Finally, a computer simulation confirms these analytical results.

To further extend these results to the case of a test that was synthesized under the expectation of detecting an FH waveform with extremely small SNR, an asymptotic analysis of a different sort is undertaken. This analysis shows how the above tests perform for the composite hypothesis problem as the minimum reliably detectable SNR of the FH waveform becomes increasingly small. Numerical results for this case are given, but a corresponding computer simulation is not possible due to the rate of increase of the number of computations required as the SNR diminishes.

2.2   Likelihood Function: One Epoch

The statistical test for the composite hypothesis problem is defined by finding an asymptotically optimal test for a binary hypothesis problem and applying that test to the composite case and accepting the resulting degradation. This simplified binary problem consists of the two hypotheses H_0, where no signal is present, and H_γ, where a signal is present with SNR γ.

For this binary hypothesis problem, Appendix 2.A contains a derivation of the SELF, which is the likelihood function Λ_i of the ith-epoch observation $y(t)$ for $iT_h < t \le (i+1)T_h$. The SELF is expressed as

$$\Lambda_i(y) \quad \triangleq \quad \mathcal{E}_k\left[\Lambda_i(y/k)\right] \tag{2.2}$$

$$= \quad \frac{e^{-\gamma}}{K} \sum_{k=0}^{K-1} I_0\left(\sqrt{2\gamma}\sqrt{P_k^2 + Q_k^2}\right) \tag{2.3}$$

where I_0 is the zeroth order modified Bessel function of the first kind and

$$P_k = \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} y(t) \cos \omega_{k_i} t \, dt$$

$$Q_k = \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} y(t) \sin \omega_{k_i} t \, dt. \tag{2.4}$$

Because of the statistical independence of their respective observations, the likelihood function of the n-epoch observation is then $\prod_{i=1}^{n} \Lambda_i$, i.e., the product of these individual SELFs.

The SELF is nicely modeled as the configuration of well-known devices, as indicated in Figure 2.1. That is, the SELF is channelized so that each channel has a matched filter that is tuned to a particular hop frequency and whose output is envelope detected and emphasized by a Bessel function non-linearity. The output of each channel, after scaling by $e^{-\gamma}/K$, is summed to produce the SELF.
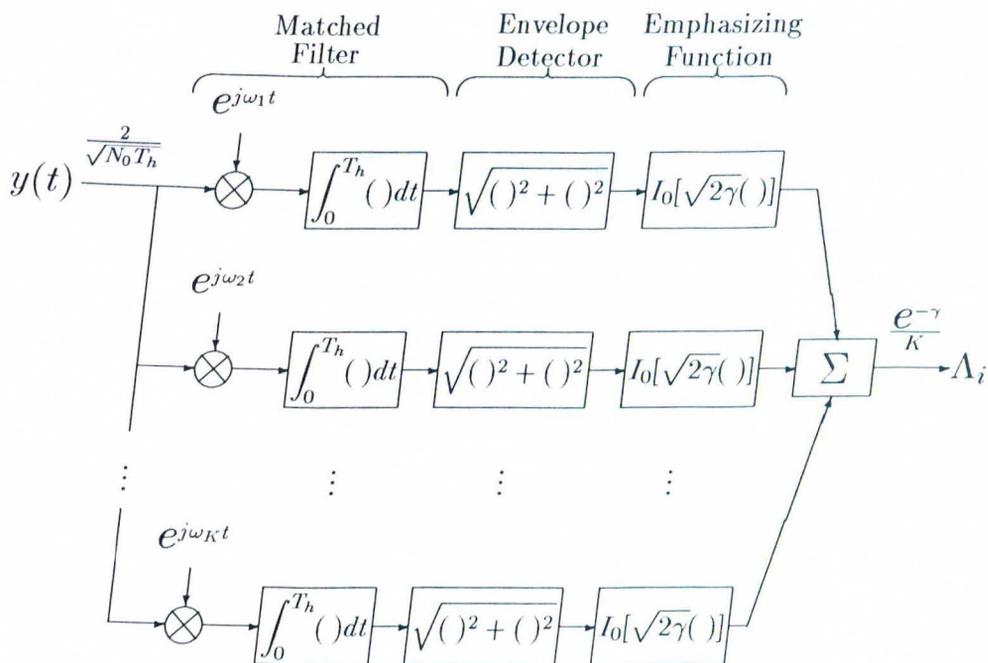
Figure 2.1: Block Diagram of Single-Epoch Likelihood Function

2.3 Asymptotic Log-Likelihood Function

The Asymptotic Log-Likelihood Function (ALLF) is asymptotic to the n-epoch likelihood function, $\prod_{i=1}^{n} \Lambda_i$, as the number of FH channels becomes large. The critical idea behind the derivation of the ALLF is the application of the central limit theorem to yield asymptotic densities for the SELF from which, using an n-epoch collection of SELFs as an equivalent observation set, the ALLF will be determined.

The SELF (2.3) was computed assuming a binary hypothesis problem, i.e., H_0 is the hypothesis that no FH waveform is present, while H_γ is the

28

hypothesis that an FH waveform exists with a known SNR γ. The following analysis will assume that an FH waveform, if present, will have a SNR γ' or equivalently an average signal energy S' that is not necessarily equal to the average signal energy S assumed known in the binary case. This generalization is not necessary for deriving the ALLF but will be needed to analyze the performance of the ALLF in the composite hypothesis problem.

Proceeding with the derivation of the ALLF, the central limit theorem is applied to the SELF to obtain an asymptotic density under all hypotheses, $0 \leq \gamma'$. The central limit theorem is justified here because the SELF's output is the sum of many channels whose statistics will be shown to be nearly independent and nearly identical. It will be shown that the degree of dependence between channels is determined by the amount of isolation between channels, which is perfect for minimally spaced channels as is the case assumed here. It also will be shown that the channel means and variances, while different for the signal-present and signal-absent cases, are of a commensurate magnitude.

2.3.1 Matched Filter Output Statistics

Because the central limit theorem requires only the mean and variance of each channel, only the statistics of the matched filter outputs need be determined exactly since the SELF's mean and variances can be determined from

these statistics alone. Assuming that the signal present is in the kth channel, then the matched filter output in the lth channel can be found from (2.4) as

$$P_l \approx \begin{cases} \sqrt{2\gamma'}\sin\theta + \nu_l & \text{for } l = k \\ \nu_l & \text{for } l \neq k \end{cases} \tag{2.5}$$

$$Q_l \approx \begin{cases} \sqrt{2\gamma'}\cos\theta + \xi_l & \text{for } l = k \\ \xi_l & \text{for } l \neq k \end{cases} \tag{2.6}$$

where

$$\nu_l \triangleq \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} n(t)\cos\omega_l t \, dt$$

$$\tag{2.7}$$

$$\xi_l \triangleq \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} n(t)\sin\omega_l t \, dt.$$

The matched filter outputs for the no-signal-present hypothesis H_0 are the special case of the above expressions for $\gamma' = 0$. Two assumptions were made in determining these approximate expressions for the matched filter outputs. The first assumption is that $\omega_k T_h$ is large and equivalent to requiring a large number of carrier cycles over a single epoch. The second assumption of orthogonally spaced channels [i.e. $(\omega_k - \omega_l)T_h/2\pi$ is an integer] implies, in essence, that the channels are isolated from one another. Another condition implying channel isolation is wide spacing between the channels [i.e. $(\omega_k - \omega_l)T_h$ is large]. In a practical implementation, smooth window functions also could have been used in the matched filter implementation to achieve the channel isolation assumed here.

Simplified expressions for the matched filter outputs are represented by (2.5) and (2.6). The statistical nature of their noise components, $\{\nu_l\}$ and

$\{\xi_l\}$, is determined next. From (2.7), it follows that the random variables

$\{\nu_l\}, \{\xi_l\}$ are Gaussian with zero mean and unity variance. Under the isolated

channel assumption, it is easy to show that

$$\begin{aligned} \mathcal{E}[\nu_m \nu_n] &= 0 &&\text{for } m \neq n & 1 \leq m,n \leq K \\ \mathcal{E}[\nu_m \xi_n] &= 0 &&\text{for all } m,n & 1 \leq m,n \leq K \\ \mathcal{E}[\xi_m \xi_n] &= 0 &&\text{for } m \neq n & 1 \leq m,n \leq K. \end{aligned} \quad (2.8)$$

Thus $\{\nu_l\}, \{\xi_l\}$ are mutually independent, since they are Gaussian. These

relations also determine the joint density of ν_l and ξ_l as

$$p_{\nu_l, \xi_l}(\nu_l, \xi_l) = \frac{1}{2\pi} e^{-\frac{1}{2}(\nu_l^2 + \xi_l^2)}. \quad (2.9)$$

The equations (2.5) and (2.6) and (2.8), along with the joint density of ν_l

and ξ_l (2.9), constitute a complete statistical description of the matched filter

outputs $\{P_l\}$ and $\{Q_l\}$.

2.3.2   SELF Moments

The statistics of $\{P_l\}$ and $\{Q_l\}$ were found in order to determine the mean

and variance of the SELF (2.3). The SELF moments are needed to apply the

central limit theorem and thus ultimately produce the ALLF. A few conditions

for the application of the central limit theorem will be established now. First,

since the random variables $\{P_l\}$ and $\{Q_l\}$ are mutually independent, each

channel output of the SELF is also independent. Furthermore, the channel

outputs are all identically distributed except for the output of the channel

with the signal present. This particular channel output will be shown to have

31

a variance comparable to that of the other channel outputs and thus the central limit theorem still applies and with it we get a density asymptotic to the actual SELF density.

To continue, we need explicit expressions for the mean and variance of the SELF. Assuming a signal is present with a relative SNR of $r = \sqrt{\gamma'/\gamma}$, then the matched filter outputs of the channel containing the signal are by (2.5) and (2.6)

$$
\begin{aligned}
P_l &= \sqrt{2\gamma'}\sin\theta + \nu_l \\
Q_l &= \sqrt{2\gamma'}\cos\theta + \xi_l.
\end{aligned}
\tag{2.10}
$$

If μ_r and σ_r^2 are defined to be the mean and variance for this channel output, then (2.9) implies

$$
\mu_r = \mathcal{E}\left[I_0\left(\sqrt{2\gamma}\sqrt{P_l^2 + Q_l^2}\right)\right]
\tag{2.11}
$$

$$
= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_0\left(\sqrt{2\gamma}\sqrt{P_l^2 + Q_l^2}\right)e^{-\frac{1}{2}(\nu_l^2 + \xi_l^2)}d\nu_l d\xi_l.
\tag{2.12}
$$

With the rectangular-to-polar conversion, $P_l = \rho\cos\phi$, $Q_l = \rho\sin\phi$, and applying the identity

$$
I_0(a) = \frac{1}{2\pi}\int_0^{2\pi} e^{a\cos\phi}\, d\phi
\tag{2.13}
$$

the integral becomes

$$
\mu_r = e^{-\gamma'}\int_0^{\infty} \rho I_0\left(\sqrt{2\gamma}\rho\right) I_0\left(\sqrt{2\gamma'}\rho\right)e^{-\frac{\rho^2}{2}}d\rho
\tag{2.14}
$$

$$
= e^{\gamma} I_0\left(2\sqrt{\gamma\gamma'}\right).
\tag{2.15}
$$

32

This integral was evaluated in [49], §13.31(1). The variance is now evaluated as follows:

$$\sigma_r^2 + \mu_r^2 = \mathcal{E}\left[I_0^2\left(\sqrt{2\gamma}\sqrt{P_l^2 + Q_l^2}\right)\right] \tag{2.16}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_0^2\left(\sqrt{2\gamma}\sqrt{P_l^2 + Q_l^2}\right)e^{-\frac{1}{2}(\nu_l^2+\xi_l^2)}d\nu_l d\xi_l \tag{2.17}$$

which becomes with rectangular-to-polar conversion, and applying (2.13),

$$\sigma_r^2 + \mu_r^2 = e^{-\gamma'}\int_0^{\infty} \rho I_0^2\left(\sqrt{2\gamma}\rho\right) I_0\left(\sqrt{2\gamma'}\rho\right) e^{-\frac{\rho^2}{2}} d\rho. \tag{2.18}$$

This integral is evaluated by applying a formula from [49], §11.41(16), which states

$$\frac{1}{\pi}\int_0^{\pi} I_0\left(\sqrt{a^2 + b^2 - 2ab\cos\phi}\right) d\phi = I_0(a)I_0(b). \tag{2.19}$$

Application of this formula and an interchange of integrations reduces the integral (2.18) to a simpler integral solved in [49], §13.31(1). The net result is

$$\sigma_r^2 + \mu_r^2 = \frac{e^{2\gamma}}{\pi}\int_0^{\pi} e^{-2\gamma\cos\phi} I_0\left(4\sqrt{\gamma\gamma'}\sin\frac{\phi}{2}\right) d\phi. \tag{2.20}$$

Summarizing, for a signal in channel l with a SNR γ', the channel moments are

$$\begin{aligned} \mu_r &= e^{\gamma} I_0(2r\gamma) \\ \sigma_r^2 &= e^{2\gamma}\left[\frac{1}{\pi}\int_0^{\pi} e^{-2\gamma\cos\phi} I_0\left(4r\gamma\sin\frac{\phi}{2}\right) d\phi - I_0^2\left(2r\gamma\right)\right]. \end{aligned} \tag{2.21}$$

where $r = \sqrt{\gamma'/\gamma}$.

The above calculations give expressions for the channel moments for a channel with a signal present. The moments for the case of a channel without

a signal present are special cases of the above with $r = 0$ and are thus denoted by μ_0 and σ_0^2. From (2.21) and the Bessel function identity (2.13), they are

$$
\begin{aligned}
\mu_0 &= e^{\gamma} \\
\sigma_0^2 &= e^{2\gamma} \left[I_0(2\gamma) - 1 \right].
\end{aligned}
\tag{2.22}
$$

Likewise, moments for the lth channel (whenever it contains a signal with strength γ) correspond to the above moments with $r = 1$ and are thus denoted by μ_1 and σ_1^2.

As previously mentioned, the application of the central limit theorem depends on the various channel means and variances having commensurate amplitudes. The relative amplitudes between the moments are computed from (2.21) and (2.22) and are

$$
\frac{\mu_r}{\mu_0} = I_0(2r\gamma)
\tag{2.23}
$$

$$
\frac{\sigma_r^2}{\sigma_0^2} = 1 + 2r^2\gamma + 2r^2\gamma^2 + \mathcal{O}(\gamma^3).
\tag{2.24}
$$

Hence, for small assumed SNRs ($\gamma \leq 1$), the mean and variance of the channel with a signal present and the mean and variance of the channels without a signal present are within a factor of three of each other.

Expressions for the mean and variance of the SELF are now immediate, since the SELF is the sum of all K channels scaled by $e^{-\gamma}/K$. The expressions are

$$
M_r = \frac{e^{-\gamma}}{K} \left[(K - 1)\mu_0 + \mu_r \right]
\tag{2.25}
$$

34

$$V_r = \frac{e^{-2\gamma}}{K^2}\left[(K-1)\sigma_0^2 + \sigma_r^2\right].\qquad(2.26)$$

Here, M_r is the mean of the SELF when a signal of strength γ' is present and M_0 and M_1 are written for the special cases of M_r, when $r = 0$ and $r = 1$, respectively. The variances V_r, V_0, and V_1 are defined similarly.

2.3.3 Derivation of the ALLF

With the first two moments of the SELF determined, the central limit theorem gives approximating densities to the SELF, Λ_i, under the composite hypothesis problem. These densities are

$$\begin{aligned}
H_0: \quad & \Lambda_i \sim \frac{1}{\sqrt{2\pi V_0}}e^{-\frac{(\Lambda_i - M_0)^2}{2V_0}} \\
\text{versus} \quad & \\
H_{\gamma'}: \quad & \Lambda_i \sim \frac{1}{\sqrt{2\pi V_r}}e^{-\frac{(\Lambda_i - M_r)^2}{2V_r}} \qquad \text{for } 0 < \gamma' \leq \gamma
\end{aligned}\qquad(2.27)$$

which give a simplified statistical characterization of the SELF. That is, the SELF outputs, $\{\Lambda_i\}$, are Gaussian i.i.d. variables whose means and variances depend on the hypothesis.

As was the procedure in deriving the SELF, the Asymptotic Log-Likelihood Function (ALLF) is designed using the simpler binary hypothesis problem. For a single-epoch, likelihood function theory and (2.27) imply a log-likelihood function of

$$L_i(\Lambda_i) = c_2\Lambda_i^2 + c_1\Lambda_i + c_0\qquad(2.28)$$

where

$$c_2 = \frac{1}{2}\left(\frac{1}{V_0} - \frac{1}{V_1}\right) \tag{2.29}$$

$$c_1 = \left(\frac{M_1}{V_1} - \frac{M_0}{V_0}\right) \tag{2.30}$$

$$c_0 = \frac{1}{2}\left(\frac{M_0^2}{V_0} - \frac{M_1^2}{V_1} + \ln\frac{V_0}{V_1}\right). \tag{2.31}$$

Independence between observations over different epochs implies that the ALLF

up to time n is

$$T_n = \sum_{i=1}^{n} L_i. \tag{2.32}$$

Now that ALLF has been found, its mean and variance will be computed as a

prelude to investigating its performance in the composite hypothesis problem.

2.3.4 Moments of the ALLF

For the analysis that follows, it is useful to derive the moments of L_i from

which the ALLF moments follow trivially from (2.32), starting with the mean

$$\mathcal{M}_r \triangleq E\left[L_i\left(\Lambda_i\right)\right] \tag{2.33}$$

$$= c_2 E(\Lambda_i^2) + c_1 E(\Lambda_i) + c_0 \tag{2.34}$$

$$= c_2\left(M_r^2 + V_r\right) + c_1 M_r + c_0 \tag{2.35}$$

which expands in terms of the SELF moments to

$$\mathcal{M}_r = \frac{1}{2} \ln \frac{V_0}{V_1} + (2V_1V_0)^{-1} \left[(M_r - M_0)^2 \, V_1 - (M_r - M_1)^2 \, V_0 + (V_1 - V_0) \, V_r \right].$$

(2.36)

Now to compute the variance of L_i.

$$\mathcal{V}_r = Var\left[L_i\left(\Lambda_i \right) \right]$$

(2.37)

which upon substitution of (2.28) yields

$$\mathcal{V}_r = Var\left[c_2 \Lambda_i^2 + c_1 \Lambda_i + c_0 \right]$$

(2.38)

$$= Var\left[(c_2 M_r^2 + c_1 M_r + c_0) + (2c_2 M_r + c_1)v + c_2 v^2 \right]$$

(2.39)

where $v = \Lambda_i - M_r$. Proceeding,

$$\mathcal{V}_r = Var\left[(2c_2 M_r + c_1)\, v + c_2 v^2 \right]$$

(2.40)

$$= (2c_2 M_r + c_1)^2 \, V_r + 2c_2^2 V_r^2$$

(2.41)

which simplifies to

$$\mathcal{V}_r = \frac{V_r^2}{2} \left(\frac{1}{V_0} - \frac{1}{V_1} \right)^2 + \left[\left(\frac{1}{V_0} - \frac{1}{V_1} \right) M_r + \left(\frac{M_1}{V_1} - \frac{M_0}{V_0} \right) \right]^2 V_r.$$

(2.42)

The special cases, $r = 1$ and $r = 0$, of the moments of L_i are respectively written as $\mathcal{M}_1$ and $\mathcal{M}_0$ for the means and as $\mathcal{V}_1$ and $\mathcal{V}_0$ for the variances.

2.3.5 Summary

A log-likelihood function for the binary hypothesis problem, designated the ALLF, has been derived that is asymptotic to the true log-likelihood function as the number of channels becomes large. The ALLF was found with the help of the likelihood function theory by considering an n-epoch collection of SELFs as a set of i.i.d. observations assumed Gaussian by the central limit theorem. The Gaussian assumption was justified by showing that each SELF was the sum of nearly independent and nearly identical random variables. Various means and variances were also derived that will prove useful in future discussions. The ALLF now will be used to design an FSS test, an SPRT, and a TST.

2.4 Test Design

The results above reduced the problem of detecting an FH waveform to that of discriminating between two sets of Gaussian i.i.d. data with different means and variances. A Fixed Sample Size (FSS) test, a Sequential Probability Ratio Test (SPRT), and a Truncated Sequential Test (TST) based on this simplified model will be discussed.

2.4.1 FSS Test Design

As the name suggests, an FSS test consists of comparing a test statistic T_L, based on a fixed number of observations L, to a threshold τ. Then, if the test statistic is greater than τ, hypothesis H_1 is chosen, while a test statistic less than τ indicates hypothesis H_0. Symbolically this is

$$T_L \begin{cases} \geq \tau & \Rightarrow & H_1 \\ < \tau & \Rightarrow & H_0. \end{cases} \tag{2.43}$$

In our case, the test statistic is the L-epoch ALLF and the test parameters L and τ are specified to correspond to prescribed false alarm P_F and detection P_D probabilities. To determine L and τ, the density of the T_L is needed for each hypothesis. Although this density equals the non-central χ^2 density, an approximate Gaussian density, derived via the central limit theorem, is used instead to yield simplified expressions for the test parameters. These densities are

$$H_0: \quad T_L \quad \sim \quad \frac{1}{\sqrt{2\pi L \mathcal{V}_0}} e^{-\frac{(T_L - L\mathcal{M}_0)^2}{2L\mathcal{V}_0}}$$
$$\text{versus} \qquad \qquad \qquad \qquad \qquad \qquad \tag{2.44}$$
$$H_1: \quad T_L \quad \sim \quad \frac{1}{\sqrt{2\pi L \mathcal{V}_1}} e^{-\frac{(T_L - L\mathcal{M}_1)^2}{2L\mathcal{V}_1}}.$$

From these densities, P_D and P_F can be computed in terms of L and τ to yield $P_F = 1 - \Phi\left(\frac{\tau - L\mathcal{M}_0}{\sqrt{L\mathcal{V}_0}}\right)$ and $P_D = 1 - \Phi\left(\frac{\tau - L\mathcal{M}_1}{\sqrt{L\mathcal{V}_1}}\right)$ where Φ^{-1} is the inverse of the distribution function of a zero-mean, unity-variance Gaussian random variable. These are solved simultaneously to arrive at

$$L = \frac{\left[\mathcal{V}_1^{\frac{1}{2}} \Phi^{-1}(1 - P_D) - \mathcal{V}_0^{\frac{1}{2}} \Phi^{-1}(1 - P_F)\right]^2}{(\mathcal{M}_1 - \mathcal{M}_0)^2} \tag{2.45}$$

$$\tau = \frac{L^{\frac{1}{2}}}{(\mathcal{M}_1 - \mathcal{M}_0)} \left[\mathcal{V}_0^{\frac{1}{2}} \mathcal{M}_1 \Phi^{-1}(1 - P_F) - \mathcal{V}_1^{\frac{1}{2}} \mathcal{M}_0 \Phi^{-1}(1 - P_D) \right]. \quad (2.46)$$

2.4.2 SPRT Design

Wald's sequential probability ratio test (SPRT) now can be defined as a test with test statistic T_n, based on n observations and two thresholds a and b. The SPRT works as follows. Upon the nth observation, if T_n is greater than a, then hypothesis H_1 is chosen. If T_n is less than b, then hypothesis H_0 is chosen. If, instead, T_n is between a and b, the test statistic is updated to include $n + 1$ observations and the process is iterated. Symbolically this test is described as

$$\text{for each } n, \ T_n \begin{cases} \geq a & \Rightarrow \quad H_1 \\ \leq b & \Rightarrow \quad H_0 \\ \in (a, b) & \Rightarrow \quad \text{take another sample.} \end{cases} \quad (2.47)$$

The threshold values a and b are assigned to give the desired Neyman-Pearson probability of detection P_D and probability of false alarm P_F. Relationships between the thresholds and these probabilities are given by Wald's approximations [48]

$$a \approx \ln\left(\frac{P_D}{P_F}\right) \quad (2.48)$$

$$b \approx \ln\left(\frac{1 - P_D}{1 - P_F}\right). \quad (2.49)$$

2.4.3 TST Design

Truncated Sequential Test (TST) is a hybrid of the above two tests. Specifically, TST follows the rules of a sequential test with test statistic T_n and with thresholds a and b, but has the added feature of forcing a decision at time L (if no decision has been made) by comparing the test statistic to a threshold τ. Symbolically,

$$\text{for each } n < L, \quad T_n \begin{cases} \geq a & \Rightarrow H_1 \\ \leq b & \Rightarrow H_0 \\ \in (a, b) & \Rightarrow \text{ take another sample} \end{cases} \tag{2.50}$$

$$\text{but for } n = L, \quad T_L \begin{cases} \geq \tau & \Rightarrow H_1 \\ < \tau & \Rightarrow H_0. \end{cases}$$

Two relations secure the specification of the TST parameters a, b, L, and τ. If P_F^* and P_D^* are the actual Neyman-Pearson probabilities for the TST, then from [43]

$$P_F^* \leq P_F^{FSS} + P_F^{SPRT} \tag{2.51}$$

$$(1 - P_D^*) \leq (1 - P_D^{FSS}) + (1 - P_D^{SPRT}) \tag{2.52}$$

where P_F^{FSS} is the probability of false alarm for the TST, if $L = \infty$, and P_F^{SPRT} is the false-alarm probability for the TST, if $a = -b = \infty$. P_D^{FSS} and P_D^{SPRT} are defined similarly. Thus the errors of the TST can be viewed as a mixture of the errors of an FSS test with parameters L and τ and an SPRT with parameters a and b. These inequalities can be verified by viewing the ALLF T_n as a discrete stochastic process with time index n and enumerating

its sample paths. For instance, a sample path leading to a false alarm must either cross threshold a before threshold b and before time L or be greater than threshold τ at time L. Since these events also correspond to false alarms in either the FSS test part or the SPRT part of the TST, the inequality (2.51) must follow.

The above inequalities can be used to specify a TST, whose actual error probabilities P_F^* and $1 - P_D^*$ are less than any specified error probabilities P_F and $1 - P_D$. Thus, the TST can be designed by partitioning the bounding errors $(1 - P_D)$ and P_F among the SPRT and FSS test parts of the TST and then using the appropriate equation to compute the parameters L, τ, a, and b for TST [43]. Specifically, this partitioning is quantified with the introduction of two constants, $0 \leq C_1 \leq 1$ and $0 \leq C_2 \leq 1$, which are defined as TST mixture constants, then

$$P_F^{FSS} = C_1 P_F \qquad (2.53)$$

$$P_F^{SPRT} = (1 - C_1) P_F \qquad (2.54)$$

$$(1 - P_D^{FSS}) = C_2(1 - P_D) \qquad (2.55)$$

$$(1 - P_D^{SPRT}) = (1 - C_2)(1 - P_D) \qquad (2.56)$$

for the error probabilities of the FSS test and SPRT parts of the TST. From the above inequalities and (2.45), (2.46), (2.48), and (2.49), the TST parameters

42

are determined as follows:

$$L = \frac{\left[\mathcal{V}_1^{\frac{1}{2}}\Phi^{-1}(1 - P_D^{FSS}) - \mathcal{V}_0^{\frac{1}{2}}\Phi^{-1}(1 - P_F^{FSS})\right]^2}{(\mathcal{M}_1 - \mathcal{M}_0)^2} \tag{2.57}$$

$$\tau = \frac{L^{\frac{1}{2}}}{(\mathcal{M}_1 - \mathcal{M}_0)}\left[\mathcal{V}_0^{\frac{1}{2}}\mathcal{M}_1\Phi^{-1}(1 - P_F^{FSS}) - \mathcal{V}_1^{\frac{1}{2}}\mathcal{M}_0\Phi^{-1}(1 - P_D^{FSS})\right]$$

$$\tag{2.58}$$

$$a = \ln\left(\frac{P_D^{SPRT}}{P_F^{SPRT}}\right) \tag{2.59}$$

$$b = \ln\left(\frac{1 - P_D^{SPRT}}{1 - P_F^{SPRT}}\right). \tag{2.60}$$

Note that (2.51) and (2.52) guarantee that the actual detection errors

$$P_F^* \leq P_F \tag{2.61}$$

$$1 - P_D^* \leq 1 - P_D. \tag{2.62}$$

The mixture constants C_1 and C_2 reflect proportions of the FSS test and SPRT parts of the TST since, if $C_1 = C_2 = 1$, a pure FSS test is defined, and if $C_1 = C_2 = 0$, a pure SPRT is defined. Criteria for choosing the mixture constants will be discussed in Section 2.6.

2.5 Performance of Tests

The problem addressed by the preceding tests (the FSS test, the SPRT,

and the TST) is the detection of an FH waveform. The detection of the FH waveform is a prelude to other interception processes, such as feature detection, channel tracking, and message extraction. Here the performance of the tests in detecting an FH waveform with variable amplitude and in the presence of white Gaussian noise is quantified.

The three tests were designed under the assumption of binary hypotheses. These hypotheses are H_0 (FH waveform is not present) and H_γ (FH waveform is present and has SNR γ). Of concern here is the performance of the three tests when the actual SNR γ' of the FH waveform is more generally $0 < \gamma' \leq \gamma$. Two parameters characterize a test's performance for a particular γ'. The first, denoted by $E(N/r, \gamma)$, is the Average Sample Number (ASN) defined as the average of the number of samples needed to reach a decision. The second parameter, denoted by $P_0(r, \gamma)$, is the Operating Characteristic (OC) defined as the probability of declaring the absence of an FH waveform. Both the ASN and OC are defined as functions of relative SNR r and the assumed SNR γ.

2.5.1 Analysis of FSS Test

For the FSS test, the ASN is obviously L, while the OC can be determined by approximating the ALLF at time L by a Gaussian random variable with

44

the same moments. This central-limit-theorem argument produces

$$P_0(r, \gamma) = \Phi\left(\frac{\tau - L\mathcal{M}_r}{\sqrt{L\mathcal{V}_r}}\right) \tag{2.63}$$

for the OC.

2.5.2 Analysis of SPRT

For the SPRT, the analysis is more difficult but can be approached as a diffusion problem. Here we approximate the test statistic by a Wiener process. Specifically, if $T(t)$ is a Wiener process with variance function $\mathcal{V}_r t$ and mean function $\mathcal{M}_r t$, then the ALLF, T_n, converges weakly to $T(t)$ at integer times $t = n$, provided n is sufficiently large. This last restriction is needed to ensure that T_n has an approximate Gaussian density as implied by the central limit theorem. In terms of the approximating Wiener process $T(t)$, the problem of finding the OC function is now the problem of finding the probability that $T(t)$ will "touch" the lower threshold b before the upper threshold a. Likewise, the problem of finding the ASN is now the problem of finding the average time that $T(t)$ first "touches" either threshold (a or b). This time is also called the average stopping time. Expressions for these quantities are given in [6] and [1] as

$$P_0(r, \gamma) = \begin{cases} \dfrac{e^{-2b\frac{\mathcal{M}_r}{\mathcal{V}_r}} - 1}{e^{-2b\frac{\mathcal{M}_r}{\mathcal{V}_r}} - e^{-2a\frac{\mathcal{M}_r}{\mathcal{V}_r}}} & \mathcal{M}_r \neq 0 \\[2ex] \dfrac{a}{a - b} & \mathcal{M}_r = 0 \end{cases} \tag{2.64}$$

45

$$
\mathcal{E}(N/r,\gamma) \;=\; \begin{cases} \dfrac{aP_0(r,\gamma) + b\left[1 - P_0(r,\gamma)\right]}{\mathcal{M}_r} & \mathcal{M}_r \neq 0 \\[3mm] -\dfrac{ab}{\mathcal{V}_r} & \mathcal{M}_r = 0 \end{cases} \;. \qquad (2.65)
$$

2.5.3 Analysis of TST

The diffusion analysis technique also applies to the TST but is more involved. The ASN is by [6]

$$
E(N/r,\gamma) = A \sum_{n=1}^{\infty} (-1)^n \frac{n}{k_n^2} B_n (e^{-k_n L} - 1) \qquad (2.66)
$$

where

$$
A \;=\; \frac{\mathcal{V}_r \pi}{(a-b)^2} \qquad (2.67)
$$

$$
B_n \;=\; e^{\frac{\mathcal{M}_r b}{\mathcal{V}_r}} \sin \frac{n\pi a}{a-b} - e^{\frac{\mathcal{M}_r a}{\mathcal{V}_r}} \sin \frac{n\pi b}{a-b} \qquad (2.68)
$$

$$
k_n \;=\; \frac{\mathcal{M}_r^2}{2\mathcal{V}_r} + \frac{\mathcal{V}_r n^2 \pi^2}{2(a-b)^2}. \qquad (2.69)
$$

The OC function is defined by [1] as

$$
\begin{aligned}
P_0(r,\gamma) \;=\; & \Phi\left(\frac{\tau - L\mathcal{M}_r}{\sqrt{L\mathcal{V}_r}}\right) \\
& - \sum_{n=1}^{\infty} e^{2\frac{\mathcal{M}_r}{\mathcal{V}_r}[na-(n-1)b]} \Phi\left(\frac{\tau - L\mathcal{M}_r - 2[na-(n-1)b]}{\sqrt{L\mathcal{V}_r}}\right) \\
& + e^{2\frac{\mathcal{M}_r}{\mathcal{V}_r}n(a-b)} \Phi\left(\frac{\tau - L\mathcal{M}_r - 2n(a-b)}{\sqrt{L\mathcal{V}_r}}\right) \\
& + e^{2\frac{\mathcal{M}_r}{\mathcal{V}_r}[nb-(n-1)a]} \Phi\left(\frac{2[nb-(n-1)a] - \tau + L\mathcal{M}_r}{\sqrt{L\mathcal{V}_r}}\right)
\end{aligned}
$$

$$-e^{2\frac{\mathcal{M}_r}{\mathcal{V}_r}n(b-a)}\Phi\left(\frac{2n(b-a)-\tau+L\mathcal{M}_r}{\sqrt{L\mathcal{V}_r}}\right). \qquad (2.70)$$

Equations (2.63) through (2.70) represent a complete characterization of the performance of the FSS test, the SPRT, and the TST.

The fact that the diffusion technique yields accurate expressions for the ASN and OC functions will not be proved here but will be verified below by computer simulation.

2.5.4 Numerical Results

The FSS test, the SPRT, and the TST were simulated by the computer to verify the assumptions of the analysis and as an independent measure of the relative performance of the three tests. The simulated detector consisted of 512 channels and each test was synthesized to ensure a probability of false alarm P_F of no more than 1% and a probability of detection P_D of at least 99%. Here the fairly relaxed probability of false alarm of 1% was chosen in order to limit the number of data needed for a decision. Under these specifications, the simulation was run until 1,000 decisions were reached for each of 11 SNRs evenly spaced between 0 and γ. The decisions that no FH waveform was present were averaged to estimate the OC, while the number of observations taken to reach a decision was averaged to estimate the ASN. Additionally, the standard deviation of ASN average was measured to indicate the ASN

estimation error.

Figure 2.2 and Figure 2.3 are, respectively, the ASNs by simulation and by theory when the assumed SNR $\gamma = 1$, while Figure 2.4 and Figure 2.5 are the corresponding curves for ($\gamma = 0.3$). As predicted, the ASN is greatly reduced, by about 57%, for the SPRT in the regions around $\gamma' = 0$ and $\gamma' = \gamma$. These curves exemplify a general property of the SPRT: to perform very well when the observation statistics are close to those assumed, but to exhibit a degraded performance, often to the point of being worse than the FSS test, when the observation statistics are different. In our context, this degradation is evidenced by a large ASN for the SPRT, when the actual SNR γ' is midway between the two assumed values 0 and γ. The TST reduces the ASN in this mid-SNR region, as shown by the figures, but it does so at the expense of performance in the regions around $\gamma' = 0$ and $\gamma' = \gamma$. Despite this performance loss, truncation is necessary for implementation reasons. It also will be shown that the TST has the desirable property of having a higher detection probability than the SPRT at small SNRs and that, through optimization of the mixture constants, the TST can regain much of what it lost in ASN around $\gamma' = 0$ and $\gamma' = \gamma$.

Focusing on the OCs (Figures 2.6 and 2.7 for $\gamma = 1$, Figures 2.8 and 2.9 for $\gamma = 0.3$),it is obvious that the FSS test has slightly higher probability of detection for small SNRs while the SPRT has degraded performance in this

48
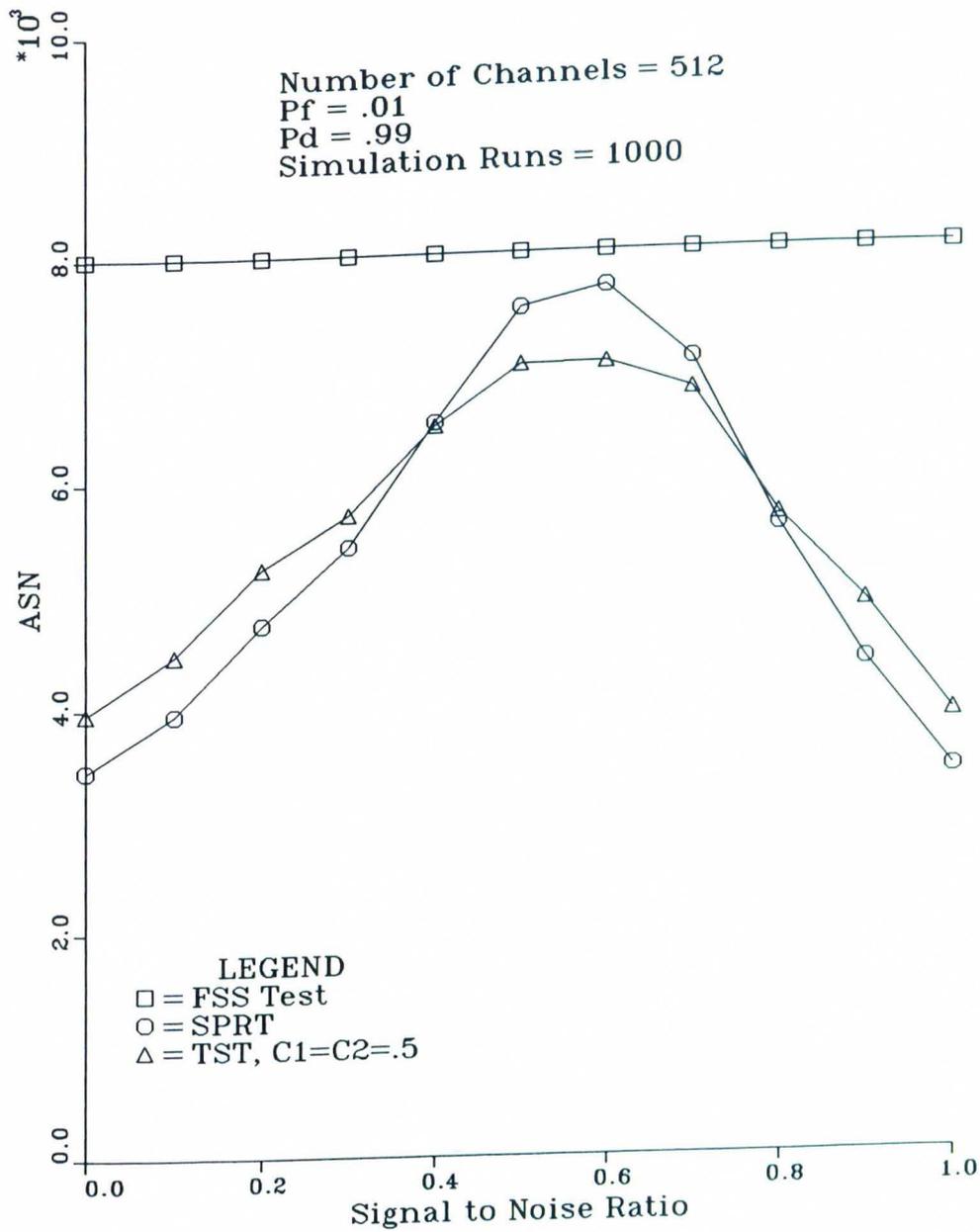
Figure 2.2: ASN from Simulation versus SNR, $\gamma = 1$
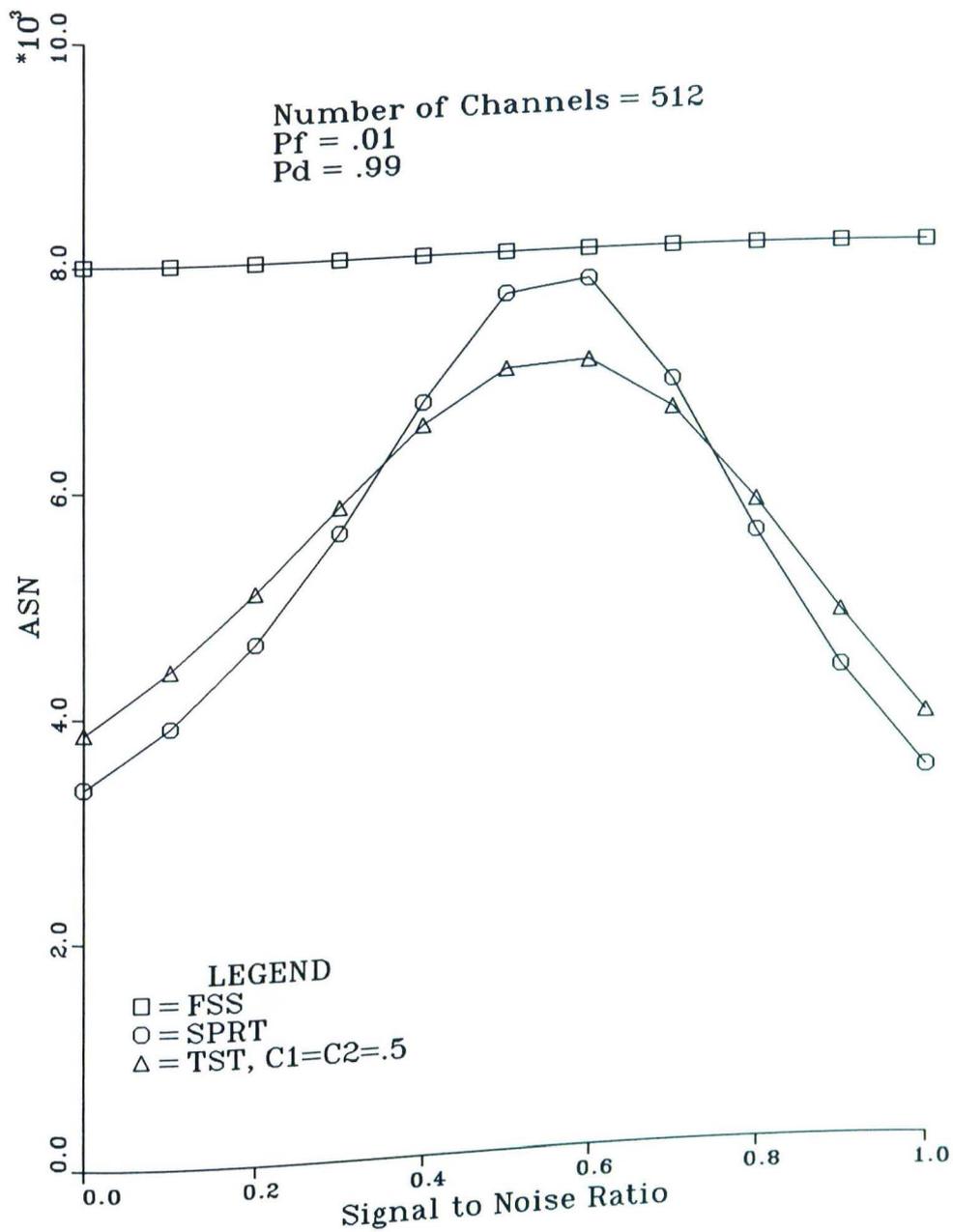
Figure 2.3: ASN from Theory versus SNR, $\gamma = 1$
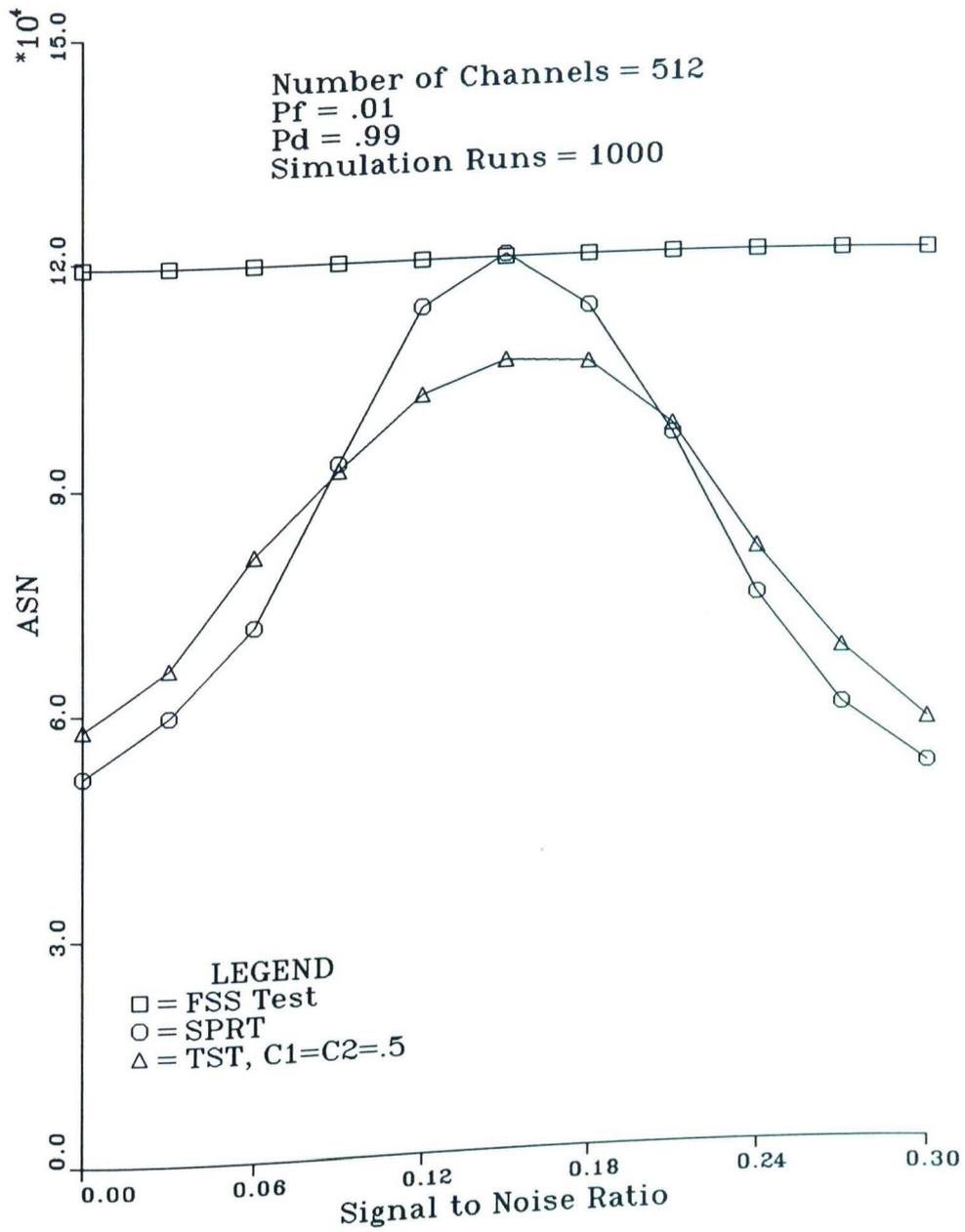
Figure 2.4: ASN from Simulation versus SNR, $\gamma = 0.3$

$*10^4$

Number of Channels = 512
Pf = .01
Pd = .99

ASN

Signal to Noise Ratio

LEGEND
□ = FSS
○ = SPRT
△ = TST, C1=C2=.5
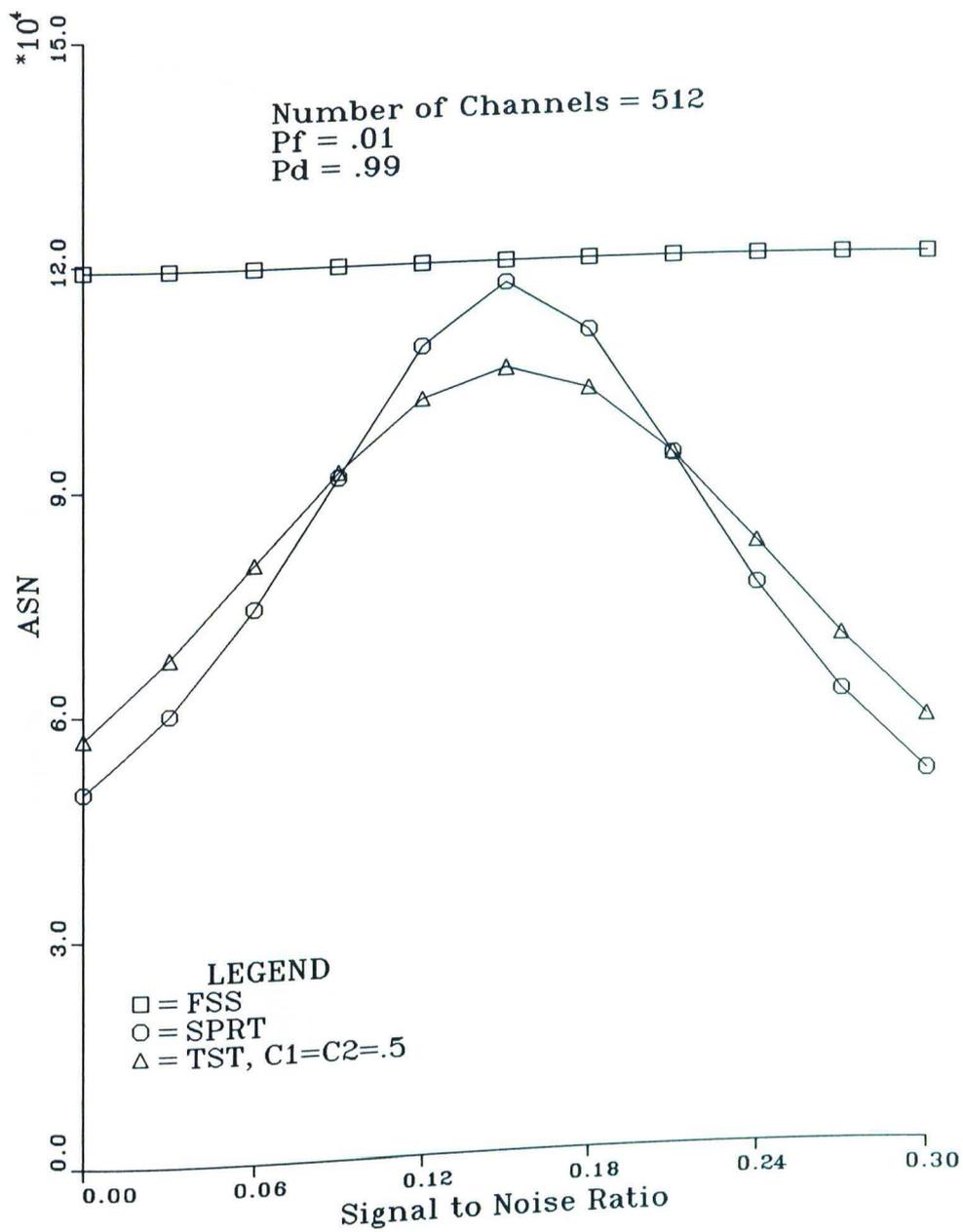
Figure 2.5: ASN from Theory versus SNR, $\gamma = 0.3$
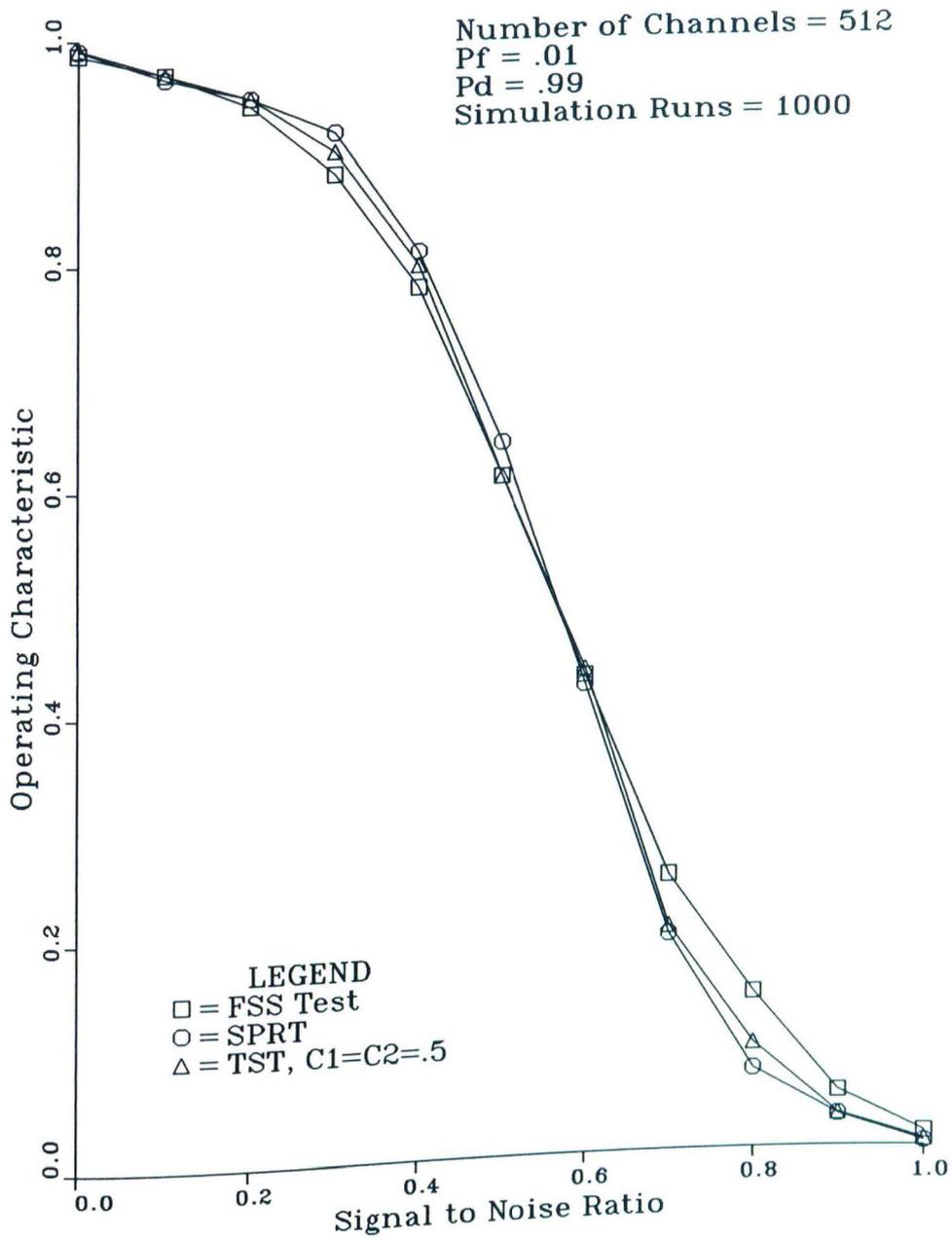
Figure 2.6: OC from Simulation, $\gamma = 1$

Number of Channels = 512
Pf = .01
Pd = .99

Operating Characteristic

LEGEND
□ = FSS Test
○ = SPRT
△ = TST, C1=C2=.5

Signal to Noise Ratio
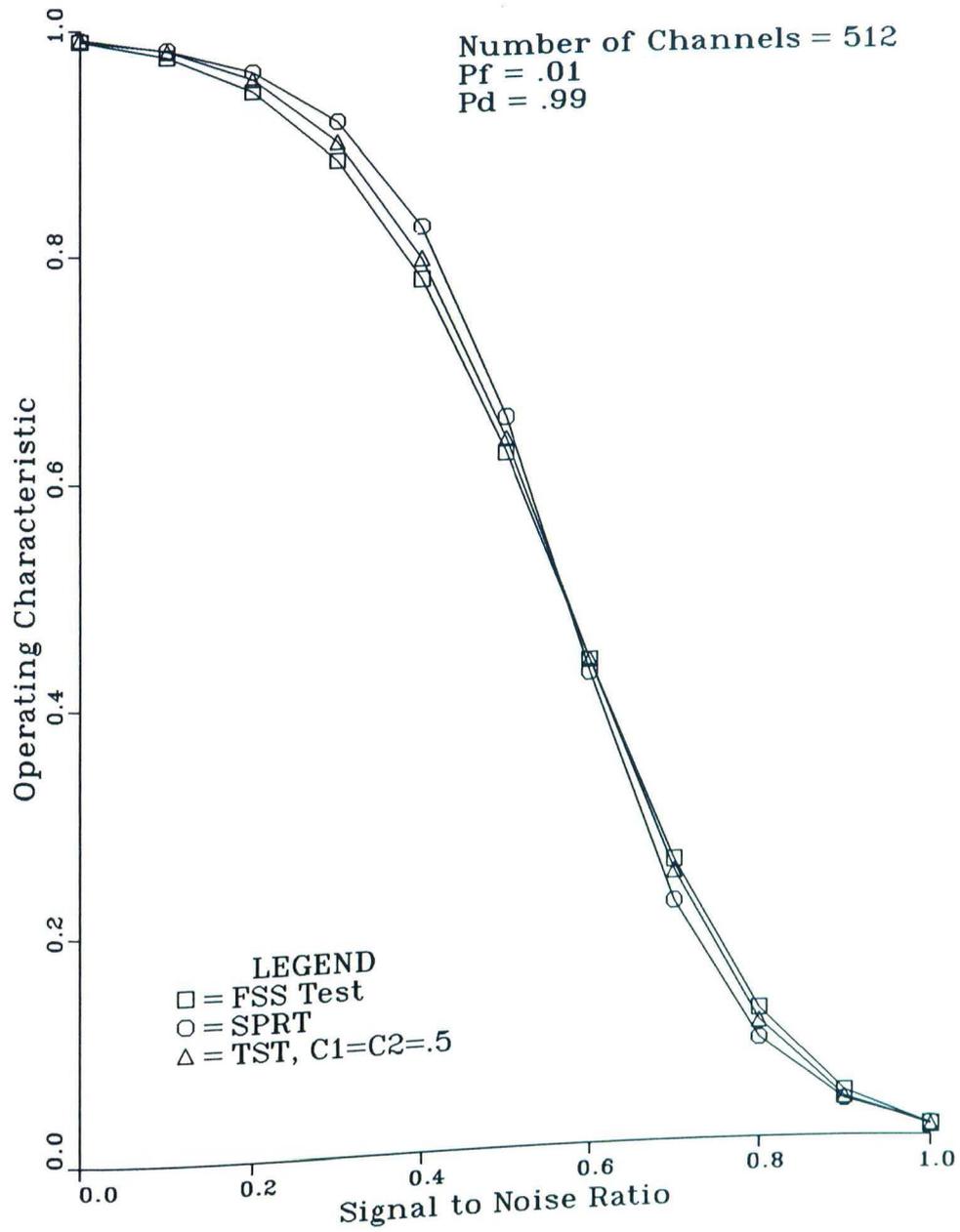
Figure 2.7: OC from Theory, $\gamma = 1$

Number of Channels = 512
Pf = .01
Pd = .99
Simulation Runs = 1000

LEGEND
□ = FSS Test
○ = SPRT
△ = TST, C1=C2=.5

Operating Characteristic

Signal to Noise Ratio
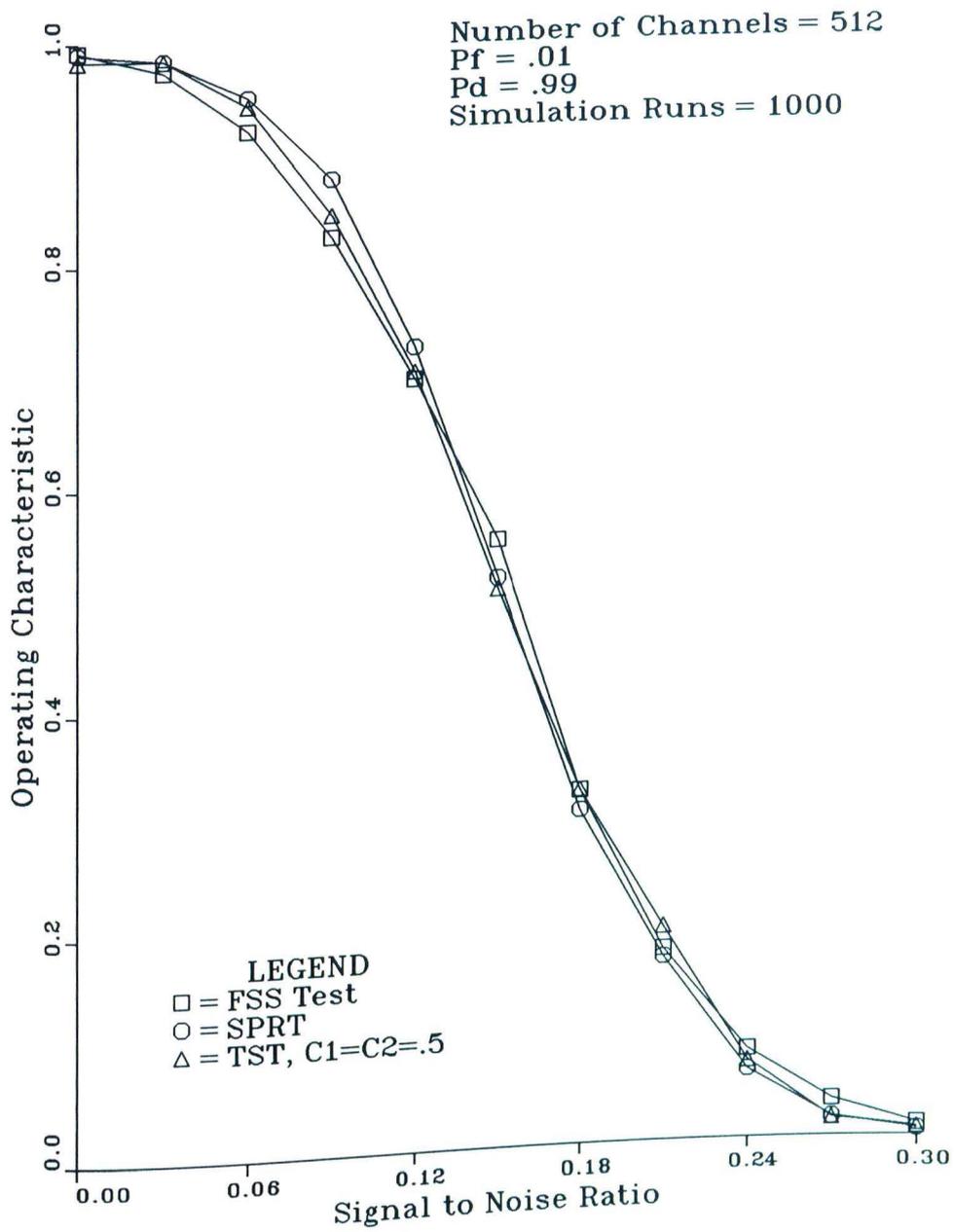
Figure 2.8: OC from Simulation, $\gamma = 0.3$

55
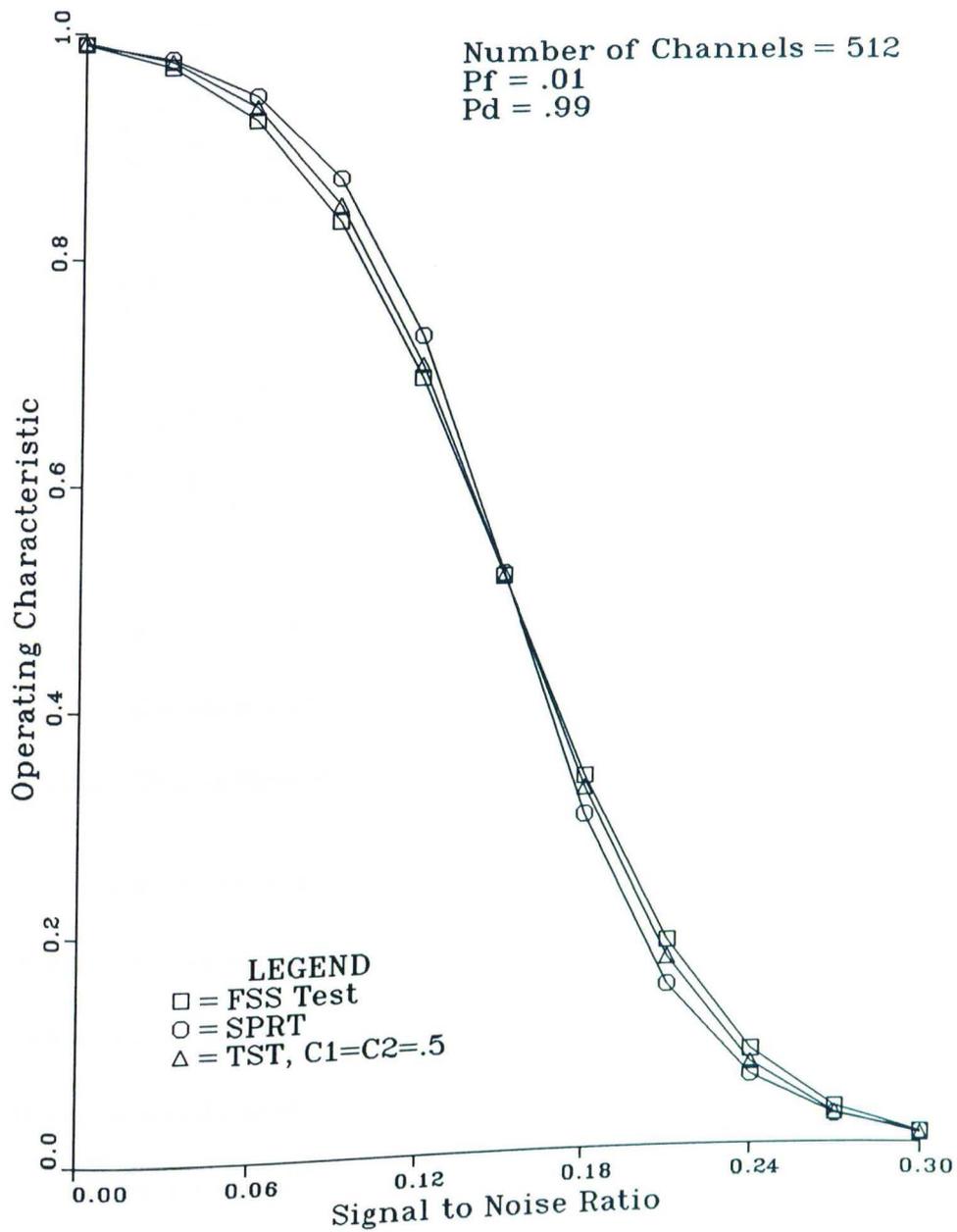
Figure 2.9: OC from Theory, $\gamma = 0.3$

Table 2.1: Comparison Between Theory and Simulation for $\gamma=1$

γ'	FSS test		SPRT		TST	
	Δ ASN	Δ OC	Δ ASN	Δ OC	Δ ASN	Δ OC
0.00	0.00	0.84	-1.19	-0.03	-1.47	0.03
0.10	0.00	1.11	-0.48	2.77	-0.76	2.03
0.20	0.00	0.45	-1.05	1.92	-1.84	1.02
0.30	0.00	0.18	1.16	-0.09	1.19	-0.18
0.40	0.00	-0.25	1.09	0.89	0.20	-0.40
0.50	0.00	0.73	0.60	0.83	-0.64	1.53
0.60	0.00	0.22	0.23	0.02	-0.14	-0.15
0.70	0.00	0.37	-1.47	1.64	-2.35	3.20
0.80	0.00	-2.16	-0.94	2.25	0.76	1.17
0.90	0.00	-1.33	-1.38	0.83	-1.88	1.02
1.00	0.00	-0.84	-1.16	3.50	-1.38	1.29

region. Notice that these test performances are reversed for SNRs close to γ. The OCs also show that the TST's actual detection errors, P_F and $1 - P_D$, are within 79% of their specified bounds, P_F^* and $1 - P_D^*$.

Throughout the analysis, various simplifying approximations were made whose accuracies were hard to quantify, especially the Wiener process approximations to the ALLF. Thus the computer simulation was compared quantitatively to results predicted by theory as a validation of assumptions made. Table 2.1 for $\gamma = 1$ and Table 2.2 for $\gamma = 0.3$ show how well the simulation of the three tests corresponds to the analysis. The quantity ΔASN is the normalized difference between the theoretical ASN and the simulation ASN, where the normalizing factor is the estimated standard deviation of the average used

Table 2.2: Comparison Between Theory and Simulation for $\gamma=0.3$

γ'	FSS test		SPRT		TST	
	Δ ASN	Δ OC	Δ ASN	Δ OC	Δ ASN	Δ OC
0.00	0.00	−0.71	−1.92	0.02	−1.04	2.09
0.03	0.00	−0.73	0.32	−1.75	1.38	−2.18
0.06	0.00	0.34	1.59	−0.82	−0.71	−1.03
0.09	0.00	0.68	−0.65	−0.49	0.07	0.28
0.12	0.00	−0.24	−1.72	0.33	−0.14	0.08
0.15	0.00	−2.17	−1.09	0.17	−0.48	0.69
0.18	0.00	0.91	−0.96	−0.19	−2.32	0.16
0.21	0.00	0.88	−1.04	−1.83	−2.49	−1.93
0.24	0.00	0.56	0.98	0.09	0.78	0.43
0.27	0.00	−0.18	1.79	2.02	1.70	2.68
0.30	0.00	−0.30	−0.70	1.67	0.63	0.20

to estimate the ASN. The ΔASN values show a good correspondence between theory and simulation, since they are within two standard deviations 86% of the time. The quantity ΔOC is the normalized difference between the theoretical OC and the simulation OC. Here the normalizing factor is the standard deviation of the OC average, assuming that the theoretical OC value is correct. In other words, the normalizing factor for a theoretical OC of $P_0(\gamma')$ and 1,000 simulation runs is $\sigma_{OC} = \sqrt{P_0(\gamma')[1 - P_0(\gamma')]/1000}$. Here again, in Tables 2.1 and 2.2, a good correspondence between theory and simulation is apparent.

The purpose of the computer simulation was to validate the assumptions made in the specification and analysis of the three tests: the FSS test, the SPRT, and the TST. The accuracy with which the analysis predicts quanti-

ties measured by simulation, as shown above, substantiates the assumptions made.

2.6   Test Extensions

The analytical expressions for the ASN and the OC of the TST, (2.66 and 2.70), can be used to determine a TST with an optimum mixture of FSS and SPRT parts. Specifically, the maximum ASN with respect to the SNR γ' varies as a function of the mixture constants C_1 and C_2. This function is graphed in Figure 2.10. The figure indicates that the optimal TST should have a greater mix of SPRT than the value one-half used in Section 2.5, since the maximum ASN of smallest value occurs for smaller values of the mixture constants C_1 and C_2. The optimal mixture constants were found numerically to be $C_1 = 0.286$ and $C_2 = 0.284$. The ASN and OC of the optimal TST are shown in Figures 2.11 and 2.12. It is interesting that, by minimizing the maximum ASN, the ASN in the extreme regions about $\gamma' = 0$ and $\gamma' = \gamma$ is also reduced. This is believed to be a consequence of the optimal TST having a greater SPRT mix than the half-and-half arbitrarily picked for the Section 2.5 simulation and, therefore, exhibiting properties closer to a pure SPRT. Of course, if the first TST was specified to have a larger SPRT mix, then optimization would have increased the ASN in the extreme regions. The optimal TST offers a

Number of Channels = 512
Pf = .01
Pd = .99
gamma = 1
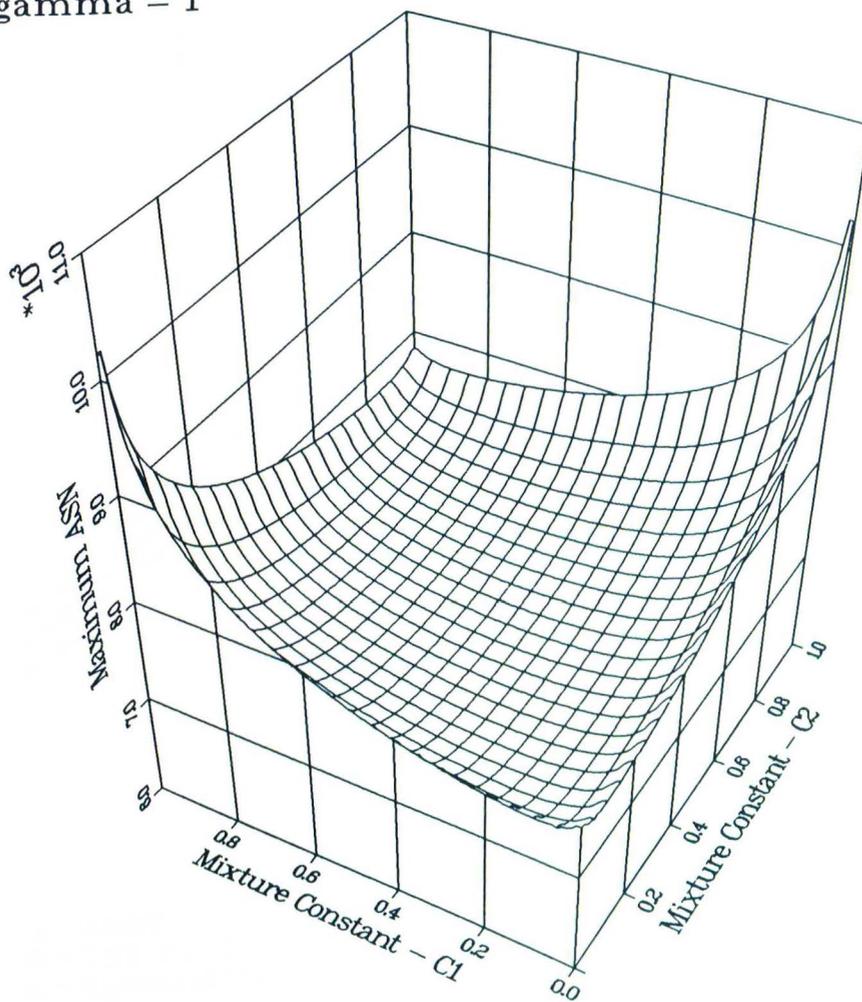
Figure 2.10: Maximum ASN versus Test Mixture Constants, C_1 and C_2
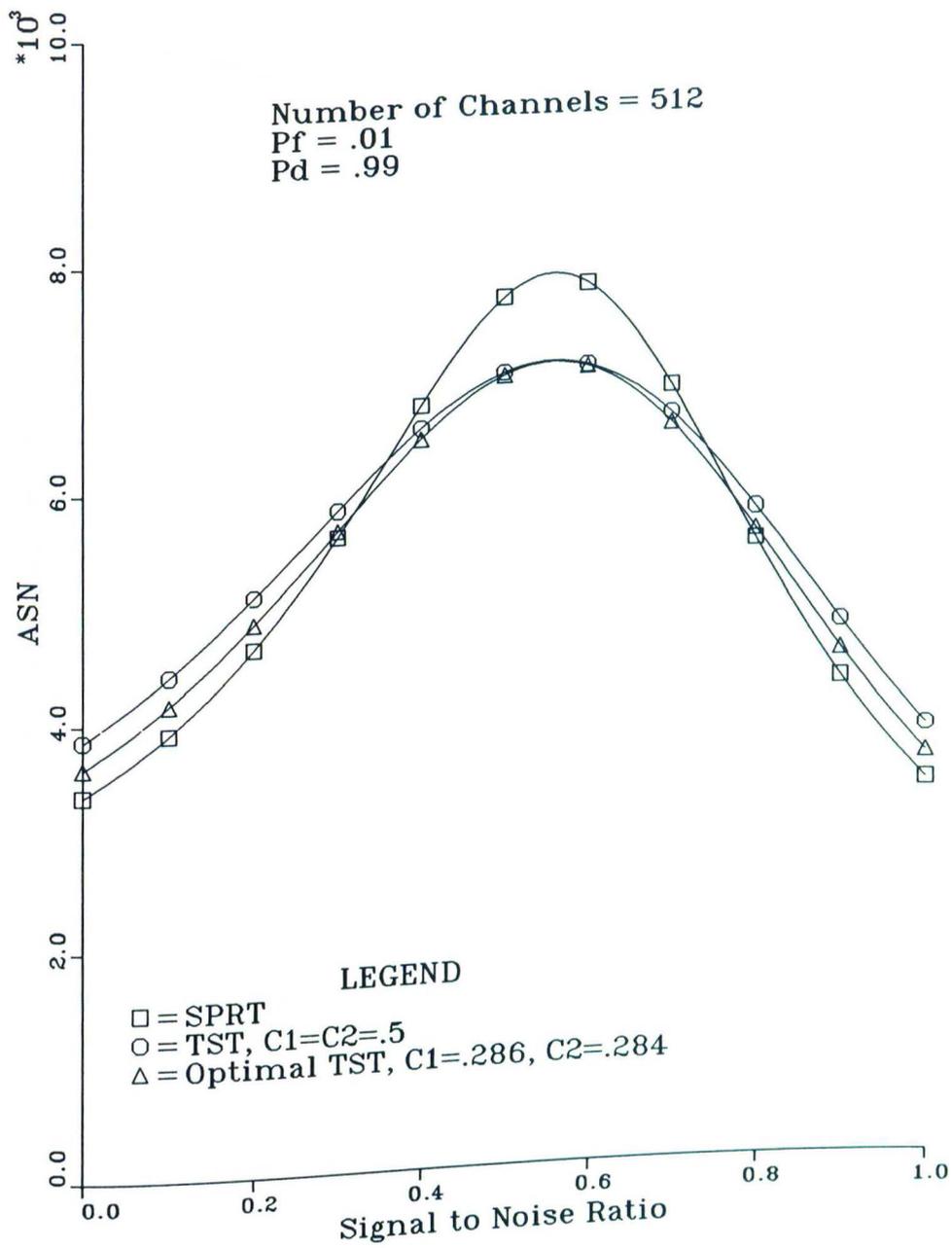
Figure 2.11: ASN of Optimal TST versus SPRT and half-mixed TST
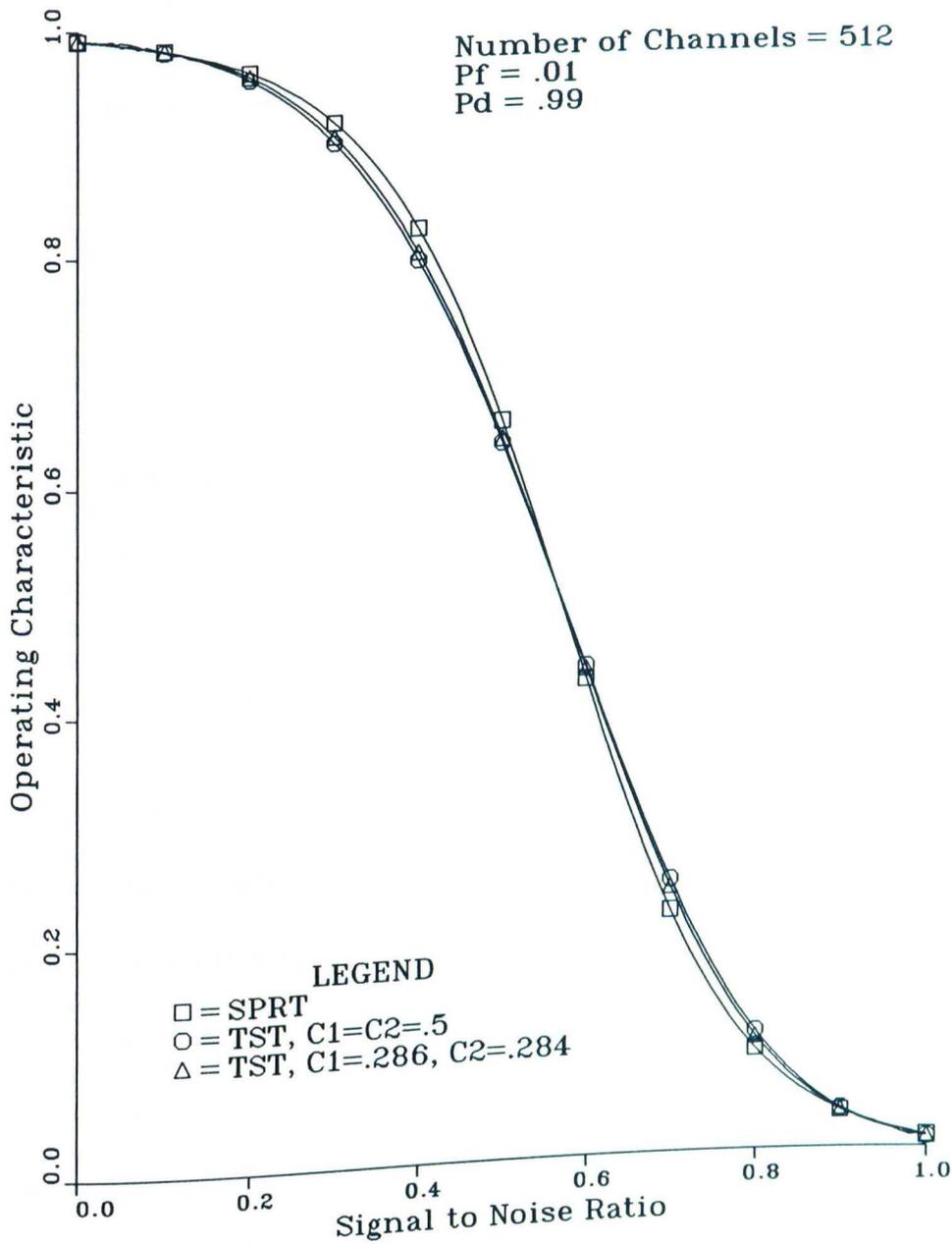
Figure 2.12: OC of Optimal TST versus SPRT and half-mixed TST

good compromise between the need to maximize the ASN performance in the extreme regions and the need to minimize the maximum ASN.

Another extension to the previously described tests is the robustification of the tests with respect to the assumed SNR γ. This can be accomplished by specifying the assumed SNR γ as the worst case, and then choosing a corresponding minimum probability of detection P_D^* that is somewhat relaxed. This procedure produces a test that adequately detects over a broader range of SNRs, and it is a way to effectively use the smaller detection times of the SPRT and TST. In this way, either a TST or an SPRT can be designed to adequately detect over a broader range of SNRs than an FSS test with the same or greater detection time.

The described tests also can be extended to the slow FH case. The detector structure itself is optimal under the fast FH assumption but is also a reasonable suboptimal structure for slow FH signals. This is especially true when there are a large number of hops over a given detection time. Even though the detector itself is suboptimal for slow FH, all the performance and design analysis developed for fast FH applies directly. This is because all design and performance analysis depended only on the chip duration being known and the interference being additive white Gaussian noise.

2.7   Asymptotic Efficiencies

The previous analysis did not include the performance of the tests when the assumed SNR γ is small. This case will be examined here. Since the ASN and OC are functions of both γ and the actual SNR γ', the ASN and OC can be recast as functions of γ and the relative SNR $r = \sqrt{\gamma'/\gamma}$. Test performance in the dwindling SNR case is captured by the limit of the ASN and OC, as γ diminishes while r is held constant. For the OC, this is a finite limit, but the ASN increases without bound. Thus, rather than comparing the ASNs directly, the limit of the ASN times γ^2 is computed. In other words, a quantity, identified as the asymptotic ASN $\tilde{E}(N/r)$, is defined as

$$\tilde{\mathcal{E}}(N/r) = \lim_{\gamma \to 0} \gamma^2 \mathcal{E}(N/r, \gamma). \tag{2.71}$$

The asymptotic ASN is useful because it preserves the relative efficiencies between the ASNs as γ diminishes. For instance, consider the FSS test ASN, $\mathcal{E}^{FSS}(N/r, \gamma)$, and the SPRT ASN, $E^{SPRT}(N/r, \gamma)$ and write

$$\lim_{\gamma \to 0} \frac{\mathcal{E}^{FSS}(N/r, \gamma)}{\mathcal{E}^{SPRT}(N/r, \gamma)} = \frac{\tilde{\mathcal{E}}^{FSS}(N/r)}{\tilde{\mathcal{E}}^{SPRT}(N/r)} \tag{2.72}$$

where $\tilde{E}^{FSS}(N/r)$ and $\tilde{E}^{SPRT}(N/r)$ are the asymptotic ASNs of the FSS test and SPRT, respectively. The asymptotic OC is simply defined as

$$\tilde{P}_0(r) = \lim_{\gamma \to 0} P_0(r, \gamma). \tag{2.73}$$

64

As an aid in evaluating these limits, asymptotic expressions for moments of the single-epoch ALLF derived in Appendix 2.B are defined as follows:

$$\mathcal{M}_r = \tilde{\mathcal{M}}_r \gamma^2 + \mathcal{O}(\gamma^3) \tag{2.74}$$

$$\mathcal{V}_r = \tilde{\mathcal{V}}_r \gamma^2 + \mathcal{O}(\gamma^3) \tag{2.75}$$

where

$$\tilde{\mathcal{M}}_r = \frac{K+2}{K^2}\left(r^2 - \frac{1}{2}\right) \tag{2.76}$$

$$\tilde{\mathcal{V}}_r = \frac{K+2}{K^2}. \tag{2.77}$$

Throughout this discussion, the quantity $\mathcal{O}(\gamma^n)$ represents any function, say $f(\gamma)$, such that

$$\gamma^{-n} \lim_{\gamma \to 0} f(\gamma) < \infty. \tag{2.78}$$

The particular function represented by $\mathcal{O}(\gamma^n)$ is determined from the context of the equation in which it appears.

To ease the expression of the asymptotic ASN and asymptotic OC, the variables $\tilde{L}$, $\tilde{\tau}$, $\tilde{a}$, and $\tilde{b}$ are defined. They will be labeled the asymptotic test parameters. Depending on the test type, they have expressions that correspond to that test type's parameter equations, where $\mathcal{M}_r$ is replaced with $\tilde{\mathcal{M}}_r$ and, likewise, $\mathcal{V}_r$ is replaced with $\tilde{\mathcal{V}}_r$. For instance, the FSS asymptotic test

parameters from (2.45) and (2.46) are

$$\tilde{L} \;=\; \frac{\left[\tilde{\mathcal{V}}_1^{\frac{1}{2}}\Phi^{-1}(1-P_D) - \tilde{\mathcal{V}}_0^{\frac{1}{2}}\Phi^{-1}(1-P_F)\right]^2}{\left(\tilde{\mathcal{M}}_1 - \tilde{\mathcal{M}}_0\right)^2} \tag{2.79}$$

$$\tilde{\tau} \;=\; \frac{\tilde{L}^{\frac{1}{2}}}{(\tilde{\mathcal{M}}_1 - \tilde{\mathcal{M}}_0)}\left[\tilde{\mathcal{V}}_0^{\frac{1}{2}}\tilde{\mathcal{M}}_1\Phi^{-1}(1-P_F) - \tilde{\mathcal{V}}_1^{\frac{1}{2}}\tilde{\mathcal{M}}_0\Phi^{-1}(1-P_D)\right]. \tag{2.80}$$

By using the asymptotic expressions (2.74) and (2.75), we have proved that the asymptotic ASN and OC of a particular test are exactly those of a test with the corresponding asymptotic test parameters. For example, this fact and (2.63) suggest that for the FSS test the asymptotic ASN is

$$\tilde{\mathcal{E}}^{FSS}(N/r) = \tilde{L} \tag{2.81}$$

while the FSS test's asymptotic OC is

$$\tilde{P}_0(r) = \Phi\left(\frac{\tilde{\tau} - \tilde{L}\tilde{\mathcal{M}}_r}{\sqrt{\tilde{L}\tilde{\mathcal{V}}_r}}\right). \tag{2.82}$$

The ASN and OC for the three different tests are plotted and compared in Figures 2.13 and 2.14. The relative relationship among the tests' asymptotic ASNs is almost exactly like that between the ASNs for $\gamma = 1$ and $\gamma = 0.3$, shown in Figures 2.2 through 2.5. This indicates that the three tests have reached their asymptotes, even for $\gamma = 1$. This comment also applies to the OCs. The usefulness of this asymptotic analysis, beyond verifying that the relationship among tests remains the same for diminishing SNR, is that it simplifies the test parameter relationships with respect to the parameters γ
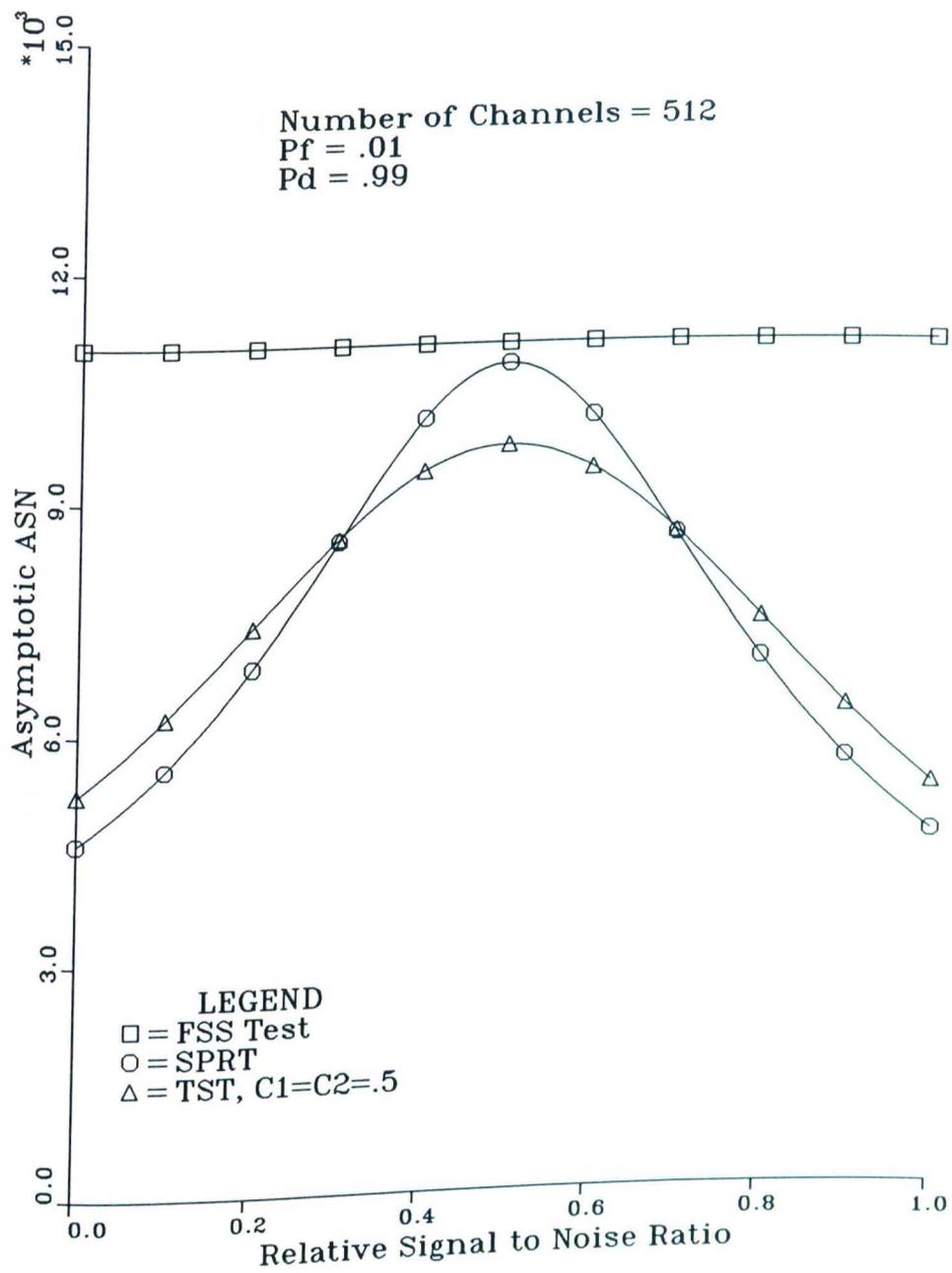
Figure 2.13: Asymptotic ASN of FSS Test, SPRT, and TST
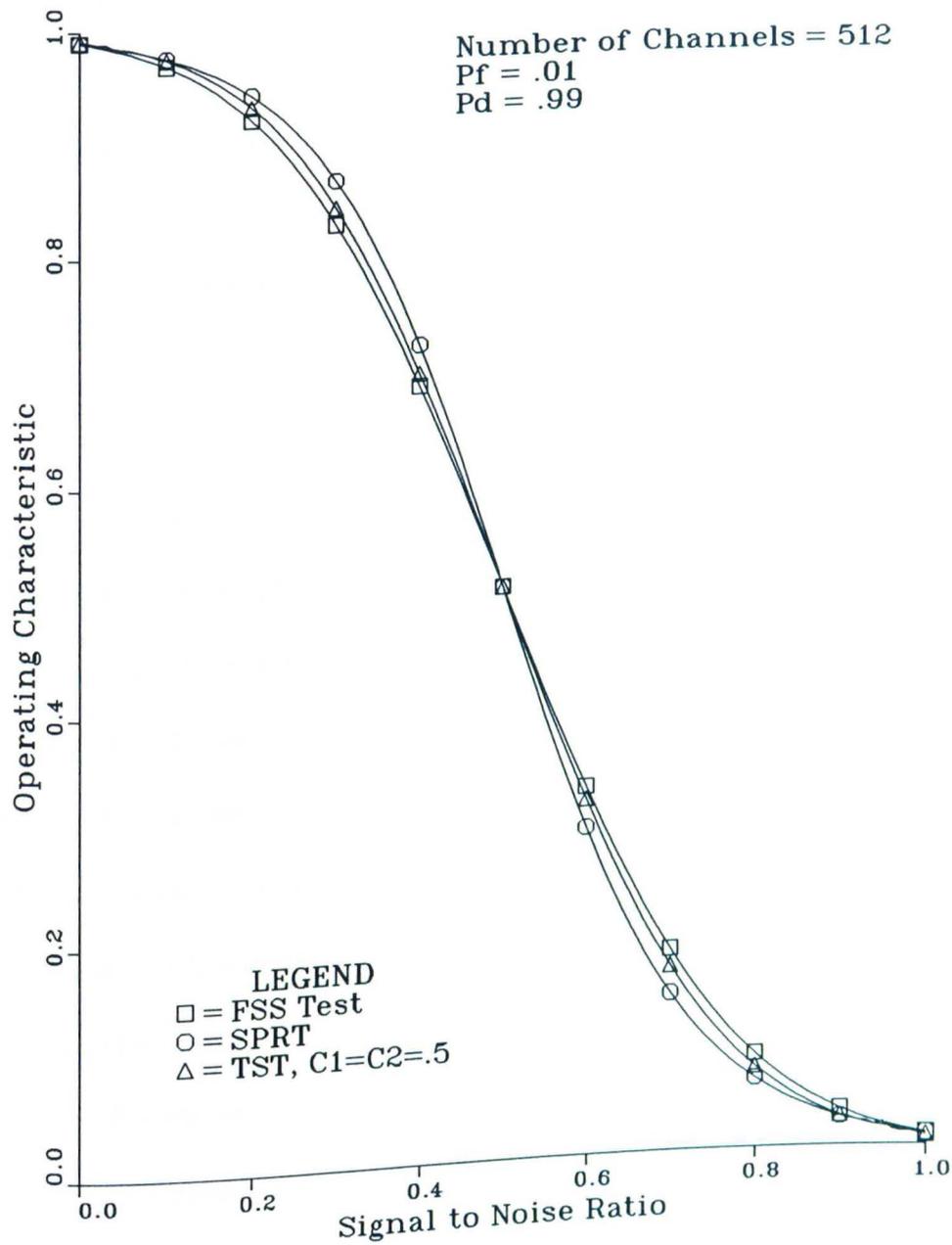
Figure 2.14: Asymptotic OC Function of FSS Test, SPRT, and TST

and r. Thus for each test, we could choose parameters $L = \tilde{L}\gamma^{-2}$, $\tau = \tilde{\tau}$, $a = \tilde{a}$, and $b = \tilde{b}$ and have comparable performance for any small γ. This feature simplifies any adaptation with respect to γ that might be added to this detection scheme.

2.8   Conclusions

Methods for the sequential detection of noncoherent fast FH waveforms have been developed. In the process, the FH waveform is modeled to have an information component that consisted of a series of chips with a known constant epoch, where each chip frequency is one of a known ensemble of frequencies. In the model, a particular chip frequency is independently determined by a uniform random variable on the frequency ensemble. The FH waveform is assumed to have an additive white-noise component. By assuming the modeled FH waveform to be a known SNR, the optimal detector based on a single-epoch observation (SELF) is developed using likelihood-function theory. SELF is the sum of many nearly identical and nearly independent random variables and thus has nearly Gaussian statistics. This central-limit argument allows a multi-epoch collection of SELFs to be considered an equivalent set of Gaussian i.i.d. variables. From these simplified observations, a log-likelihood function (ALLF) is computed that is asymptotic to the exact log-likelihood

function as the number of possible hop frequencies becomes large. The ALLF becomes the test statistic on which three detection tests are based. The tests are the FSS test, SPRT, and TST. These are defined to ensure that detection errors are below desired levels. By modeling the ALLF as a Wiener process, diffusion theory yields the performance of the three tests not only for an FH waveform of the assumed SNR but also for all SNRs below the one assumed. This analysis compares favorably with a computer simulation of the detector that validates the analysis. The analysis also becomes a tool used to numerically optimize performance of the TST, when the actual FH SNR deviates from that assumed. In order to study the performance of tests synthesized by assuming an extremely small FH SNR, expressions for the asymptotic test efficiencies are computed. This asymptotic analysis also yields simplified test parameter expressions applicable to the small SNR case.

A significant feature of the SPRT exposed by the analysis is that, with the same error probabilities, an FH waveform with a given SNR can be detected in less than half the time of the corresponding FSS test. This reduction in detection time is especially significant for Low Probability of Intercept (LPI) applications in which the transmissions are purposely short. For the pure SPRT, detection time increases whenever the observed SNR differs from that assumed in the test's synthesis. For SNRs midway between zero and the assumed value, it is even comparable to the corresponding FSS test. The

TST significantly improves this anomaly while sacrificing little performance over that of the purely sequential test, and what little performance is lost is largely regained by the optimal TST. The decrease in the detection time of the sequential tests can be used to robustify the test with respect to the input SNR, while maintaining better performance than that of the non-robust FSS test. The simplified test parameter expressions derived by asymptotic methods may be useful for any schemes to adapt these tests for varying FH SNRs. The three tests and their corresponding design and performance analysis also apply to the slow FH case. The detector structure is suboptimal for slow FH but it is believed that the performance loss is small, especially for detection times that include a large number of hops.

It is apparent that other simplifications and extensions to these results are possible. For instance, it is assumed that the starting time and duration of the chip epoch are known. This first restriction might be relaxed by redefining the SELF to perform sliding window integration instead of the integrate-and-dump operation now performed. This, of course, would degrade the detector's performance for some values of epoch starting time, but probably would exhibit a better average performance. There are also possible simplifications to the SELF to improve its implementability; for example, removal of the emphasizing function, would make the detector structure suboptimal but probably still asymptotically optimal for small assumed SNRs. Another simplification could

be coarse subband preselection, where the total spread-spectrum bandwidth is subdivided into subbands, each containing a large number of chip frequencies. An algorithm then could be used to select a subset of the subbands that most likely would contain the intercepted signal. Detailed processing on these preselected bands then could be done with the methods described in this chapter.

2.A   Derivation of SELF

Proceeding from Appendix B of [50], the likelihood function, given the carrier phase θ and the channel k, the conditional likelihood function for the ith epoch is

$$\Lambda_i\left(y/k,\theta\right) = e^{-\frac{E}{N_0}} e^{\frac{2}{N_0} \int_{iT_h}^{(i+1)T_h} y(t)x_i(t)dt} \tag{2.83}$$

where E is the single-epoch energy of the FH signal, i.e.,

$$E = \int_{iT_h}^{(i+1)T_h} 2S\sin^2(\omega_{k_i}t+\theta)dt. \tag{2.84}$$

But

$$E \approx ST_h \qquad \text{for } \omega_{k_i}T_h \gg 1 \tag{2.85}$$

which, upon substitution into the conditional likelihood function (2.83) and expanding $x_i(t)$, yields

$$\Lambda_i\left(y/k,\theta\right) = e^{-\gamma} e^{\sqrt{2\gamma}(P_k\sin\theta + Q_k\cos\theta)} \tag{2.86}$$

72

where

$$P_k = \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} y(t) \cos \omega_{k_i} t \, dt \qquad (2.87)$$

$$Q_k = \frac{2}{\sqrt{N_0 T_h}} \int_{iT_h}^{(i+1)T_h} y(t) \sin \omega_{k_i} t \, dt.$$

Taking Expectations with respect to θ defines

$$\Lambda_i(y/k) \triangleq \mathcal{E}_\theta \left[\Lambda_i(y/k, \theta) \right] \qquad (2.88)$$

which is the likelihood function conditional only on the channel. This expectation can be evaluated as

$$\mathcal{E}_\theta \left[\Lambda_i(y/k, \theta) \right] = \frac{e^{-\gamma}}{2\pi} \int_0^{2\pi} e^{\sqrt{2\gamma}(P_k \sin \theta + Q_k \cos \theta)} d\theta \qquad (2.89)$$

$$= \frac{e^{-\gamma}}{2\pi} \int_0^{2\pi} e^{\sqrt{2\gamma}\sqrt{P_k^2 + Q_k^2} \sin(\theta + \psi)} d\theta \qquad (2.90)$$

where $\psi = \mathrm{Arg}(P_k + jQ_k)$. Now by the periodicity of the integrand

$$\mathcal{E}_\theta \left[\Lambda_i(y/k, \theta) \right] = \frac{e^{-\gamma}}{2\pi} \int_0^{2\pi} e^{\sqrt{2\gamma}\sqrt{P_k^2 + Q_k^2} \cos \theta} d\theta \qquad (2.91)$$

$$= e^{-\gamma} I_0 \left(\sqrt{2\gamma}\sqrt{P_k^2 + Q_k^2} \right) \qquad (2.92)$$

by the identity

$$I_0(a) = \frac{1}{2\pi} \int_0^{2\pi} e^{a \cos \theta} d\theta \qquad (2.93)$$

where I_0 is the zeroth-order modified Bessel function of the first kind. Taking expectations with respect to the channel k yields the single-epoch likelihood function (SELF)

$$\Lambda_i(y) \triangleq \mathcal{E}_k \left[\Lambda_i(y/k) \right] \qquad (2.94)$$

$$= \frac{e^{-\gamma}}{K} \sum_{k=0}^{K-1} I_0 \left(\sqrt{2\gamma}\sqrt{P_k^2 + Q_k^2} \right). \qquad (2.95)$$

73

2.B   Asymptotic Expressions of the ALLF Moments

We want to examine the behavior of the ALLF moments when SNR γ diminishes while the relative SNR r is held constant. The asymptotic expressions derived here encapsulate this behavior. To derive asymptotic expressions for the mean and variance of the ALLF, consider the mean and variance $\mathcal{M}_r$ and $\mathcal{V}_r$ of the single-epoch ALLF only. To this end, it will be useful to derive asymptotic expression of two functions of the channel moments: μ_r/σ_0 and σ_r^2/σ_0^2. Starting with the first expression and substituting (2.21) and (2.22) forms

$$\frac{\mu_r^2}{\sigma_0^2} = \frac{I_0^2(2r\gamma)}{[I_0(2\gamma) - 1]}. \tag{2.96}$$

We will need a partial power series expansion of $I_0(x)$, i.e.,

$$I_0(x) = 1 + \frac{x^2}{4} + \frac{x^4}{64} + \mathcal{O}(x^6). \tag{2.97}$$

Here and throughout this discussion, the quantity $\mathcal{O}(x^n)$ represents any function, say $f(x)$, such that

$$x^{-n} \lim_{x \to 0} f(x) < \infty. \tag{2.98}$$

The particular function represented by $\mathcal{O}(x^n)$ is determined from the context of the equation in which it appears. With the above power series for I_0, (2.96)

becomes

$$\frac{\mu_r^2}{\sigma_0^2} = \frac{[1 + r^2\gamma^2 + \mathcal{O}(\gamma^4)]^2}{\gamma^2 + \frac{1}{4}\gamma^4 + \mathcal{O}(\gamma^6)} \tag{2.99}$$

$$= \frac{1 + 2r^2\gamma^2 + \mathcal{O}(\gamma^4)}{\gamma^2 + \frac{1}{4}\gamma^4 + \mathcal{O}(\gamma^6)} \tag{2.100}$$

$$= \gamma^{-2} + \left(2r^2 - \tfrac{1}{4}\right) + \mathcal{O}(\gamma^2) \tag{2.101}$$

$$= \left[\gamma^{-1} + \left(r^2 - \tfrac{1}{8}\right)\gamma + \mathcal{O}(\gamma^3)\right]^2 \tag{2.102}$$

thus

$$\frac{\mu_r}{\sigma_0} = \gamma^{-1} + \left(r^2 - \tfrac{1}{8}\right)\gamma + \mathcal{O}(\gamma^3). \tag{2.103}$$

Now let us evaluate the second channel moment function. Using (2.21) and (2.97) plus the power series for e^x, after carrying out the multiplications we get

$$\frac{\sigma_r^2}{\sigma_0^2} = [I_0(2\gamma) - 1]^{-1}\left[\frac{1}{\pi}\int_0^\pi e^{-2\gamma\cos\phi}I_0(4r\gamma\sin\tfrac{\phi}{2})d\phi - I_0^2(2r\gamma)\right] \tag{2.104}$$

$$\frac{\sigma_r^2}{\sigma_0^2} = \left[\gamma^2 + \tfrac{1}{4}\gamma^4 + \mathcal{O}(\gamma^6)\right]^{-1}\left\{\frac{1}{\pi}\int_0^\pi\right.$$
$$\left[1 - 2\gamma\cos\phi + 2\gamma^2\cos^2\phi - \tfrac{4}{3}\gamma\cos^3\phi^3 + \tfrac{2}{3}\cos^4\phi\gamma^4 + \mathcal{O}(\gamma^5)\right] \times$$
$$\left[1 + 4r^2\gamma^2\sin^2\tfrac{\phi}{2} + 4r^4\gamma^4\sin^4\tfrac{\phi}{2} + \mathcal{O}(\gamma^6)\right]d\phi \tag{2.105}$$
$$\left. - \left[1 + r^2\gamma^2 + \tfrac{1}{4}r^4\gamma^4 + \mathcal{O}(\gamma^6)\right]^2\right\}$$

75

$$\frac{\sigma_r^2}{\sigma_0^2} = \left[\gamma^2 + \tfrac{1}{4}\gamma^4 + \mathcal{O}(\gamma^6)\right]^{-1} \left\{ \frac{1}{\pi}\int_0^\pi \left[1 - 2\cos\phi\,\gamma \right.\right.$$

$$+ \left(4r^2\sin^2\tfrac{\phi}{2} + 2\cos^2\phi\right)\gamma^2 - \left(8r^2\cos\phi\sin^2\tfrac{\phi}{2} + \tfrac{4}{3}\cos^3\phi\right)\gamma^3$$

$$+ \left(\tfrac{2}{3}\cos^4\phi + 8r^2\cos^2\phi\sin^2\tfrac{\phi}{2} + 4r^4\sin^4\tfrac{\phi}{2}\right)\gamma^4 + \mathcal{O}(\phi,\gamma^5) \Big]\, d\phi$$

$$\left. - \left[1 + 2r^2\gamma^2 + \tfrac{3}{2}r^4\gamma^4 + \mathcal{O}(\gamma^6)\right] \right\}. \tag{2.106}$$

Let $f(\phi,\gamma^5)$ be the particular function represented by the symbol $\mathcal{O}(\phi,\gamma^5)$ under the integral, then

$$\lim_{\gamma\to 0} \gamma^{-5}\int_0^\pi f(\phi,\gamma^5) = \int_0^\pi \lim_{\gamma\to 0} \gamma^{-5} f(\phi,\gamma^5) < \infty \tag{2.107}$$

implying that $\int_0^\pi f(\phi,\gamma^5) \in \mathcal{O}(\gamma^5)$. The interchange of the limit and integration is justified as follows. The function $f(\phi,\gamma^5)$ inherits continuity on the compact set $\{\phi,\gamma : \phi \in [0,\pi] \text{ and } \gamma \in [0,1]\}$ from the integrand. Therefore, $\gamma^{-5}f(\phi,\gamma^5)$, which has a finite limit at the origin, is also continuous on this compact set and hence is bounded, say by B, on this set. The function B is integrable on $\phi \in [0,\pi]$, from which the interchange follows by the Lebesgue Dominated Convergence theorem. The interchange implies that (2.106) can be integrated term wise to yield

$$\frac{\sigma_r^2}{\sigma_0^2} = \left[\gamma^2 + \tfrac{1}{4}\gamma^4 + \mathcal{O}(\gamma^6)\right]^{-1} \times$$

$$\left\{ \left[1 + (2r^2+1)\gamma^2 + 2r^2\gamma^3 + (\tfrac{1}{4} + 2r^2 + \tfrac{3}{2}r^4)\gamma^4 + \mathcal{O}(\gamma^5)\right] \right. \tag{2.108}$$

$$\left. - \left[1 + 2r^2\gamma^2 + \tfrac{3}{2}r^4\gamma^4 + \mathcal{O}(\gamma^6)\right] \right\}$$

which simplifies to

$$\frac{\sigma_r^2}{\sigma_0^2} = 1 + 2r^2\gamma + 2r^2\gamma^2 + \mathcal{O}(\gamma^3). \tag{2.109}$$

With these asymptotic expressions for μ_r/σ_0 and σ_r^2/σ_0^2, we proceed with the derivation of asymptotic expressions for the ALLF moments. The ALLF mean is expressed in terms of the SELF moments as

$$\mathcal{M}_r = \frac{1}{2} \ln \frac{V_0}{V_1} + [2V_1 V_0]^{-1} \left\{ [M_r - M_0]^2 V_1 - [M_r - M_1]^2 V_0 + [V_1 - V_0] V_r \right\}.$$

(2.110)

The last three terms can be evaluated as follows:

$$[2V_1 V_0]^{-1} \left\{ [M_r - M_0]^2 V_1 - [M_r - M_1]^2 V_0 + [V_1 - V_0] V_r \right\}$$

$$= \frac{1}{2K} \left[K - 1 + \frac{\sigma_1^2}{\sigma_0^2} \right]^{-1} \left\{ \left[\frac{\mu_r}{\sigma_0} - \frac{\mu_0}{\sigma_0} \right]^2 \left[K - 1 + \frac{\sigma_1^2}{\sigma_0^2} \right] \right.$$

$$\left. -K \left[\frac{\mu_r}{\sigma_0} - \frac{\mu_1}{\sigma_0} \right]^2 + \left[K - 1 + \frac{\sigma_r^2}{\sigma_0^2} \right] \left[\frac{\sigma_1^2}{\sigma_0^2} - 1 \right] \right\}$$

(2.111)

$$= \frac{1}{2K} [K + 2\gamma + \mathcal{O}(\gamma^2)]^{-1} \left\{ [r\gamma + \mathcal{O}(\gamma^3)]^2 [K + \mathcal{O}(\gamma)] \right.$$

$$-K \left[\left(r^2 - 1 \right) \gamma + \mathcal{O}(\gamma^3) \right]^2$$

$$\left. + \left[K + 2r^2\gamma + \mathcal{O}(\gamma^2) \right] \left[2\gamma + 2\gamma^2 + \mathcal{O}(\gamma^3) \right] \right\}$$

(2.112)

$$= \frac{\gamma}{K} + \left(\frac{K + 2}{K^2} r^2 + \frac{K - 4}{2K^2} \right) \gamma^2 + \mathcal{O}(\gamma^3).$$

(2.113)

Now for the final term. We will need the following power series expansion for $\ln(x)$

$$\ln(1 + x) = x - \frac{x^2}{2} + \mathcal{O}(x^3)$$

(2.114)

$$\frac{1}{2} \ln\left(\frac{V_0}{V_1}\right) = -\frac{1}{2} \ln\left[1 - \frac{1}{K} + \frac{1}{K}\frac{\sigma_1^2}{\sigma_0^2}\right] \quad (2.115)$$

$$= -\frac{1}{2} \ln\left[1 + \frac{2}{K}\gamma + \frac{2}{K}\gamma^2 + \mathcal{O}(\gamma^3)\right] \quad (2.116)$$

$$= -\frac{\gamma}{K} + \left(\frac{1}{K^2} - \frac{1}{K}\right)\gamma^2 + \mathcal{O}(\gamma^3) \quad (2.117)$$

combining (2.113) and (2.117)

$$\mathcal{M}_r = \frac{K+2}{K^2}\left(r^2 - \frac{1}{2}\right)\gamma^2 + \mathcal{O}(\gamma^3). \quad (2.118)$$

Now to proceed with the variance

$$\mathcal{V}_r = \frac{V_r^2}{2}\left(\frac{1}{V_0} - \frac{1}{V_1}\right)^2 + \left[\left(\frac{1}{V_0} - \frac{1}{V_1}\right)M_r + \left(\frac{M_1}{V_1} - \frac{M_0}{V_0}\right)\right]^2 V_r. \quad (2.119)$$

The first term can be evaluated as

$$\frac{V_r^2}{2}\left(\frac{1}{V_0} - \frac{1}{V_1}\right)^2 = \frac{1}{2K^2}\left\{\frac{\left[\frac{\sigma_1^2}{\sigma_0^2} - 1\right]\left[K - 1 + \frac{\sigma_r^2}{\sigma_0^2}\right]}{\left[K - 1 + \frac{\sigma_1^2}{\sigma_0^2}\right]}\right\}^2 \quad (2.120)$$

$$= \frac{1}{2K^2}\left\{\frac{[2\gamma + \mathcal{O}(\gamma^2)][K + \mathcal{O}(\gamma)]}{[K + \mathcal{O}(\gamma)]}\right\}^2 \quad (2.121)$$

$$= \frac{2\gamma^2}{K^2} + \mathcal{O}(\gamma^3). \quad (2.122)$$

The second term of (2.119)

$$\left[\left(\frac{1}{V_0} - \frac{1}{V_1}\right)M_r + \left(\frac{M_1}{V_1} - \frac{M_0}{V_0}\right)\right]^2 V_r$$

$$= \frac{1}{K^2} \frac{\left[K - 1 + \frac{\sigma_r^2}{\sigma_0^2}\right]}{\left[K - 1 + \frac{\sigma_1^2}{\sigma_0^2}\right]^2} \left\{\left[\frac{\sigma_1^2}{\sigma_0^2} - 1\right]\left[(K-1)\frac{\mu_0}{\sigma_0} + \frac{\mu_r}{\sigma_0}\right] + K\left[\frac{\mu_1}{\sigma_0} - \frac{\mu_0}{\sigma_0}\frac{\sigma_1^2}{\sigma_0^2}\right]\right\}^2$$

$$\tag{2.123}$$

$$\tag{2.124}$$
$$= \frac{1}{K^2}\frac{[K + \mathcal{O}(\gamma)]}{[K + \mathcal{O}(\gamma)]^2}\left\{\left[2\gamma + 2\gamma^2 + \mathcal{O}(\gamma^3)\right]\left[K\gamma^{-1} + \mathcal{O}(\gamma)\right]\right.$$
$$+ K\left[\gamma^{-1} + \tfrac{7}{8}\gamma + \mathcal{O}(\gamma^3)\right]$$
$$\left. - K\left[\gamma^{-1} - \tfrac{1}{8}\gamma + \mathcal{O}(\gamma^3)\right]\left[1 + 2\gamma + 2\gamma^2 + \mathcal{O}(\gamma^3)\right]\right\}^2$$

$$= \frac{1}{K^2}\frac{[K + \mathcal{O}(\gamma)]}{[K + \mathcal{O}(\gamma)]^2}\left\{\left[2K + 2K\gamma + \mathcal{O}(\gamma^2)\right] + K\left[-2 - \gamma + \mathcal{O}(\gamma^2)\right]\right\}^2 \tag{2.125}$$

$$\tag{2.126}$$
$$= \frac{\gamma^2}{K} + \mathcal{O}(\gamma^3).$$

Combining (2.122) and (2.126) yields

$$\mathcal{V}_r = \frac{K+2}{K^2}\gamma^2 + \mathcal{O}(\gamma^3). \tag{2.127}$$

Summarizing these results

$$\mathcal{M}_r = \frac{K+2}{K^2}\left(r^2 - \frac{1}{2}\right)\gamma^2 + \mathcal{O}(\gamma^3) \tag{2.128}$$

$$\mathcal{V}_r = \frac{K+2}{K^2}\gamma^2 + \mathcal{O}(\gamma^3). \tag{2.129}$$

Chapter 3

The Optimal Interception of Frequency-Hopped Waveforms via a Compressive Receiver

3.1   Introduction

The goal of the intercept receiver is to detect deceptive electromagnetic sources and follow up by extracting waveform features for use in the jamming or exploitation of that source. With the advent of frequency-hopped and other spread-spectrum signals, the search bandwidth that assures a reasonable probability of intercept is increased greatly with a corresponding increase in the complexity of the intercept problem. For these cases, wideband interceptors prove unacceptable, while high-performing channelized interceptors prove virtually unimplementable. The compressive receiver, which simultaneously estimates frequency components over a wide, set band, has promise as an interceptor with both the simplicity of a wideband device and the performance

of the channelized device.

The literature is rich with intercept methods for frequency-hopped waveforms (see [40], [45], [32], [41], and [39]). There are also some analyses of the detection performance of the compressive receiver (see [4] and [14]). However, very little has been written on the application of the compressive receiver to the interception of spread-spectrum signals ([20] is an exception) and even less on the interception of frequency-hopped waveforms. To fill the void, this chapter fully develops an optimal and a simplified suboptimal method for the detection of frequency-hopped waveforms. The chapter further exhibits a structure for hop frequency estimation.

The chapter models the compressive-receiver input as consisting of either stationary Gaussian noise of known autocorrelation or of noise plus a hopped signal of known hop epoch, unknown phase, and energy above a minimum detectable level. Approximate transfer relationships for signal and noise are developed separately and used to translate the detection problem to an equivalent one on the compressive-receiver output. Likelihood function theory is applied to the equivalent problem and yields a locally optimal (i.e. optimal for small signal-to-noise ratio) detector. The locally optimal detector has an unwieldy structure that defeats the motivation of using a compressive receiver: that of simplicity plus high performance. Therefore, a time-multiplexed detector is developed that, at the expense of duty cycle, can achieve performance

as close to optimal as desired. Asymptotic statistics of the detector's output are derived and used to quantify performance. A hop frequency estimator is presented and its probability of track estimated.

3.2 Preliminaries

Several areas need elucidation before a useful interceptor can be developed, in particular, a precise statistical model of the compressive-receiver input and a precise model of the compressive-receiver itself. In order to statistically model the compressive receiver output, transfer relationships are needed for both the noise and signal.

3.2.1 Input Signal Model

The signal model is for a composite hypothesis problem. Specifically, given the observation $y(t)$, the problem is one of choosing between H_0, which is the hypothesis that an FH waveform is not present, and $H_{\gamma'}$, which is the hypothesis that an FH waveform is present with an SNR γ' greater than some minimum SNR γ. Exactly, the model is

$$\begin{aligned} H_0 &: \quad y(t) &=& \quad n(t) \\ \text{versus} \quad H_{\gamma'} &: \quad y(t) &=& \quad s(t) + n(t) \qquad \gamma < \gamma' \end{aligned} \qquad (3.1)$$

where the frequency-hopped signal $s(t)$ is given by (1.1) and $n(t)$ is stationary, colored Gaussian noise with variance σ_i^2 and with autocorrelation function
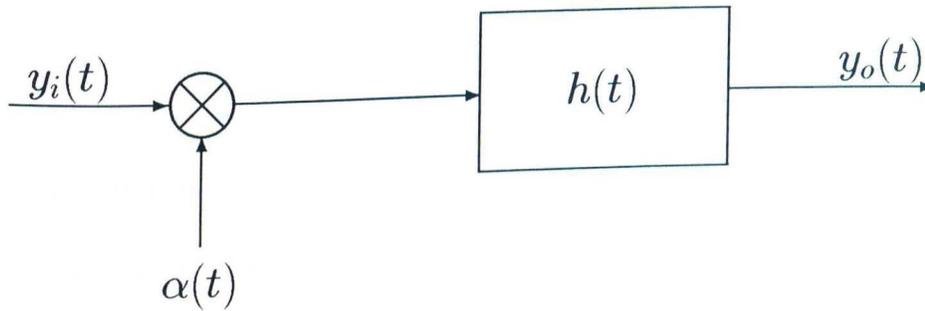
Figure 3.1: Block Diagram of Compressive Receiver Model

$\sigma_i^2 R_i(t)$. The hypothesized SNR γ' is related to the other model parameters by $\gamma' = S'T_h/\sigma_i^2$, while similarly the minimum SNR $\gamma = ST_h/\sigma_i^2$.

Significantly, the signal model allows for colored noise and is, therefore, quite general. Note that the model assumes that all signal parameters except amplitude and hop frequency are known.

3.2.2 Receiver Model

Figure 3.1 blocks out the compressive-receiver model. The compressive receiver mixes the input signal $y_i(t)$ with a linearly frequency-modulated signal

$$\alpha(t) = \cos(\omega_0 t - \beta t^2) \qquad 0 \leq t \leq T_s \qquad (3.2)$$

that scans downward in frequency from ω_0 to $\omega_0 - 2\beta T_s$. Here T_s is the scan time. The scanned waveform is input to a pulse compression filter, hence, the

name compressive receiver. The filter has impulse response

$$h(t) = \cos(\omega_0 t + \beta t^2)\, w(t) \qquad 0 \le t \le T_c \tag{3.3}$$

where $w(t)$ is a weighting function used to minimize energy spillover between signals of different frequencies. The output of the compressive receiver now follows as

$$y_o(t) \;=\; \int_0^{T_c} \alpha(t - \tau) y_i(t - \tau) h(\tau)\, d\tau \qquad T_c \le t \le T_s \tag{3.4}$$

$$\;=\; x_o(t) + n_o(t) \tag{3.5}$$

where

$$x_o(t) \;\triangleq\; \int_0^{T_c} \alpha(t - \tau) x_i(t - \tau) h(\tau)\, d\tau \qquad T_c \le t \le T_s \tag{3.6}$$

$$n_o(t) \;\triangleq\; \int_0^{T_c} \alpha(t - \tau) n_i(t - \tau) h(\tau)\, d\tau \qquad T_c \le t \le T_s. \tag{3.7}$$

3.2.3 Output due to Signal

Using (3.2), (3.3), and the commuted version of (3.7), the output of the compressive receiver can be expressed as

$$x_o(t) = \int_{t-T_c}^{t} x(\tau) \cos(\omega_0 \tau - \beta\tau^2) \cos\left[\omega_0(t - \tau) + \beta(t - \tau)^2\right] w(t - \tau)\, d\tau \tag{3.8}$$

whenever $T_c \le t \le T_s$. Trigonometric manipulation leads to

$$x_o(t) \;=\; \frac{1}{2}\cos(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} x(\tau) \cos(2\beta t\tau) w(t - \tau)\, d\tau$$

$$+ \frac{1}{2} \sin(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} x(\tau) \sin(2\beta t \tau) w(t - \tau) \, d\tau + \epsilon. \quad (3.9)$$

Application of Lemma 2 shows that error term

$$|\epsilon| \leq \frac{P_w P_x}{\omega_0} \quad (3.10)$$

where P_w is the positive variation of the window $w(t)$ on $t \in [0, T_c]$ and where P_x is the positive variation of the input $x(t)$ on $t \in [0, T_s]$. The definition of positive variation appears as Definition 1 in Appendix 3.B.

The error bound has special meaning when $X(t)$ is a sine wave of angular frequency ω. In this case, $P_x \approx \omega T_c / \pi \leq 2\beta T_s T_c / \pi$ and hence

$$|\epsilon| \leq 2 \frac{P_w \beta T_s T_c}{\pi \omega_0} \quad (3.11)$$

which is very small for typical values of ω_0, T_c, T_s, and β.

3.2.4 Output due to Noise

As shown in Appendix 3.A (3.95 and 3.99), the normalized autocorrelation (divided by σ_i^2) of the compressive-receiver output is

$$R_o(t, d) = \frac{1}{8} \int_{-T_c}^{T_c} R_i(u_1 - d) \left\{ \right.$$

$$\int_{|u_1|}^{2T_c - |u_1|} \cos\left[(\omega_0 - 2\beta t + \beta u_2)(u_1 - d)\right] \cos\left[(\omega_0 + \beta u_2)u_1\right]$$

$$\left. w\left(\frac{u_2 + u_1}{2}\right) w\left(\frac{u_2 - u_1}{2}\right) du_2 \right\} du_1 + \epsilon. \quad (3.12)$$

The error term ϵ is bounded as

$$|\epsilon| \leq \frac{1}{8} P_w^2 B T_c \qquad (3.13)$$

where

$$B = \frac{1}{2\beta T_c} + \frac{1}{2\omega_0 - 2\beta T_s} + \frac{2}{\omega_0 - 2\beta T_s} + \frac{2}{\omega_0 - \beta T_c} \qquad (3.14)$$

with P_w being the positive variation of $w(t)$ defined by Definition 1 in Appendix 3.B.

Under typical operating constraints, the error bound can be simplified further. The term $2\beta(T_s - T_c)$ represents the total frequency spanned by the compressive filter, which is very large (typically on the order of megahertz). Additionally, the frequency ω_0 is usually in the tens to hundreds of megahertz range, hence $\omega_0 \gg 10^6$. These two facts, along with the fact that the scan time is typically twice the compression time (i.e. $T_s = 2T_c$), imply that $B < 1/\beta T_c$. Under these assumptions, the error is bounded as

$$|\epsilon| < \frac{1}{8} \frac{P_w^2}{\beta}. \qquad (3.15)$$

Of interest are special cases of the autocorrelation. When the input noise is white, meaning[1] that $\sigma_i^2 R_i(t) = \delta(t) N_0/2$, then the output noise is stationary

---

[1]Since the variance of a white noise process is undefined, arbitrarily let $\sigma_i^2 = N_0$ where N_0 is the single-sided spectral density of the white noise process. This choice make the signal-to-noise ratio, γ, consistent with other definitions in the literature.

and has autocorrelation

$$R_o(d) = \frac{1}{16} \int_{|d|}^{2T_c - |d|} \cos\left[|d|(\omega_0 + \beta u_2)\right] w\left(\frac{u_2 + |d|}{2}\right) w\left(\frac{u_2 - |d|}{2}\right) du_2$$

(3.16)

whenever $|d| \leq T_c$, otherwise $R_o(d) = 0$. If the window function w is rectangular, then

$$R_o(d) = \frac{1}{8} \frac{\sin\left[\beta|d|(T_c - |d|)\right]}{\beta|d|} \cos\left[|d|(\omega_0 + \beta T_c)\right]$$

(3.17)

whenever $|d| \leq T_c$, otherwise $R_o(d) = 0$. Regressing to the case of general stationary noise but now considering only rectangular windows,

$$R_o(t, d) = \frac{1}{8} \int_{-T_c}^{T_c} R_i(u_1 - d) g(u_1, t, d)\, du_1$$

(3.18)

where

$$g(u_1, t, d,) = \frac{\sin\left[(\beta d - 2\beta u_1)(T_c - |u_1|)\right]}{(\beta d - 2\beta u_1)}$$

$$\times \cos\left[\omega_0 d + 2\beta t(u_1 - d) - 2\omega_0 u_1 + (\beta d - 2\beta u_1)T_c\right]$$

$$+ \frac{\sin\left[\beta d(T_c - |u_1|)\right]}{\beta d} \cos\left[\omega_0 d + 2\beta t(u_1 - d) + \beta d T_c\right].$$

(3.19)

3.3 Locally Optimal Detector

We aim to develop a locally optimal detector of frequency-hopped waveforms based on a compressive-receiver output. Keying on the fact that the

optimal detector of frequency-hopped waveforms integrates coherently over a single hop period [41], we conjecture that an optimally configured compressive receiver should integrate over a period commensurate with the hop epoch T_h. But because we are also interested in the detector's performance in estimating the hop frequency, we want to eliminate interference from adjacent hops. We thus choose $T_s = T_h$ and assume that the compressive receiver is synchronized to frequency hops. This is not a realistic assumption in the pure detection problem but it will lead to an optimal detector whose performance degrades gracefully upon relaxing this assumption.

Because the interfering noise is typically of much larger bandwidth than the hop rate, the correlation between hops is negligible and so the optimal multihop detection statistic is simply the sum of the optimal single-hop detection statistics. We thus confine ourselves to the problem of using the compressive receiver to optimally detect, given an observation period of T_s, a sine wave of unknown amplitude and phase and whose frequency is one of the known hop frequencies.

Based on the above assumptions, the detection problem is now

$$
\text{versus} \quad
\begin{aligned}
H_0 : \quad & x_i(t) = n_i(t) \\
H_{\gamma'} : \quad & x_i(t) = \sqrt{2S}\sin(\omega_k t + \theta) + n_i(t)
\end{aligned}
\quad (3.20)
$$

for $\gamma < \gamma'$ and $T_c \leq t \leq T_s$. The parameters θ, γ, γ', ω_k, and $n_i(t)$ are as defined in Section 3.2.1.

88

By using the results in Sections 3.2.3 and 3.2.4, the detection problem based on the output of the compressive receiver is

$$
\begin{aligned}
H_0: \quad x_o(t) &= n_o(t) \\
H_{\gamma'}: \quad x_0(t) &= \sqrt{2S'}\cos\theta\, y_c(t,\omega_k) + \sqrt{2S'}\sin\theta\, y_s(t,\omega_k) + n_o(t)
\end{aligned}
$$

versus

$$
\tag{3.21}
$$

for $\gamma < \gamma'$ and $T_c \leq t \leq T_s$ and where $n_o(t)$ is stationary, colored Gaussian noise with autocorrelation function $R_o(t)$ as defined by (3.12) and

$$
\begin{aligned}
y_c(t,\omega_k) \; \triangleq \; & \frac{1}{2}\cos(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} \cos(\omega_k \tau)\cos(2\beta t \tau) w(t-\tau)\, d\tau \\[2mm]
+ \; & \frac{1}{2}\sin(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} \cos(\omega_k \tau)\sin(2\beta t \tau) w(t-\tau)\, d\tau
\end{aligned}
$$

$$
\tag{3.22}
$$

$$
\begin{aligned}
y_s(t,\omega_k) \; \triangleq \; & \frac{1}{2}\cos(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} \sin(\omega_k \tau)\cos(2\beta t \tau) w(t-\tau)\, d\tau \\[2mm]
+ \; & \frac{1}{2}\sin(\omega_0 t + \beta t^2) \int_{t-T_c}^{t} \sin(\omega_k \tau)\sin(2\beta t \tau) w(t-\tau)\, d\tau.
\end{aligned}
$$

$$
\tag{3.23}
$$

From [47], the conditional log-likelihood function for this problem becomes

$$
\ln \Lambda[x_o(t)/\omega_k, \theta, \gamma] =
$$

$$
\sqrt{\frac{2\gamma}{\sigma_i^2 T_h}}\cos\theta \int_{T_c}^{T_s} x_o(\tau) g_c(\tau,\omega_k)\, d\tau + \sqrt{\frac{2\gamma}{\sigma_i^2 T_h}}\sin\theta \int_{T_c}^{T_s} x_o(\tau) g_s(\tau,\omega_k)\, d\tau -
$$

$$2\frac{\gamma}{T_h} \int_{T_c}^{T_s} [\cos\theta \; y_c(\tau,\omega_k) + \sin\theta \; y_s(\tau,\omega_k)] \, [\cos\theta \; g_c(\tau,\omega_k) + \sin\theta \; g_s(\tau,\omega_k)] \; d\tau$$

$$(3.24)$$

where the functions $g_c(t,\omega_k)$ and $g_s(t,\omega_k)$ are respectively defined by the integral equations

$$\int_{T_c}^{T_s} R_o\left[\frac{\tau+t}{2}, \tau-t\right] g_c(\tau,\omega_k) \, d\tau \;=\; y_c(t,\omega_k) \qquad (3.25)$$

$$\int_{T_c}^{T_s} R_o\left[\frac{\tau+t}{2}, \tau-t\right] g_s(\tau,\omega_k) \, d\tau \;=\; y_s(t,\omega_k) \qquad (3.26)$$

for $T_c \leq t \leq T_s$. Since we are interested in a locally optimal test (i.e. small γ), we neglect the last term of (3.24) and say

$$\ln \Lambda\,[x_o(t)/\omega_k, \theta, \gamma] \quad \approx$$

$$\sqrt{\frac{2\gamma}{\sigma_i^2 T_h}} \cos\theta \int_{T_c}^{T_s} x_o(\tau) g_c(\tau,\omega_k) \, d\tau + \sqrt{\frac{2\gamma}{\sigma_i^2 T_h}} \sin\theta \int_{T_c}^{T_s} x_o(\tau) g_s(\tau,\omega_k) \, d\tau.$$

$$(3.27)$$

Averaging this approximate likelihood ratio over θ and ω_k yields

$$\Lambda\,[x_o(t)/\gamma] = \sum_{k=1}^{K} I_0\left[\sqrt{\frac{2\gamma}{\sigma_i^2 T_h}} \left|\int_{T_c}^{T_s} x_o(\tau) G_k(\tau)\right|\right] \qquad (3.28)$$

where I_0 is the modified Bessel function of the first kind and zero order and where the complex-valued function $G_k(t)$ is defined as

$$G_k(t) \triangleq g_c(t,\omega_k) + i \; g_s(t,\omega_k). \qquad (3.29)$$

90
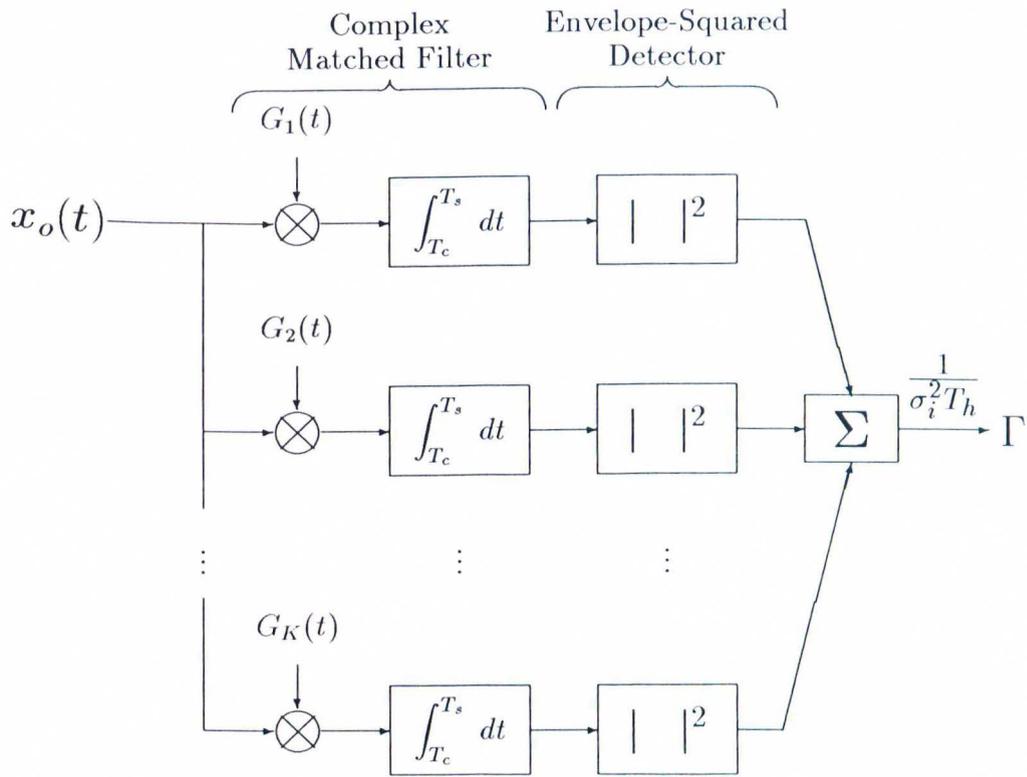
Figure 3.2: Locally Optimal Single-Epoch Detector

Consider again the small γ case and note that $I_0(x) \approx 1 + x^2/4$ for small x.

Conjure (3.28) into a locally optimal statistic

$$\Gamma = \frac{1}{\sigma_i^2 T_h} \sum_{k=1}^{K} \left| \int_{T_c}^{T_s} x_o(\tau) G_k(\tau) \, d\tau \right|^2 \qquad (3.30)$$

where the scale factor $1/\sigma_i^2 T_h$ is added for convenience in future analyses.

Figure 3.2 blocks out (3.30). To complete the detector, Γ is compared against a

threshold v, whose value determines the probability of false alarm P_F. (Section

3.5 shows the exact relationship between v and P_F.) The statistic γ being

locally optimal will, for small signal-to-noise ratios, yield the greatest possible

probability of detection, hence it is locally the most powerful test in terms of

91

signal-to-noise ratio.

3.4   Time-Multiplexed Detector

The locally optimal detector of the previous section efficiently detects frequency-hopping waveforms. As will be shown, it rivals the optimal detector that directly observes the original time waveform. Unfortunately, it also rivals the optimal detector in implementation complexity and thus undermines the attractive simplicity of the compressive receiver. In this section, we construct a detector based on the locally optimal detector that maintains simplicity of implementation for a small performance cost.

The time-multiplexed detector depicted in Figure 3.3 is both efficient and simple. The detector consists of a complex filter whose impulse response $G_\Sigma(t)$ is constructed from the pseudo-signals $G_k(t)$ by the equation

$$G_\Sigma(t) = \sum_{j=1}^{K} G_j \left[T_s - t + (j-1)\Delta T \right]. \tag{3.31}$$

The envelope of the filter's squared output is then sampled at times,

$$t = T_s + (k-1)\Delta T \qquad \text{for } k = 1, \cdots, K \tag{3.32}$$

then summed and scaled by $1/\sigma_i^2 T_h$ to produce the test statistic $\widehat{\Gamma}$. It is easily shown that

$$\widehat{\Gamma} = \frac{1}{\sigma_i^2 T_h} \sum_{k=1}^{K} \left| \int_{T_c}^{T_s} x_o(\tau) G_\Sigma \left[T_s + (k-1)\Delta T - \tau \right] \, d\tau \right|^2 \tag{3.33}$$

92
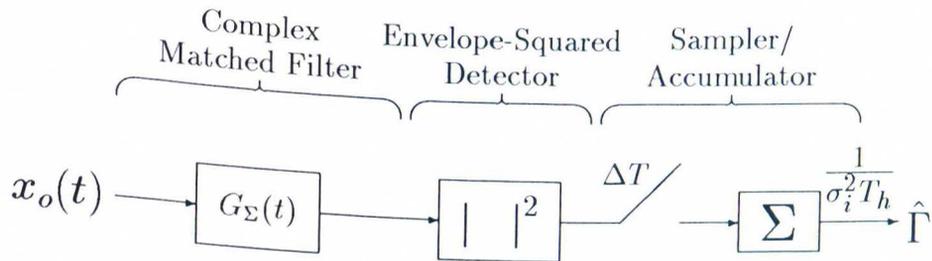
Figure 3.3: Time-Multiplexed Detector

which by defining

$$\widehat{G_k}(t) \triangleq G_\Sigma\left[T_s - t + (k-1)\Delta T\right] \qquad \text{for } T_c \leq t \leq T_s \qquad (3.34)$$

yields the alternate expression

$$\widehat{\Gamma} = \frac{1}{\sigma_i^2 T_h} \sum_{k=1}^{K} \left| \int_{T_c}^{T_s} x_o(\tau) \widehat{G}_k(\tau) \, d\tau \right|^2 . \qquad (3.35)$$

Note that (3.31) transforms (3.34) to

$$\widehat{G}_k(t) = \sum_{j=1}^{K} G_j\left[(j-k)\Delta T + t\right] \qquad (3.36)$$

which, for $\Delta T \geq T_s - T_c$, implies $\widehat{G}_k(t) = G_k(t)$, since the G_js are zero outside the range $T_c \leq t \leq T_s$. Hence the time-multiplexed detector is equivalent to locally optimal detector for this choice of ΔT. We aim to show that ΔT can be made significantly smaller for a small performance cost, thus concluding that the time-multiplexed detector captures most of the detectability of the locally

93

optimal detector but maintains the simplicity of the compressive-receiver configuration.

3.5   Detector Performance Analysis

In both the locally optimal detector and the time-multiplexed detector, the test statistic is the sum of squares of a large number of weakly correlated random variables. Namely, for the locally optimal detector,

$$\Gamma = \sum_{j=1}^{2K} \zeta_j^2 \tag{3.37}$$

where

$$\zeta_j \triangleq \frac{1}{\sqrt{\sigma_i^2 T_h}} \int_{T_c}^{T_s} x_o(\tau) h_j(\tau)\, d\tau \tag{3.38}$$

and where $h_j(t)$ is defined as below with $1 \le m \le K$,

$$h_{2m-1}(t) = g_c(t, \omega_m) \tag{3.39}$$

$$h_{2m}(t) = g_s(t, \omega_m). \tag{3.40}$$

Similarily for the time-multiplexed detector, the test is

$$\widehat{\Gamma} = \sum_{j=1}^{2K} \hat{\zeta}_j^2 \tag{3.41}$$

where

$$\hat{\zeta}_j \triangleq \frac{1}{\sqrt{\sigma_i^2 T_h}} \int_{T_c}^{T_s} x_o(\tau) \widehat{h}_j(\tau)\, d\tau \tag{3.42}$$

and for $1 \leq m \leq K$,

$$\widehat{h}_{2m-1}(t) \;\triangleq\; \widehat{g}_c(t, \omega_m) \tag{3.43}$$

$$\widehat{h}_{2m}(t) \;\triangleq\; \widehat{g}_s(t, \omega_m). \tag{3.44}$$

Here, $\widehat{g}_c(t, \omega_m) \triangleq \Re\left[\widehat{G}_m(t)\right]$ and $\widehat{g}_s(t, \omega_m) \triangleq \Im\left[\widehat{G}_m(t)\right]$. In the analysis to follow, the hat notation will be dropped since the results apply to each detector in the same way. In other words, to get the result for the time-multiplexed detector, add hats to the appropriate variables.

Because the test statistic is the sum of a large number of weakly correlated random variables, there is reason to believe, despite the correlation, that the statistic has approximately Gaussian distribution. We proceed under this assumption with justification to follow later. To specify the asymptotic distribution of Γ, we need its mean and variance under the signal-present and signal-absent hypotheses. For this purpose, define $z_j(t)$, $1 \leq j \leq 2K$, as

$$z_{2m-1}(t) \;=\; y_c(t, \omega_m) \tag{3.45}$$

$$z_{2m}(t) \;=\; y_s(t, \omega_m) \tag{3.46}$$

with m ranging between 1 and K, while $\bar{z}_j(t)$, for $1 \leq j \leq 2K$, is defined as

$$\bar{z}_{2m-1}(t) \;=\; \int_{T_c}^{T_s} R_o\left[\frac{t+\tau}{2}, t - \tau\right] g_c(t, \omega_m)\, d\tau \tag{3.47}$$

$$\bar{z}_{2m}(t) \;=\; \int_{T_c}^{T_s} R_o\left[\frac{t+\tau}{2}, t - \tau\right] g_s(t, \omega_m)\, d\tau \tag{3.48}$$

95

with m also ranging between 1 and K. For the same reason as above, we define the time cross correlations

$$\xi_{m,n} = \frac{1}{T_h} \int_{T_c}^{T_s} z_m(\tau) h_n(\tau) \, d\tau \tag{3.49}$$

$$\bar{\xi}_{m,n} = \frac{1}{T_h} \int_{T_c}^{T_s} \bar{z}_m(\tau) h_n(\tau) \, d\tau \tag{3.50}$$

for $1 \leq m, n \leq K$. (For the locally optimal detector case, note that $\bar{\xi}_{j,k} = \xi_{j,k}$.) Assume now that the signal is at frequency ω_l. Then

$$\zeta_k = \sqrt{2\gamma'} \cos \theta \; \xi_{2l-1,k} + \sqrt{2\gamma'} \sin \theta \; \xi_{2l,k} + \eta_k \tag{3.51}$$

where the random variable θ is uniformly distributed on $[0, 2\pi]$ and the η_ks are zero-mean Gaussian with covariances $\mathcal{E}(\eta_j \eta_k) = \bar{\xi}_{j,k}$. From (3.132) in Appendix 3.C

$$\mu_{k/l} \triangleq \mathcal{E}(\zeta_k^2/\omega_l) = \gamma'(\xi_{2l-1,k}^2 + \xi_{2l,k}^2) + \bar{\xi}_{k,k}. \tag{3.52}$$

When averaged over l and summed over k, the mean of Γ is

$$M_{\gamma'} = \frac{\gamma'}{K} \sum_{k=1}^{2K} \sum_{l=1}^{2K} \xi_{k,l}^2 + \sum_{k=1}^{2K} \bar{\xi}_{k,k}. \tag{3.53}$$

Use (3.51) and (3.136) to construct the covariance between the jth and kth terms of Γ when the signal is at frequency ω_l. The result is

$$\nu_{j,k/l} = 2(\gamma')^2 \, \xi_{2l-1,j} \, \xi_{2l,j} \, \xi_{2l-1,k} \, \xi_{2l,k} + \frac{(\gamma')^2}{2} \, \xi_{2l-1,j}^2 \, \xi_{2l-1,k}^2 + \frac{(\gamma')^2}{2} \, \xi_{2l,j}^2 \, \xi_{2l,k}^2$$

$$- \frac{(\gamma')^2}{2} \, \xi_{2l-1,j}^2 \, \xi_{2l,k}^2 - \frac{(\gamma')^2}{2} \, \xi_{2l,j}^2 \, \xi_{2l-1,k}^2 + 4\gamma' \xi_{2l-1,j} \xi_{2l-1,k} \bar{\xi}_{j,k}$$

$$+ \; 4\gamma' \xi_{2l,j} \xi_{2l,k} \bar{\xi}_{j,k} + 2 \, \bar{\xi}_{j,k}^2 \tag{3.54}$$

96

which upon averaging over l and summing over j and k, becomes

$$V_{\gamma'} =$$

$$\frac{(\gamma')^2}{K} \sum_{j=1}^{2K} \sum_{k=1}^{2K} \sum_{l=1}^{K} \left(2\, \xi_{2l-1,j}\, \xi_{2l,j}\, \xi_{2l-1,k}\, \xi_{2l,k} - \frac{1}{2}\, \xi_{2l-1,j}^2\, \xi_{2l,k}^2 - \frac{1}{2}\, \xi_{2l,j}^2\, \xi_{2l-1,k}^2 \right)$$

$$+ \frac{\gamma'}{K} \sum_{j=1}^{2K} \sum_{k=1}^{2K} \sum_{l=1}^{2K} \left(\frac{\gamma'}{2}\, \xi_{l,j}^2\, \xi_{l,k}^2 + 4\, \xi_{l,j}\, \xi_{l,k}\, \bar{\xi}_{j,k} \right) + 2 \sum_{j=1}^{2K} \sum_{k=1}^{2K} \bar{\xi}_{j,k}^2. \tag{3.55}$$

Of course, for the signal-not-present case, the mean and variance are simply (3.53) and (3.55) with the signal-to-noise ratio $\gamma' = 0$.

Since the test statistic Γ has an approximately Gaussian distribution, the threshold v and probability of detection P_D, for a given probability of false alarm, follow as

$$v = \sqrt{V_0}\, \Phi^{-1}(1 - P_F) + M_0 \tag{3.56}$$

and

$$P_F = 1 - \Phi \left(\frac{\sqrt{V_0}\, \Phi^{-1}(1 - P_F) - M_{\gamma'} + M_0}{\sqrt{V_{\gamma'}}} \right) \tag{3.57}$$

where $\Phi(x)$ is the distribution function of the standard Gaussian.

We now justify the above use of the central limit theorem. First let,

$$\boldsymbol{\Sigma} \triangleq \begin{bmatrix} \bar{\xi}_{1,1} & \cdots & \bar{\xi}_{1,2K} \\ \vdots & & \vdots \\ \bar{\xi}_{2K,1} & \cdots & \bar{\xi}_{2K,2K} \end{bmatrix} \tag{3.58}$$

which is the covariance matrix of the ζ_is, and

$$\boldsymbol{E}_{\theta,l} \triangleq \begin{bmatrix} \xi_{2l-1,1} \cos\theta + \xi_{2l,1} \sin\theta \\ \vdots \\ \xi_{2l-1,2K} \cos\theta + \xi_{2l,2K} \sin\theta \end{bmatrix} \tag{3.59}$$

97

which are $1/\sqrt{2\gamma'}$ times the means of the ζ_is under the condition that the signal has phase θ and frequency ω_l. We note that, since Σ is nonnegative definite and symmetric, there exists a square-root matrix $\Sigma^{\frac{1}{2}}$ such that $\Sigma^{\frac{T}{2}}\Sigma^{\frac{1}{2}} = \Sigma$. Consider also the diagonalization of $\Sigma^{-\frac{T}{2}}\Sigma^{-1}\Sigma^{-\frac{1}{2}} = T^T \Lambda T$ where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{2K} \end{bmatrix} \tag{3.60}$$

is the matrix of eigenvalues of Σ and T is an orthogonal matrix of eigenvectors. Use the above diagonalization and $\Sigma^{\frac{1}{2}}$ to rewrite the test statistic as

$$\Gamma = \left(G + \sqrt{2\gamma'}M_{\theta,l}\right)^T \Lambda \left(G + \sqrt{2\gamma'}M_{\theta,l}\right) \tag{3.61}$$

where

$$G = \begin{bmatrix} g_1 \\ \vdots \\ g_2 \end{bmatrix} \tag{3.62}$$

with $\{g_i\}$ independent, zero mean, unity variance, and Gaussian and where

$$M_{\theta,l} \triangleq \begin{bmatrix} m_{1,\theta,l} \\ \vdots \\ m_{n,\theta,l} \end{bmatrix} = T^T \Sigma^{-\frac{1}{2}} E_{\theta,l}. \tag{3.63}$$

The test statistic Γ is now the sum of squares of independent Gaussian variables. Through application of the Berry-Esseén Theorem (see [25]), Γ, conditioned on θ and l, is approximately Gaussian distributed with an error no more than $4c/\sigma$ where

$$c = \max_{i,\theta,l} \lambda_i \frac{48\gamma' m_{i,\theta,l}^2 + 7}{8\gamma' m_{i,\theta,l}^2 + 2} \tag{3.64}$$

$$\sigma^2 = \sum_{i=1}^{2K} \lambda_i^2 (8\gamma' m_{i,\theta,l}^2 + 2). \tag{3.65}$$

If this error bound is small, a fact that must be established numerically, then the CLT applies uniformly to the conditional distribution of Γ. If, in addition, the overall mean $M_{\gamma'}$ and variance $V_{\gamma'}$ remain essentially constant with respect to l and θ, then the CLT applies to the unconditional distribution of Γ as well. This fact must also be established numerically.

The above analysis using the asymptotic distribution of Γ is supplemented with upper- and lower-bounding distributions. Specifically,

$$1 - Q_K \left[\sqrt{2\gamma' e_{\max}}, \sqrt{\frac{a}{\lambda_{\max}}} \right] \leq \Pr\left[\Gamma \leq a\right] \leq 1 - Q_K \left[\sqrt{2\gamma' e_{\min}}, \sqrt{\frac{a}{\lambda_{\min}}} \right] \quad (3.66)$$

where $\lambda_{\max}$, $\lambda_{\min}$ are respectively the maximum and minimum eigenvalues of the covariance $\boldsymbol{\Sigma}$ and where $e_{\max}$, $e_{\min}$ are respectively the maximum and minimum over all eigenvalues of the matrices

$$\boldsymbol{B}_l \triangleq \boldsymbol{A}_l^T \boldsymbol{\Sigma} \boldsymbol{A}_l, \quad 1 \leq l \leq K \quad (3.67)$$

with

$$\boldsymbol{A}_l \triangleq \begin{bmatrix} \xi_{2l-1,1} & \xi_{2l,1} \\ \vdots & \vdots \\ \xi_{2l-1,2K} & \xi_{2l,2K} \end{bmatrix} \quad (3.68)$$

and finally where Q_m is the generalized Marcum Q-function defined as

$$Q_m(\alpha, \beta) \triangleq \int_\beta^\infty x \left(\frac{x}{\alpha}\right)^{m-1} e^{-\frac{1}{2}(x^2 + \alpha^2)} I_{m-1}(\alpha x) \, dx. \quad (3.69)$$

Of interest is that the upper bound equals the lower bound only when the channel outputs are i.i.d. and the sum of the square magnitude of the
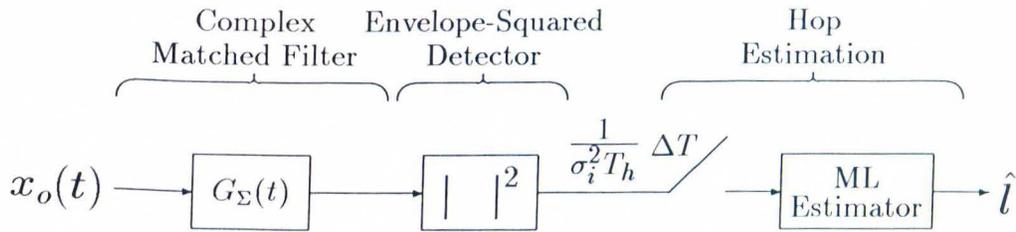
Figure 3.4: Hop Frequency Estimator

signal component across the channels is independent of signal phase. In a sense, the bounds give an indication of how well the detector fits the i.i.d. assumption, since the upper bound corresponds to the detector distribution under the i.i.d. and phase independence assumptions, but with an increased noise level, while the lower bound has the same interpretation, but with a decreased noise level. These bounds, when averaged, approximate the detector distribution, the usefulness of which will be studied and compared with the asymptotic distribution in Section 3.7.

3.6   Hop Frequency Estimator

The time-multiplexed detector can be modified to estimate hop frequency as depicted in Figure 3.4. Here, instead of accumulating samples of the square

envelope, the maximum sample is found as

$$|G_{\hat{i}}|^2 = \max_{1 \leq l \leq K} |G_l|^2. \tag{3.70}$$

Then $\omega_{k_{\hat{i}}}$ is declared the hop frequency estimate.

We need the performance of this maximum likelihood estimator. For this purpose, the probability of track P_T is defined as the probability that, given the presence of a signal, the hop estimate matches the actual hop frequency. Because of the lack of symmetry and independence between the squared envelope samples, an exact expression of P_T that is also computable is very difficult to find, although it can be expressed as a K-dimensional integral. We will be satisfied to tightly bound P_T when it is large. The results will be accurate under the useful operating conditions of the estimator.

Proceeding in this fashion, we will need to define the distribution function of each sample

$$F_k(x) \triangleq \Pr\left(|G_k|^2 \leq x\right) \tag{3.71}$$

and note that its density $f_k(x)$ exists and will be computed exactly later. Now suppose the signal is at hop l; then the probability that $|G_l|^2$ is between x and $x + \Delta x$ is approximately $f_l(x)\Delta x$ and, given that it is between these two numbers, the probability that it is the maximum, hence correctly chosen, is

approximately

$$T_l(x) \triangleq \Pr\left[\bigcap_{\substack{k=1 \\ k \neq l}}^{K} \left(|G_k|^2 \leq x\right)\right]. \tag{3.72}$$

Therefore, upon letting $\Delta x \to 0$ and averaging over l,

$$P_T = \frac{1}{K} \sum_{l=1}^{K} \int_0^\infty f_l(x) T_l(x) \, dx. \tag{3.73}$$

In order to produce the promised bounds on P_T, we will need to produce bounds on $T_l(x)$. Using De Morgan's law,

$$T_l(x) = 1 - \Pr\left[\bigcup_{\substack{k=1 \\ k \neq l}}^{K} \left(|G_k|^2 > x\right)\right] \tag{3.74}$$

$$= 1 - \sum_{\substack{\text{singles} \\ \sim l}} P_1(x) + \sum_{\substack{\text{pairs} \\ \sim l}} P_2(x) + \cdots \tag{3.75}$$

where $P_1(x)$ is the probability that a single sample exceeds x and, similarly, $P_2(x)$ is the probability that both of a given pair of samples exceed x. The above expression is useful, because, for the case of large x, the events that samples other than the lth exceed x are approximately disjoint. In this case, the first two terms of the series accurately approximate $T_l(x)$. With enough terms, this series can determine $T_l(x)$ to any desired degree of accuracy and, being alternating, implies bounds on $T_l(x)$ since the series out to positive terms is above the actual probability and, similarly, the series out to negative terms is below. We will use the first three terms of the series in just this way along with (3.71) to write

$$b_l(x) \leq T_l(x) \leq b_l(x) + \epsilon_l(x) \tag{3.76}$$

where

$$b_l(x) = 1 - \sum_{\substack{k=1 \\ k \neq l}}^{K} [1 - F_k(x)] \tag{3.77}$$

$$\epsilon_l(x) = \sum_{\substack{k=1 \\ k \neq l}}^{K} \sum_{\substack{j=1 \\ j \neq l,k}}^{K} \{1 - \max [F_k(x), F_j(x)]\}. \tag{3.78}$$

Equations (3.73) and (3.76) now determine

$$P_T \geq \frac{1}{K} \sum_{l=1}^{K} \int_0^\infty f_l(x) \max [b(x), 0] \ dx \tag{3.79}$$

$$P_T \leq \frac{1}{K} \sum_{l=1}^{K} \int_0^\infty f_l(x) \min [b(x) + \epsilon(x), 1] \ dx \tag{3.80}$$

which, along with explicit expressions for $f_l(x)$ and $F_l(x)$ below, complete the

performance analysis of the hop estimator:

$$F_l(x) = \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_0^{\sqrt{x}} \int_0^{2\pi} \int_0^{2\pi} r e^{-\frac{1}{2}(\boldsymbol{Z}-\boldsymbol{M})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}-\boldsymbol{M})} \ d\phi d\theta dr \tag{3.81}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \bar{\xi}_{2l-1,2k-1} & \bar{\xi}_{2l-1,2k} \\ \bar{\xi}_{2l,2k-1} & \bar{\xi}_{2l,2k} \end{bmatrix} \tag{3.82}$$

and

$$\boldsymbol{Z} = \begin{bmatrix} r \cos \phi \\ r \sin \phi \end{bmatrix} \tag{3.83}$$

and

$$\boldsymbol{M} = \sqrt{2\gamma'} \begin{bmatrix} \hat{\xi}_{2l-1,2k-1} \cos \theta + \hat{\xi}_{2l,2k-1} \sin \theta \\ \hat{\xi}_{2l-1,2k} \cos \theta + \hat{\xi}_{2l,2k} \sin \theta \end{bmatrix} \tag{3.84}$$

following upon differentiation is

$$f_l(x) = \frac{1}{2|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_0^{2\pi} \int_0^{2\pi} e^{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{M})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{M})} \ d\phi d\theta \tag{3.85}$$

where

$$Y = \left[\begin{array}{c} \sqrt{x}\cos\phi \\ \sqrt{x}\sin\phi \end{array} \right].$$
(3.86)

3.7 Performance Comparisons

This section graphically compares the performance of the locally optimal detector, based on the compressive-receiver output, to that of the optimal detector, based on receiver input. Also evaluated is the performance of the time-multiplexed detector as a function of the sampling ΔT.

The parameters chosen to make comparisons are: $T_c = 50$ μs, $T_s = 1000$ μs, $\omega_0 = 2\pi \times 40$ Mhz, the minimum hop frequency equals 2 Mhz and the maximum hop frequency equals 4 Mhz, 100 hop frequencies, and a hop rate of 20 Khops/sec.

Figure 3.5 shows how the locally optimal detector compares with the optimal detector based on the original observations. As expected, for low SNRs the locally optimal and the optimal compare favorably. On the other hand, there is about a 3 db difference between the performances in the high region. Two factors are responsible. One is that the analysis of the optimal receiver used the CLT, a poor model when one channel dominates as in the high SNR case. It is, in fact, optimistic. Another, as explained later, is a modeling
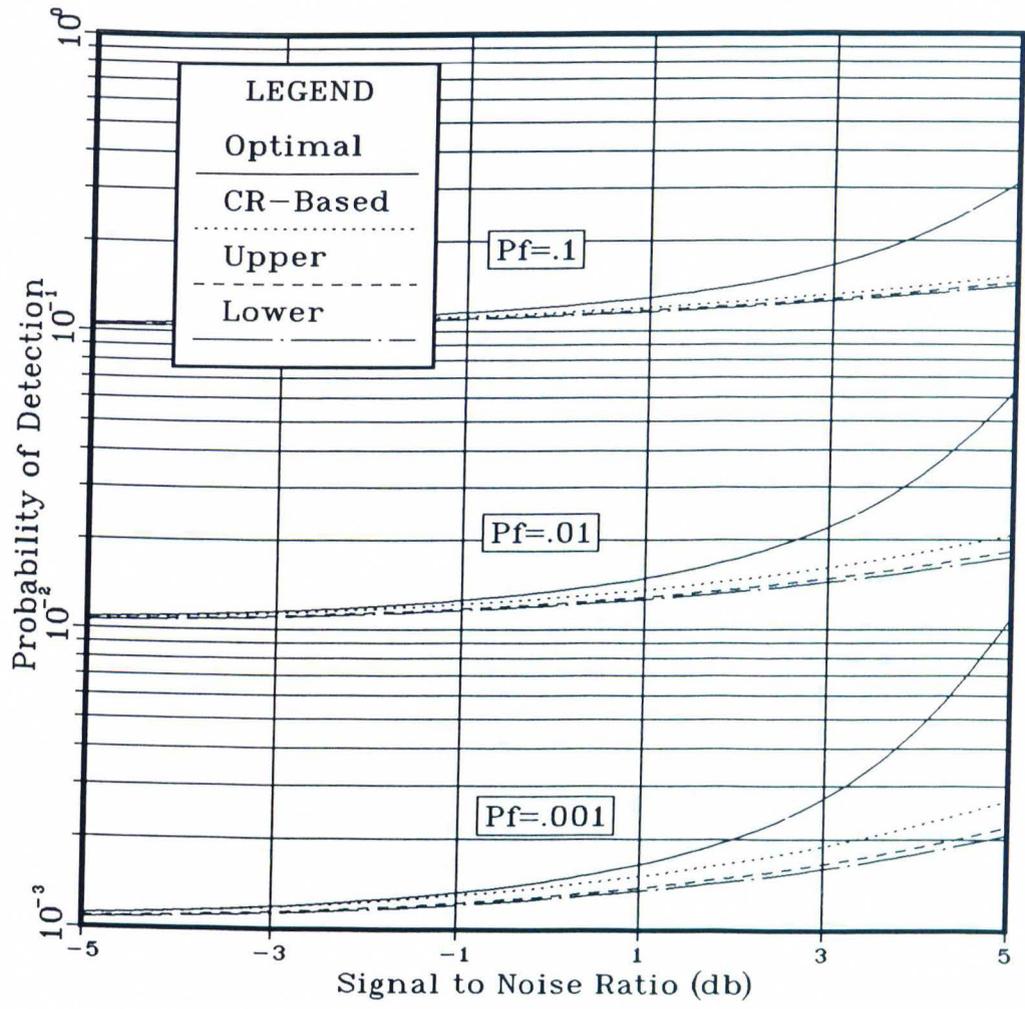
Figure 3.5: Performance of Locally Optimal, Compressive Receiver-Based Detector versus Optimal Detector Based on Direct Observations

phenomenon that can account for up to 3-db error between predicted and actual performances. Also note how the distribution bounds are pessimistic relative to the CLT analysis. This also may be due to the inappropriateness of the CLT when one channel dominates.

Figure 3.6 shows how performance degrades immediately as ΔT becomes less than the compression time T_c. This indicates that any tradeoff between performance and duty cycle would not be worth the degradation.

Figure 3.7 differs from the previous case in that the noise is bandpass instead of white. It has the same general character as far as the degradation with respect to ΔT but is approximately 3 db better. This is not an actual performance difference but a modeling phenomenon. As far as noise analysis is concerned, the compressive receiver is a mixer followed by a narrow bandpass filter. If the input noise is white, there will be uncorrelated noise contributions from the sum and difference frequencies produced by the mixer, hence the 3 db. A more realistic scenario has noise of uniform spectral density across the analysis band of the compressive receiver. In this case, the sum contribution is filtered out. This reasoning also explains the difference between the optimal and the locally optimal detectors.
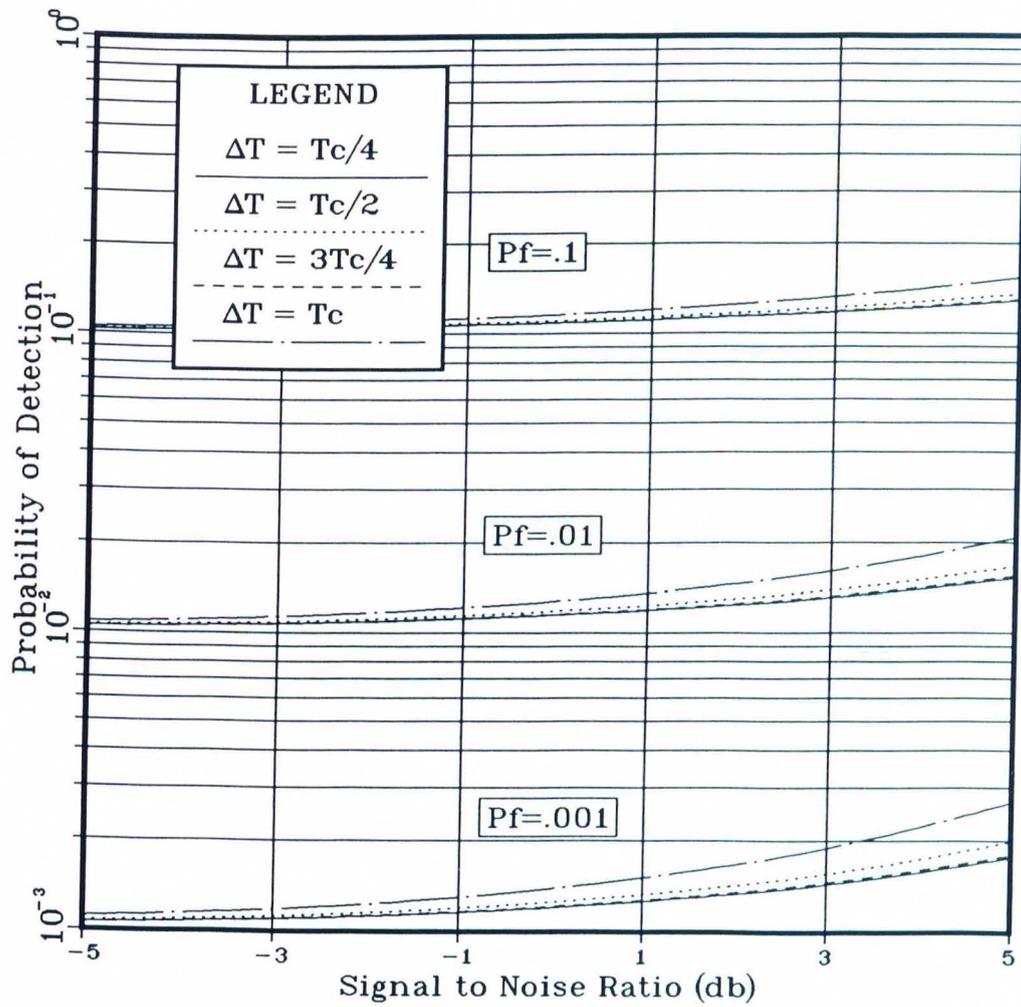
Figure 3.6: Performance of Locally Optimal Detector Compared with Time-Multiplexed Detector, White-Noise Case
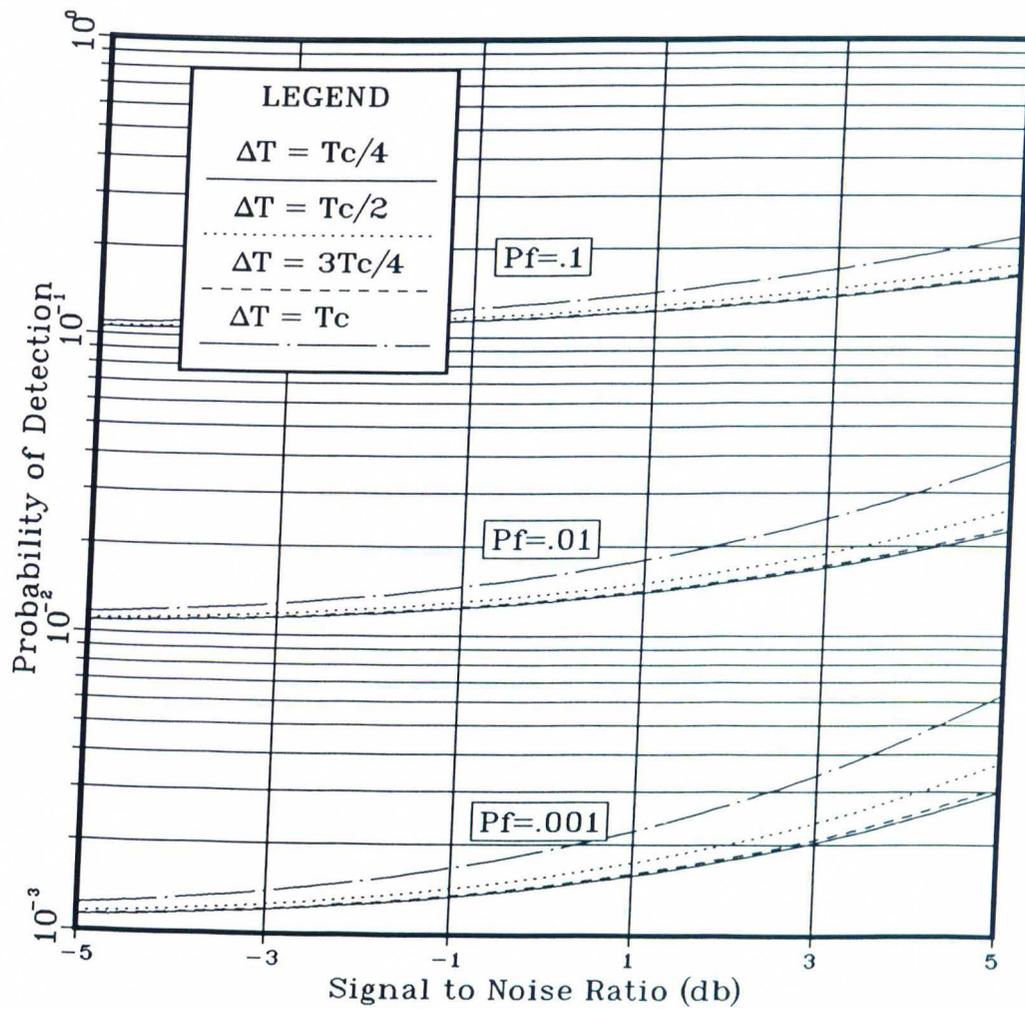
Figure 3.7: Performance of Locally Optimal Detector Compared with Time-Multiplexed Detector, Bandpass Noise Case

3.8 Conclusions

Presented were two detectors of frequency-hopped waveforms based on the compressive receiver. The first was developed by applying the likelihood-ratio theory to the observed compressive-receiver output and yielded a locally optimal (low-SNR) detector. The second, motivated by the simplicity of implementation, was a time-multiplexed version of the first that, through the choice of a parameter, either could, at the expense of a low duty cycle, achieve the detectability of the first or, at the expense of degraded performance, achieve higher duty cycles. The second detector was modified into a maximum likelihood estimator of hop frequency. Performance of both detectors and the hop frequency estimator were analyzed and compared.

The compressive receiver fulfills its promise as a simple, yet high-performing interceptor. The performance of the locally optimal detector shows that relatively little detectability is lost in the compressive receiver processing. Most of the discrepancy is due to the difference in coherent integration time (one-half for the parameters used). Furthermore, for a small performance cost, the simplicity of the compressive-receiver approach can be retained by the time-multiplexed detector. The hop frequency estimator again compares favorably with the corresponding device that used raw input instead of compressive-receiver output.

3.A Derivation of Compressive-Receiver Autocorrelation

The normalized autocorrelation of the compressive receiver output is defined to be

$$R_o(t,d) = \frac{1}{\sigma_i^2} E\left[n_o\left(t + \frac{d}{2}\right) n_o\left(t - \frac{d}{2}\right)\right] \tag{3.87}$$

under the restrictions that

$$T_c \leq t \leq T_s \tag{3.88}$$

and

$$\frac{|d|}{2} \leq \min(t - T_c, \ T_s - t). \tag{3.89}$$

Substitute the expression for the output noise, (3.7), interchange expectations and integration, and use the definition of the normalized input correlation; $R_i(\tau) = \mathcal{E}[n_i(t)n_i(t + \tau)]/\sigma_i^2$, to get

$$R_o(t,d) =$$

$$\frac{1}{\sigma_i^2} \int_0^{T_c} \int_0^{T_c} R_i(\tau_1 - \tau_2 - d)$$

$$\cos\left[\omega_0\left(t + \frac{d}{2} - \tau_1\right) - \beta\left(t + \frac{d}{2} - \tau_1\right)^2\right] \cos\left[\omega_0\tau_1 + \beta\tau_1^2\right] w(\tau_1)$$

$$\cos\left[\omega_0\left(t - \frac{d}{2} - \tau_2\right) - \beta\left(t - \frac{d}{2} - \tau_2\right)^2\right] \cos\left[\omega_0\tau_2 + \beta\tau_2^2\right] w(\tau_2) \ d\tau_2 \ d\tau_1.$$

$$\tag{3.90}$$

To exploit the stationarity of the input noise, first transform the above integral with

$$u_1 \quad = \quad \tau_1 - \tau_2 \qquad\qquad (3.91)$$

$$u_2 \quad = \quad \tau_1 + \tau_2 \qquad\qquad (3.92)$$

and then reduce the cosine components with multiple applications of the identity

$$\cos(A)\cos(B) = 1/2 \cos(A + B) + 1/2 \cos(A - B) \qquad (3.93)$$

to form

$$R_o(t, d) \quad = \frac{1}{16} \int_{-T_c}^{T_c} R_i(u_1 - d) \sum_{j=1}^{8} \Bigg\{$$

$$\int_{|u_1|}^{2T_c - |u_1|} \cos(\theta_j + \omega_j u_2 + \beta_j u_2^2) w\left(\frac{u_2 + u_1}{2}\right) w\left(\frac{u_2 - u_1}{2}\right) \, du_2 \Bigg\} du_1$$

$$(3.94)$$

where θ_j, ω_j, and β_j are given by Table 3.1. Use more trigonometry and apply Lemma 2 to terms 3–8 to rewrite the autocorrelation as

$$R_o(t, d) \quad = \frac{1}{8} \int_{-T_c}^{T_c} R_i(u_1 - d) \Bigg\{$$

$$\int_{|u_1|}^{2T_c - |u_1|} \cos\left[(\omega_0 - 2\beta t + \beta u_2)(u_1 - d)\right] \cos\left[(\omega_0 + \beta u_2)u_1\right]$$

$$w\left(\frac{u_2 + u_1}{2}\right) w\left(\frac{u_2 - u_1}{2}\right) \, du_2 \Bigg\} du_1 + \epsilon \qquad (3.95)$$

Table 3.1: Coefficients of (3.94)

j	θ_j	ω_j	β_j
1	$\omega_0 t - 2\beta t^2 - \dfrac{\beta}{2}d^2 + \beta d u_1$	$2\beta t$	0
2	$-\omega_0 t + 2\beta t^2 + \dfrac{\beta}{2}d^2 - \beta d u_1 + \beta u_1^2$	$2\omega_0 - 2\beta t$	β
3	$-(\omega_0 - 2\beta t)(u_1 - d) + \dfrac{\beta}{2}u_1^2$	$\omega_0 + \beta d - \beta u_1$	$\dfrac{\beta}{2}$
4	$-\omega_0 t + 2\beta t^2 + \dfrac{\beta}{2}d^2 - (\omega_0 + \beta d)u_1 + \dfrac{\beta}{2}u_1^2$	$\omega_0 - 2\beta t - \beta u_1$	$\dfrac{\beta}{2}$
5	$(\omega_0 t - 2\beta t)(u_1 - d) + \dfrac{\beta}{2}u_1^2$	$\omega_0 - \beta d + \beta u_1$	$\dfrac{\beta}{2}$
6	$-\omega_0 t + 2\beta t^2 + \dfrac{\beta}{2}d^2 + (\omega_0 - \beta d)u_1 + \dfrac{\beta}{2}u_1^2$	$\omega_0 - 2\beta t + \beta u_1$	$\dfrac{\beta}{2}$
7	$\omega_0 d + 2\beta t(u_1 - d)$	βd	0
8	$(\omega_0 - 2\beta t)d + (-2\omega_0 + 2\beta t)u_1$	$\beta d - 2\beta u_1$	0

with error term

$$|\epsilon| \le \frac{1}{8} P_w^2 \sum_{j=1}^{6} \int_{-T_c}^{T_c} \frac{|R_i(u_1 - d)|}{(\omega_j + 2\beta_j|u_1|)}\, du_1 \tag{3.96}$$

where P_w^2 is the positive variation of $w(t)$ defined by Definition 1 of Appendix 3.B. Simplify the error bound by minimizing each term, $\omega_j + 2\beta_j|u_1|$, with respect to u_1, d, and t, while noting the restrictions (3.88) and (3.89). The result is

$$|\epsilon| \le \frac{1}{8} P_w^2 B \int_{-T_c}^{T_c} |R_i(u_1 - d)|\, du_1 \tag{3.97}$$

where

$$B = \frac{1}{2\beta T_c} + \frac{1}{2\omega_0 - 2\beta T_s} + \frac{2}{\omega_0 - 2\beta T_s} + \frac{2}{\omega_0 - \beta T_c}. \tag{3.98}$$

The relation $|R(t)| \le 1$ to further simplifies the bound as

$$|\epsilon| \le \frac{1}{8} P_w^2 B T_c. \tag{3.99}$$

3.B Bounds on Integrals of Linearly Frequency-Modulated Sinusoids

The first bound, Lemma 1, is a tool used only to prove the second bound, Lemma 2, which is used for derivations concerning the output of the compressive receiver: namely, the derivation of the noise autocorrelation and the derivation of a simplified expression for the output signal component.

Lemma 1 *Under the restrictions that $b > a$, $\omega > 0$, $\beta \ge 0$, and $\omega + 2\beta a > 0$,*

$$\left| \int_a^b \cos(\theta + \omega t + \beta t^2) \, dt \right| \le \frac{2}{\omega + 2\beta a}. \tag{3.100}$$

Proof

Define the function

$$\Upsilon(\theta, \omega, \beta) = \int_0^{t_\pi} \sin(\tilde{\theta} + \omega t + \beta t^2) \, dt \tag{3.101}$$

where

$$t_\pi = \begin{cases} \dfrac{-1 + \sqrt{1 + 4\beta(\pi - \tilde{\theta})}}{2\beta} & \beta > 0 \\[3ex] \dfrac{\pi - \tilde{\theta}}{\omega} & \beta = 0 \end{cases} \tag{3.102}$$

$$\tilde{\theta} = \theta \bmod \pi. \tag{3.103}$$

This is simply the integral of $\sin(\tilde{\theta} + \omega t + \beta t^2)$ from $t = 0$ to its first zero crossing.

Preliminarily, three facts need to be proven: first, that $\Upsilon(\theta, \omega, \beta)$ decreases with respect to β; second, that it decreases with respect to ω; and third, that with $\beta = 0$ it decreases with respect to $\tilde{\theta}$.

Beginning with the first fact we will show that $\Upsilon(\theta, \omega, \beta)$ decreases with respect to β by proving its partial derivative to be negative. Employ the chain rule to get

$$\frac{\delta\Upsilon(\theta, \omega, \beta)}{\delta\beta} = \int_0^{t_\pi} t^2 \cos(\tilde{\theta} + \omega t + \beta t^2)\, dt. \tag{3.104}$$

Let $x = t/t_\pi$ and observe that $\beta t_\pi^2 = \pi - \tilde{\theta} - \omega t_\pi$; then

$$\frac{\delta\Upsilon(\theta, \omega, \beta)}{d\beta} = t_\pi^3 \int_0^1 x^2 \cos[\tilde{\theta} + \omega t_\pi x + (\pi - \tilde{\theta} - \omega t_\pi)x^2]\, dx. \tag{3.105}$$

To tightly bound the above integral, find its supremum by observing that from $\omega t_\pi > 0$ follows $\tilde{\theta} + \omega t_\pi x + (\pi - \tilde{\theta} - \omega t_\pi)x^2 > \pi x^2$ on $x \in [0, 1]$, from which

$$\int_0^1 x^2 \cos\left[\tilde{\theta} + \omega t_\pi x + (\pi - \tilde{\theta} - \omega t_\pi)x^2\right]\, dx < \int_0^1 x^2 \cos(\pi x^2)\, dx \tag{3.106}$$

upon noting that the cosine argument is always within the region $[0, \pi]$, a region where the cosine decreases. Substitution $u = x^2$ yields

$$\int_0^1 x^2 \cos(\pi x^2)\, dx = 2^{-1}\int_0^1 u^{\frac{1}{2}} \cos(\pi u)\, du \tag{3.107}$$

$$\leq 2^{-\frac{3}{2}}\int_0^1 \cos(\pi u)\, du = 0 \tag{3.108}$$

114

after observing that $u^{\frac{1}{2}} \cos(\pi u) \leq 2^{-\frac{1}{2}} \cos(\pi u)$ on $u \in [0,1]$. The function $\Upsilon(\theta, \omega, \beta)$ is decreasing with respect to $\beta \geq 0$, since (3.105), (3.106), and (3.108) imply that the partial derivative of $\Upsilon(\theta, \omega, \beta)$ with respect to β is negative.

The second preliminary fact, that $\Upsilon(\theta, \omega, \beta)$ decreases with respect to ω, will be proven similarly by applying the chain rule to compute

$$\frac{\delta \Upsilon(\theta, \omega, \beta)}{\delta \omega} = \int_0^{t_\pi} t \, \cos(\tilde{\theta} + \omega t + \beta t^2) \, dt. \tag{3.109}$$

Again let $x = t/t_\pi$, observe that $\beta t_\pi^2 = \pi - \tilde{\theta} - \omega t_\pi$, and apply the same reasoning leading to (3.106); then

$$\frac{\delta \Upsilon(\theta, \omega, \beta)}{\delta \omega} < t_\pi^2 \int_0^1 x \cos(\pi x^2) \, dx = 0. \tag{3.110}$$

The partial derivative $\delta \Upsilon(\theta, \omega, \beta)/\delta \omega < 0$ implies the promised result.

The third preliminary fact, that $\Upsilon(\omega, \theta, \beta)$ decreases with respect to $\tilde{\theta}$, follows because $\tilde{\theta} = \pi - \omega t_\pi$ implies

$$\Upsilon(\theta, \omega, 0) = \int_0^{t_\pi} \sin(\tilde{\theta} + \omega t) \, dt = \frac{\cos \tilde{\theta} + 1}{\omega}. \tag{3.111}$$

The fact that $\tilde{\theta} \in [0, \pi)$, a region on which the cosine decreases, clearly demonstrates that $\Upsilon(\theta, \omega, \beta)$ also decreases.

Now with the preliminary facts established, consider $\int_a^b \cos(\theta + \omega t + \beta t^2) dt$ and let $\{\eta_i\}_{i=1}^n$ be, in order, the zeros of the integrand on (a, b). In other words,

$a < \eta_{i-1} < \eta_i < b$ and, whenever $t \in (a,b)$, $\cos(\theta + \omega t + \beta t^2) = 0$, if and only if $t = \eta_i$ for some i (if there are no zeros then set $\eta_1 = b$). Decompose the integral into subintegrals between the zeros and get

$$\left| \int_a^b \cos(\theta + \omega t + \beta t^2)\, dt \right| = \left| \sum_{i=0}^n (-1)^i e_i \right| \tag{3.112}$$

where

$$e_i \begin{cases} = \Upsilon\left(\theta - \dfrac{\pi}{2}, \omega + 2\beta a, \beta\right) & \text{for } i = 0 \\[2mm] = \Upsilon(0, \omega + 2\beta\eta_i, \beta) & \text{for } 0 < i < n \\[2mm] \leq \Upsilon(0, \omega + 2\beta\eta_n, \beta) & \text{for } i = n. \end{cases} \tag{3.113}$$

Since the e_is are an alternating sequence of elements whose magnitude, after the first element, decreases,

$$\left| \int_a^b \cos(\theta + \omega t + \beta t^2)\, dt \right| \leq \max(e_0, e_1). \tag{3.114}$$

Use the fact that $\Upsilon(\theta, \omega, \beta)$ decreases with respect to β to maximize e_0, e_1 by putting $\beta = 0$. Next, maximize with respect to the other arguments to show

$$\max(e_0, e_1) \leq \Upsilon(0, \omega + 2\beta a, 0) = \frac{2}{\omega + 2\beta a}. \tag{3.115}$$

The conclusion of the lemma now follows from (3.115) and (3.114).

Definition 1 *Given the partition $\mathcal{P} = [a = t_0 < t_1 \cdots t_{n-1}, t_n = b]$, the positive variation of $x(t)$ on $[a,b]$ is*

$$P_x = \sup_{\mathcal{P}} \sum_{i=1}^n [x(t_i) - x(t_{i-1})]^+ + x(a)^+ + x(b)^+ \tag{3.116}$$

116

where r^+ has the value r, if $r > 0$, and the value zero, otherwise.

Definition 2 *Given the partition* $\mathcal{P} = [a = t_0 < t_1 \cdots t_{n-1}, t_n = b]$, *the nega-tive variation of* $x(t)$ *on* $[a, b]$ *is*

$$N_x = \inf_{\mathcal{P}} \sum_{i=1}^{n} [x(t_i) - x(t_{i-1})]^- + x(a)^- + x(b)^- \qquad (3.117)$$

where r^- *has the value* r, *if* $r < 0$, *and the value zero, otherwise.*

A function is said to be of bounded variation if both its positive and neg-ative variations are finite.

Lemma 2 *Under the restrictions that* $b > a$, $\omega > 0$, $\omega + 2\beta a > 0$, *and that* $x(t)$ *and* $y(t)$ *are piecewise continuous and of bounded variation on* $[a, b]$,

$$\int_a^b \cos(\theta + \omega t + \beta t^2) x(t) y(t) \, dt \leq \frac{2 P_x P_y}{\omega + 2\beta a} \qquad (3.118)$$

where P_x *and* P_y *are the positive variations of* $x(t)$ *and* $y(t)$.

Since the function $x(t)$, is of bounded variation, it is integrable implying that, for arbitrary $\epsilon > 0$, there exists a step function $x_s(t) = \sum_{i=1}^{n} x(t_{i-1}) I_{[t_{i-1}, t_i]}$ with corresponding partition $[a = t_0 < t_1 \cdots t_{n-1}, t_n = b]$, such that

$$\int_a^b |x(t) - x_s(t)| \, dt < \epsilon. \qquad (3.119)$$

The step function represented above is the sum of nonoverlapping steps. We want to reconstruct it as the sum of overlapping steps with the property that

the accumulated absolute amplitudes of the steps are minimal. We do this in an iterated fashion by ordering the step amplitudes $\rho_0 = x(t_{i_0}) \leq \cdots p_{j-1} = x(t_{i_{j-1}})$, $p_j = 0$, $p_{j+1} = x(t_{i_{j+1}}) \cdots \leq \rho_n = x(t_{i_n})$ with the zero amplitude included and defining the increments $l_j = [\rho_{j-1}, \rho_j]$. Starting with the kth step, an increment l_j is considered "open" if $l_j \in [0, \sum_{i=0}^{k} x(t_i)]$ and "closed" otherwise. Whenever an increment l_j transitions from open to closed, define a step of amplitude $r_{k,j} = \mathcal{L}(l_j)$ and of duration $d_{k,j}$ ranging from t_k to the time when the increment was last opened. Proceeding in this manner, the step function now has the form

$$x_s(t) = \sum_{j=0}^{n} \sum_{k=0}^{n} r_{k,j} I_{[t_k, t_k - d_{k,j}]} \tag{3.120}$$

upon setting $r_{k,j} = 0$ for previously undefined values.

At each stage in the iteration, notice that the sum of the lengths of increments either opened or closed is equal to the variation of the step function at that point, implying that $\sum_{j=0}^{n} r_{k,j} = x(t_j) - x(t_{j-1})$. Notice also that an increment opened by an increase/decrease in the step function can be closed only by a future opposite decrease/increase, meaning every step is uniquely associated with a point of increase. These facts mean that

$$\sum_{j=0}^{n} \sum_{k=0}^{n} r_{k,j} = \sum_{i=1}^{n} [x(t_i) - x(t_{i-1})]^+ + [x(t_0)]^+ \tag{3.121}$$

$$\leq P_x \tag{3.122}$$

where the last relation follows from the definition of positive variation. Prop-

118

erly define c_l and τ_l in terms of $r_{k,j}$, t_k, and $d_{k,j}$ to write

$$x_s(t) = \sum_{l=1}^{m} c_l I_{[\tau_{l-1}, \tau_l]} \tag{3.123}$$

with $\sum_{l=1}^{m} c_l \leq P_x$ and $\int_a^b |x(t) - x_s(t)| dt < \epsilon$. A similar step function exists for $y(t)$, namely,

$$y_s(t) = \sum_{l=1}^{n} b_l I_{[u_{l-1}, u_l]} \tag{3.124}$$

with $\sum_{l=1}^{n} b_l \leq P_y$ and $\int_a^b |y(t) - y_s(t)| dt < \epsilon$.

With these two step functions in hand, compute

$$\int_a^b \sin(\theta + \omega t + \beta t^2) x(t) y(t) \, dt = \int_a^b \sin(\theta + \omega t + \beta t^2) x_s(t) y_s(t) \, dt + e(t) \tag{3.125}$$

where $|e(t)| \leq \epsilon(M_{|x|} + M_{|y|}) + \epsilon^2$, $M_{|x|} = \sup_t |x(t)| < \infty$, and $M_{|y|} = \sup_t |y(t)| < \infty$. Putting (3.123) and (3.124) into (3.125) yields

$$\sum_{k=1}^{m} \sum_{l=1}^{n} b_k c_l \int_a^b \sin(\theta + \omega t + \beta t^2) I_{[\tau_{l-1}, \tau_l] \cap [u_{k-1}, u_k]} \, dt. \tag{3.126}$$

Applying Lemma 1 and maximizing the bound by replacing the starting time for each step with the worst case a forms

$$\int_a^b \cos(\theta + \omega t + \beta t^2) x(t) y(t) \, dt \leq \frac{2 P_x P_y}{\omega + 2\beta a} - e(t) \tag{3.127}$$

after noting that $\sum_{l=1}^{m} c_l \leq P_x$ and $\sum_{k=1}^{m} b_k \leq P_y$. Let $\epsilon \to 0$, then $|e(t)| \to 0$ and the lemma is proved.

119

3.C Moments between Squares of Correlated Gaussian Random Variables with Random Phase Component

We have in this section two random variables

$$P = \alpha \cos\theta + \beta \sin\theta + \nu \tag{3.128}$$

$$Q = \gamma \cos\theta + \delta \sin\theta + \eta \tag{3.129}$$

where the random variable θ is uniformly distributed on the on $[0, 2\pi]$ and the Gaussian random variables ν and η are zero-mean with covariances σ_ν^2, σ_η^2, and $\sigma_{\nu\eta}^2$. We want to compute the mean and variance of P^2, the mean and variance of P^2, and the covariance between P^2 and Q^2.

Our calculations will be assisted by a formula of the general fourth moment between the Gaussian random variables x_0, x_1, x_2, and x_3 with means $m_i = \mathcal{E}(x_i)$ and covariances $\sigma_{ij}^2 = \mathcal{E}[(x_i - m_i)(x_j - m_j)]$. The formula is

$$\mathcal{E}(x_0 x_1 x_2 x_3) = \sigma_{01}^2 \sigma_{23}^2 + \sigma_{02}^2 \sigma_{13}^2 + \sigma_{03}^2 \sigma_{12}^2 + \sigma_{01}^2 m_2 m_3 + \sigma_{02}^2 m_1 m_3$$

$$+ \sigma_{03}^2 m_1 m_2 + \sigma_{12}^2 m_0 m_3 + \sigma_{13}^2 m_0 m_2 + \sigma_{23}^2 m_0 m_1$$

$$+ m_0 m_1 m_2 m_3. \tag{3.130}$$

We will now compute the mean of P^2 and Q^2. Equation (3.128) implies

$$\mathcal{E}[P^2/\theta] = \alpha^2 \cos^2\theta + \beta^2 \sin^2\theta + 2\alpha\beta \cos\theta \sin\theta + \sigma_\nu^2 \qquad (3.131)$$

which, upon averaging over θ, becomes

$$\mathcal{E}[P^2] = \frac{\alpha^2}{2} + \frac{\beta^2}{2} + \sigma_\nu^2. \qquad (3.132)$$

Similarly,

$$\mathcal{E}[Q^2] = \frac{\gamma^2}{2} + \frac{\delta^2}{2} + \sigma_\eta^2. \qquad (3.133)$$

Onward to the covariances. Equations (3.128), (3.129), and (3.130) imply that

$$
\begin{aligned}
\mathcal{E}[P^2 Q^2/\theta] &= \sigma_\nu^2 \sigma_\eta^2 + 2\sigma_{\nu\eta}^4 + \sigma_\nu^2 (\gamma\cos\theta + \delta\sin\theta)^2 \\[2ex]
&+ 4\sigma_{\nu\eta}^2 (\alpha\cos\theta + \beta\sin\theta)(\gamma\cos\theta + \delta\sin\theta) + \sigma_\eta^2 (\alpha\cos\theta + \beta\sin\theta)^2 \\[2ex]
&+ (\alpha\cos\theta + \beta\sin\theta)^2 (\gamma\cos\theta + \delta\sin\theta)^2 \qquad (3.134)
\end{aligned}
$$

which, upon averaging over θ, becomes

$$
\begin{aligned}
\mathcal{E}[P^2 Q^2] &= \sigma_\nu^2 \sigma_\eta^2 + 2\sigma_{\nu\eta}^4 + \frac{\sigma_\nu^2}{2}(\gamma^2 + \delta^2) + \frac{\sigma_\eta^2}{2}(\alpha^2 + \beta^2) + 2\sigma_{\nu\eta}^2(\alpha\gamma + \beta\delta) \\[2ex]
&+ \frac{3}{8}(\alpha^2\gamma^2 + \beta^2\delta^2) + \frac{1}{8}(\gamma^2\beta^2 + 4\alpha\beta\gamma\delta + \alpha^2\delta^2). \qquad (3.135)
\end{aligned}
$$

With (3.132), (3.133), and (3.135), we conclude that

$$\mathrm{cov}[P^2, Q^2] = 2\sigma_{\nu\eta}^4 + 2\sigma_{\nu\eta}^2(\alpha\gamma + \beta\delta) + \frac{1}{8}(\alpha^2\gamma^2 + \beta^2\delta^2 - \gamma^2\beta^2 - \alpha^2\delta^2) + \frac{1}{2}\alpha\beta\gamma\delta$$

$$(3.136)$$

which specializes to

$$\text{var}[P^2] = 2\sigma_\nu^4 + 2\sigma_\nu^2(\alpha^2 + \beta^2) + \frac{1}{8}(\alpha^4 + \beta^4) + \frac{1}{4}\alpha^2\beta^2 \qquad (3.137)$$

$$\text{var}[Q^2] = 2\sigma_\eta^4 + 2\sigma_\eta^2(\gamma^2 + \delta^2) + \frac{1}{8}(\gamma^4 + \delta^4) + \frac{1}{4}\gamma^2\delta^2. \qquad (3.138)$$

3.D Derivation of Upper and Lower Bounding Distributions for the Sum of Squares of Correlated Gaussian Random Variables with Random Phase Component

Theorem 1 *Define the n-dimensional vector $\boldsymbol{P}^T \triangleq [p_1, \cdots, p_n]$ to have components*

$$p_i \triangleq a_i \cos\theta + b_i \sin\theta + w_i, \quad 1 \le i \le n \qquad (3.139)$$

where each a_i and b_i is a constant, θ is a uniformly distributed random variable on $[0, 2\pi]$, and $\{w_i\}_{i=1}^n$ is a sequence of zero-mean Gaussian random variables with an invertible covariance matrix $\boldsymbol{\Sigma}$. Then, with $\boldsymbol{P}$ defined as,

$$1 - Q_{\frac{n}{2}}\left[\sqrt{e_{\max}}, \sqrt{\frac{k}{\lambda_{\max}}}\right] \le Pr\left[\boldsymbol{P}^T\boldsymbol{P} \le k\right] \le 1 - Q_{\frac{n}{2}}\left[\sqrt{e_{\min}}, \sqrt{\frac{k}{\lambda_{\min}}}\right] \qquad (3.140)$$

where $\lambda_{\max}$, $\lambda_{\min}$ are respectively the maximum and minimum eigenvalues of $\boldsymbol{\Sigma}$ and where $e_{\max}$, $e_{\min}$ are respectively the maximum and minimum eigenvalues of the matrix

$$B \triangleq A^T\boldsymbol{\Sigma}A \qquad (3.141)$$

122

with

$$A \triangleq \begin{bmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{bmatrix} \tag{3.142}$$

and, finally, where Q_m is the generalized Marcum Q-function defined as

$$Q_m(\alpha, \beta) \triangleq \int_\beta^\infty x \left(\frac{x}{\alpha} \right)^{m-1} e^{-\frac{1}{2}(x^2 + \alpha^2)} I_{m-1}(\alpha x) \, dx. \tag{3.143}$$

Proof

The proof consists of three parts. In the first part, the conditional probability $\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k/\theta]$ is expressed as an integral of a multidimensional Gaussian density over a spheroid centered at the origin. Upon transformation with a decorrelating matrix, the region of integration becomes ellipsoidal and the Gaussian density becomes independent with each of its marginal densities having unity variance. Then through eigenvalue analysis, the ellipsoidal region of integration is inscribed and circumscribed with spheroids yielding corresponding bounds on the integral. The second part of the proof shows that the integral of an independent Gaussian distribution over an arbitrary spheroid depends only on the magnitude of the mean vector and decreases with respect to it. This fact will enable further bounding in the third part after computing the minimum and maximum of the mean as a function of θ. An aftereffect will be the removal of θ dependence in the bounds allowing their direct application to the unconditional probability $\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k]$. Next the bounds, which are still expressed as integrals, are evaluated in closed form via the generalized

Marcum Q-function.

Part I

By applying the expression for a multivariate Gaussian density, the conditional probability

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \le k/\theta] = \frac{1}{2\pi \, |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_{[\boldsymbol{P}^T \boldsymbol{P} \le k]} e^{-\frac{1}{2}[\boldsymbol{P} - \boldsymbol{A}\boldsymbol{C}]^T \boldsymbol{\Sigma}^{-1} [\boldsymbol{P} - \boldsymbol{A}\boldsymbol{C}]} \, d\boldsymbol{P}$$

$$(3.144)$$

where

$$\boldsymbol{C} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}. \tag{3.145}$$

The matrix $\boldsymbol{\Sigma}^{-\frac{1}{2}}$ with the property $\boldsymbol{\Sigma}^{-\frac{T}{2}} \boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{\Sigma}^{-1}$ is guaranteed to exist, since $\boldsymbol{\Sigma}$ is an invertible covariance matrix. Furthermore, for the same reason, there exists an orthogonal matrix $\boldsymbol{T}$, such that $\boldsymbol{\Sigma} = \boldsymbol{T}^T \boldsymbol{\Lambda} \boldsymbol{T}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}$. We can now define the transformation $\boldsymbol{X} = \boldsymbol{T} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}$, from which follows

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \le k/\theta] = \frac{1}{2\pi} \int_{[\boldsymbol{X}^T \boldsymbol{\Lambda} \boldsymbol{X} \le k]} e^{-\frac{1}{2}[\boldsymbol{X} - \boldsymbol{M}_\theta]^T [\boldsymbol{X} - \boldsymbol{M}_\theta]} \, d\boldsymbol{X} \tag{3.146}$$

where

$$\boldsymbol{M}_\theta = \boldsymbol{T} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{A}\boldsymbol{C}. \tag{3.147}$$

Now, since $\boldsymbol{\Sigma}$ is an invertible covariance matrix, each entry of $\boldsymbol{\Lambda}$ (i.e. eigenvalues of $\boldsymbol{\Sigma}$) is positive. Hence

$$\left[\lambda_{\max} \boldsymbol{X}^T \boldsymbol{X} \le k\right] \subset \left[\boldsymbol{X}^T \boldsymbol{\Lambda} \boldsymbol{X} \le k\right] \subset \left[\lambda_{\min} \boldsymbol{X}^T \boldsymbol{X} \le k\right] \tag{3.148}$$

from which

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k/\theta] \geq \frac{1}{2\pi} \int_{\left[\boldsymbol{X}^T \boldsymbol{X} \leq k/\lambda_{\max}\right]} e^{-\frac{1}{2}[\boldsymbol{X}-\boldsymbol{M}_\theta]^T [\boldsymbol{X}-\boldsymbol{M}_\theta]} \, d\boldsymbol{X} \quad (3.149)$$

and

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k/\theta] \leq \frac{1}{2\pi} \int_{\left[\boldsymbol{X}^T \boldsymbol{X} \leq k/\lambda_{\min}\right]} e^{-\frac{1}{2}[\boldsymbol{X}-\boldsymbol{M}_\theta]^T [\boldsymbol{X}-\boldsymbol{M}_\theta]} \, d\boldsymbol{X}. \quad (3.150)$$

Part II

We aim to show that, for a given r,

$$R \triangleq \frac{1}{2\pi} \int_{\left[\boldsymbol{X}^T \boldsymbol{X} \leq r\right]} e^{-\frac{1}{2}[\boldsymbol{X}-\boldsymbol{M}]^T [\boldsymbol{X}-\boldsymbol{M}]} \, d\boldsymbol{X} \quad (3.151)$$

depends only on the magnitude of $\boldsymbol{M}$ and decreases as $|\boldsymbol{M}|$ gets larger.

There exists an orthogonal matrix $\boldsymbol{U}$, such that $\boldsymbol{U}\boldsymbol{M} = [|\boldsymbol{M}|, 0, \cdots, 0]^T$. The matrix $\boldsymbol{U}$ is simply a change of orthonormal basis to one that includes $\boldsymbol{M}/|\boldsymbol{M}|$ as its first member. Now $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{X}$ transforms (3.151) to

$$R = \frac{1}{2\pi} \int_{\left[\boldsymbol{Y}^T \boldsymbol{Y} \leq r\right]} e^{-\frac{1}{2}\left(y_1 - |\boldsymbol{M}|\right)^2 - \frac{1}{2} \sum_{i=2}^{n} y_i^2} \, d\boldsymbol{Y}. \quad (3.152)$$

As promised for a given r, (3.152) depends only on the magnitude of $\boldsymbol{M}$, hence the notation $R(|\boldsymbol{M}|)$.

We now show that $R(|\boldsymbol{M}|)$ decreases with respect to the magnitude of $\boldsymbol{M}$ by showing that, for any positive increment $\Delta|\boldsymbol{M}|$, the corresponding difference $\Delta R(|\boldsymbol{M}|) \triangleq R(|\boldsymbol{M}| + \Delta|\boldsymbol{M}|) - R(|\boldsymbol{M}|)$ is negative.

125

Make the respective substitutions $z_1 = y_1 - |\boldsymbol{M}|$, $z_i = y_i$, for $2 \le i \le n$, to $R(|\boldsymbol{M}|)$ and $z_1 = y_1 - |\boldsymbol{M}| - \Delta|\boldsymbol{M}|$, $z_i = y_i$, for $2 \le i \le n$, to $R(|\boldsymbol{M}| + \Delta|\boldsymbol{M}|)$. Then

$$\Delta R(|\boldsymbol{M}|) = \frac{1}{2\pi} \int_G e^{-\frac{1}{2}\boldsymbol{Z}^T\boldsymbol{Z}} \, d\boldsymbol{Z} - \frac{1}{2\pi} \int_H e^{-\frac{1}{2}\boldsymbol{Z}^T\boldsymbol{Z}} \, d\boldsymbol{Z} \qquad (3.153)$$

where the sets G and H are

$$G = \left[\boldsymbol{Z} : (z_1 + |\boldsymbol{M}| + \Delta|\boldsymbol{M}|)^2 + \sum_{i=2}^{n} z_i^2 \le r \right] \qquad (3.154)$$

$$H = \left[\boldsymbol{Z} : (z_1 + |\boldsymbol{M}|)^2 + \sum_{i=2}^{n} z_i^2 \le r \right]. \qquad (3.155)$$

Cancel out the common points of G and H then

$$\Delta R(|\boldsymbol{M}|) = \frac{1}{2\pi} \int_{G \sim H} e^{-\frac{1}{2}\boldsymbol{Z}^T\boldsymbol{Z}} \, d\boldsymbol{Z} - \frac{1}{2\pi} \int_{H \sim G} e^{-\frac{1}{2}\boldsymbol{Z}^T\boldsymbol{Z}} \, d\boldsymbol{Z}. \qquad (3.156)$$

Let $v_1 = -z_1 - 2|\boldsymbol{M}| - \Delta|\boldsymbol{M}|$, $v_i = z_i$, for $2 \le i \le n$, in the second integral above; then $H \sim G$ is mapped to $G \sim H$ and the integrals can be combined to yield

$$\Delta R(|\boldsymbol{M}|) = \frac{1}{2\pi} \int_{G \sim H} e^{-\frac{1}{2}\sum_{i=1}^{n} z_i^2} \left[1 - e^{-\frac{1}{2}(2|\boldsymbol{M}|+\Delta|\boldsymbol{M}|)^2 - z_1(2|\boldsymbol{M}|+\Delta|\boldsymbol{M}|)} \right] \, d\boldsymbol{Z}. $$
$$(3.157)$$

The coordinate z_1 is in $G \sim H$, if and only if

$$(z_1 + |\boldsymbol{M}|)^2 + 2\Delta|\boldsymbol{M}|(z_1 + |\boldsymbol{M}|) + (\Delta|\boldsymbol{M}|)^2 + \sum_{i=2}^{n} z_i \le r \qquad (3.158)$$

and

$$(z_1 + |\boldsymbol{M}|)^2 + \sum_{i=2}^{n} z_i > r. \qquad (3.159)$$

126

Upon subtracting both relations, z_1 must satisfy

$$z_1 < -|M| - \frac{\Delta|M|}{2}. \tag{3.160}$$

Using this relationship in (3.157) implies that $\Delta R(|M|) < 0$, meaning that $R(|M|)$ decreases with increasing magnitude of M.

Part III

Further bound (3.149) and (3.150) by respectively maximizing and minimizing magnitude of the mean over the random phase θ. From (3.147), the magnitude of the mean is

$$|M_\theta|^2 = M_\theta{}^T M_\theta = C^T A^T \Sigma^{-1} A C \tag{3.161}$$

following from the fact that T, being orthogonal, satisfies $T^T T = I$. The matrix $A^T \Sigma^{-1} A$, being symmetric, ensures that it can be diagonalized making

$$|M_\theta|^2 = [\cos\theta, \sin\theta] U^T \begin{bmatrix} e_{\max} & 0 \\ 0 & e_{\min} \end{bmatrix} U \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \tag{3.162}$$

where U is an orthogonal matrix and $e_{\max}$ and $e_{\min}$ are the eigenvalues of $A^T \Sigma^{-1} A$. Since U is orthogonal, it rotates the plane by some angle ϕ. This means that

$$U \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} \cos(\theta - \phi) \\ \sin(\theta - \phi) \end{bmatrix} \tag{3.163}$$

and hence that

$$|M_\theta|^2 = e_{\max} \cos^2(\theta - \phi) + e_{\min} \sin^2(\theta - \phi). \tag{3.164}$$

The eigenvalues $e_{\max}$ and $e_{\min}$ are nonnegative since $\boldsymbol{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{A}$ has a square root, namely, $\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{A}$. This fact, along with (3.164), implies

$$\sqrt{e_{\min}} \leq |\boldsymbol{M}_\theta| \leq \sqrt{e_{\max}}. \tag{3.165}$$

Use (3.165) and (3.152) to deduce from (3.150) and (3.149) the bounds

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k] \geq \frac{1}{2\pi} \int_{\left[\boldsymbol{Y}^T \boldsymbol{Y} \leq k/\lambda_{\max}\right]} e^{-\frac{1}{2}\sum_{i=2}^{n} y_i^2 - \frac{1}{2}(y_1 - \sqrt{e_{\max}})^2} \, d\boldsymbol{Y} \tag{3.166}$$

$$\Pr[\boldsymbol{P}^T \boldsymbol{P} \leq k] \leq \frac{1}{2\pi} \int_{\left[\boldsymbol{Y}^T \boldsymbol{Y} \leq k/\lambda_{\min}\right]} e^{-\frac{1}{2}\sum_{i=2}^{n} y_i^2 - \frac{1}{2}(y_1 - \sqrt{e_{\min}})^2} \, d\boldsymbol{Y}. \tag{3.167}$$

The integrals in the above bounds are simplified by showing that they are the distribution functions, evaluated respectively at $k/\lambda_{\max}$ and $k/\lambda_{\min}$ of the sum of n non-central χ^2 random variables with noncentrality $e_{\max}$, $e_{\min}$. An explicit expression for this distribution is given in [45] and leads to the conclusion of the theorem:

$$1 - Q_{\frac{n}{2}}\left[\sqrt{e_{\max}}, \sqrt{\frac{k}{\lambda_{\max}}}\right] \leq \Pr\left[\boldsymbol{P}^T \boldsymbol{P} \leq k\right] \leq 1 - Q_{\frac{n}{2}}\left[\sqrt{e_{\min}}, \sqrt{\frac{k}{\lambda_{\min}}}\right]. \tag{3.168}$$

Chapter 4

The Detection of Frequency-Hopped Waveforms via the Amplitude Distribution Function

4.1   Background and Motivation

The unfriendly detection and interception of secure communications is a topic of much current research. Secure communications usually involve some variety of spread-spectrum modulation, whose purpose is to add ambiguity or "randomness" to the communication waveform as a measure against un-intended detection or interception. The usual procedure for randomizing the waveform is pseudo-random variation of transmission times (time hopping), phases (direct sequence), or frequencies (frequency hopping). The development of a method to detect frequency-hopped waveforms is the subject of this chapter.

The use of the Amplitude Distribution Function (ADF) for detection is a new idea with potentially many diverse applications. However, although this work focuses solely on the detection of frequency-hopped waveforms, the general ADF technique can be applied to related areas such as radar or sonar detection. The central idea of the technique is that the ADF of an observed signal in additive noise is the convolution between the individual ADFs of signal and noise. We have shown through the use of deconvolution techniques that the signal component can be separated and thus detected even for small signal levels.

There are previous works that, in essence, use the ADF (see [26], [15], [34], and [36]) but none has given a precise definition and mathematical development like those offered here. Moreover, to our knowledge there is no reference that directly uses the ADF idea for detection.

The ADF indicates the time fraction that a waveform is below a given amplitude, much like a probability distribution function measures the probability that a random variable is below a given value. Previous researchers have used this concept but failed to give a precise definition of the ADF as it applies to both deterministic and stochastic signals. Our definition is given, for any signal $X(t)$, as

$$F_X(a) = \lim_{T \to \infty} \frac{1}{T} \mathcal{E}\{\mathcal{L}[t : X(t) \leq a, \, 0 \leq t < T]\} \qquad (4.1)$$
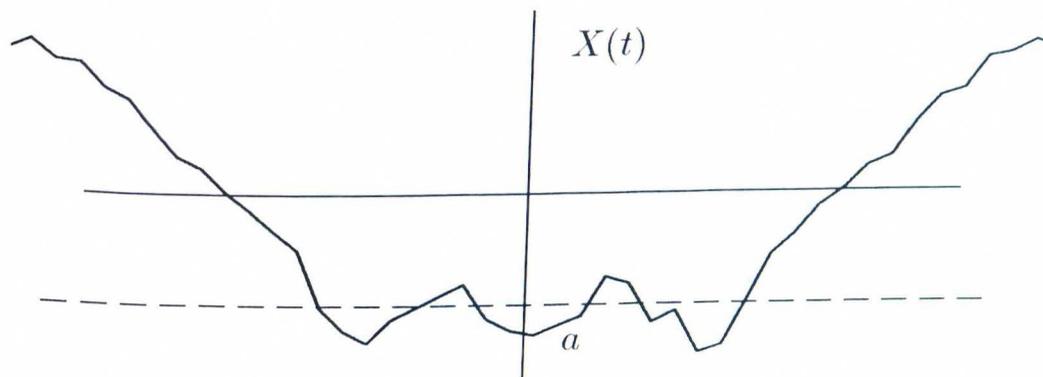
Figure 4.1: Definition of ADF

where $\mathcal{L}$ is set function giving length. Figure 4.1 illustrates that the ADF is simply the time fraction that $X(t)$ is below a given threshold a. With this definition, we have proved that, under very general conditions, the ADF of signal plus noise is the convolution of the signal ADF and the noise ADF individually. This result would not have been possible without a definition that applied to both deterministic and stochastic signals.

The adf or amplitude density function is the density, if it exists, implied by the ADF. There is, of course, a corresponding convolutional relationship between the adf of signal and noise and the individual adf's of signal and noise. Some examples exemplify this convolutional relationship and hint at the potential of the ADF in signal detection. Figure (4.2) shows the ADF

Signal
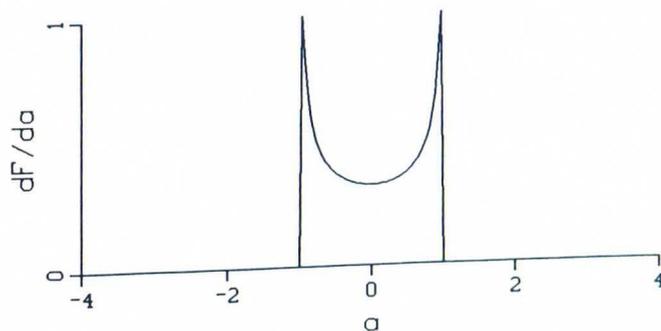
Noise
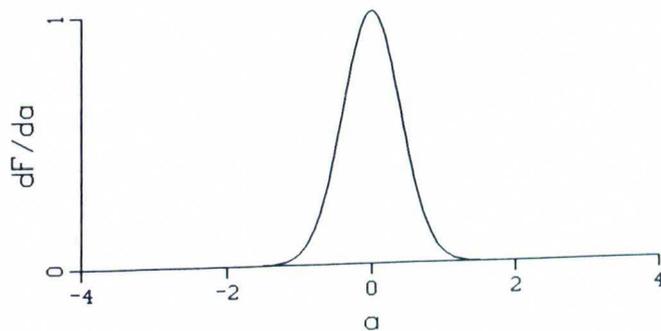
Signal plus Noise
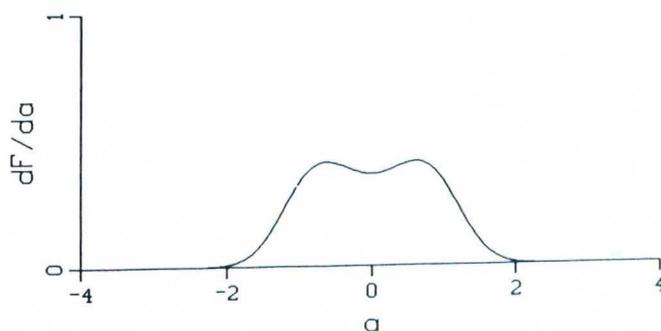
Figure 4.2: The adf of Modulated Sine Wave and Noise

of a modulated sine wave (the signal), the ADF of noise, and the ADF of signal plus noise. The main point here is that the ADF of a sine wave is invariant, under most phase and frequency modulations, but these are exactly the modulations used to thwart a potential interceptor. Therefore, the most typical spreading modulations will not degrade the performance of an ADF detector. Also of interest is the complex ADF, that is, the two-dimensional distribution in amplitude of the signal's in-phase (I) and quadrature phase (Q) components (Figure 4.3).

Both of the last two examples illustrate the convolutional spreading of the ADF due to additive noise. Like a photograph taken while the camera is out of focus, noise smears the signal part of the ADF. By a technique borrowed from image processing, the picture (i.e., ADF) can be refocused to reveal the underlying picture detail (i.e., signal). Deconvolution, as this process is called, (see [2]) involves convolving the picture with a kernel function $k(x)$ that has been specially constructed from noise details. Figure 4.4 shows the result of the process, an out-of-focus picture before and after deconvolution.

Our original idea allows the direct application of the deconvolution techniques to signal detection. Suppose a noisy signal $X(t)$, observed over the interval T, is transformed into the function $\hat{F}_X(x)$ by

$$\hat{F}_X(x) = \frac{1}{T} \int_0^T k[x - X(t)]dt \qquad (4.2)$$

133
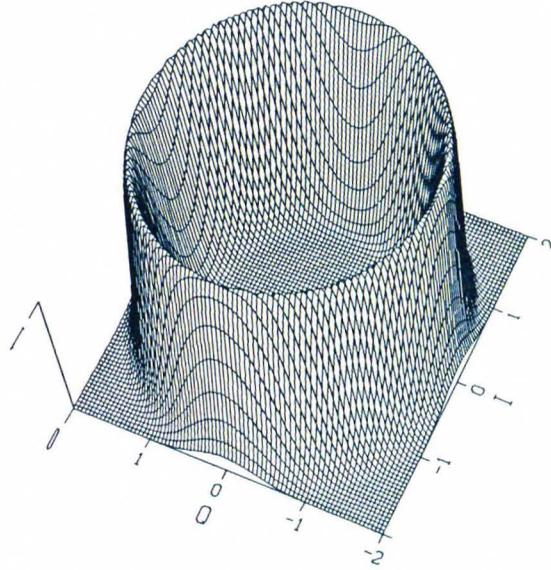
Signal

Signal plus Noise
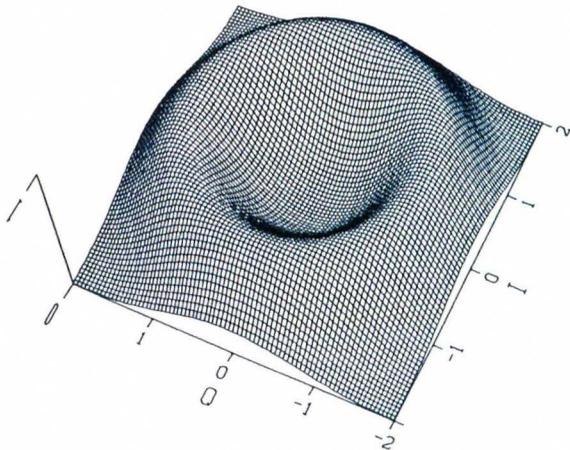
Figure 4.3: Complex adf

Before

After
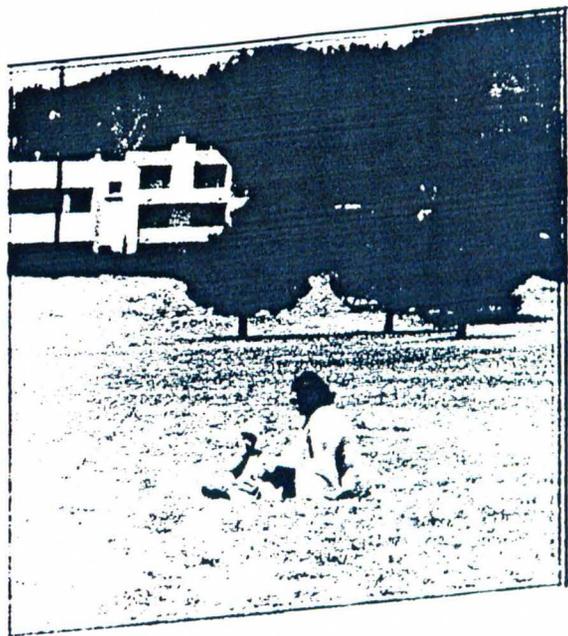
Figure 4.4: Deconvolution Applied to Out-of-Focus Picture

where $k(x)$ is the deconvolution kernel. We have shown that, for large T, $\hat{F}_X(x)$ converges directly to the deconvolved ADF. For this reason, we call this the "Deconvolution Statistic". Furthermore, we have shown that samples of the Deconvolution Statistic are approximately jointly Gaussian, to which well-known optimal detection techniques apply.

In summary, our approach consisted of precisely defining the ADF and proving the existence of an intuitive relationship between signal and noise, that of convolution. Borrowing techniques from image processing, we showed that the effects of the noise could be separated from the signal by a process called deconvolution. By transforming the observed waveform, we generated a random process that converges directly to the deconvolved ADF and upon which standard detection techniques apply.

4.2   Mathematical Tools for the ADF

The ADF, as defined here, is an original concept and thus needs a firm mathematical foundation. This section precisely defines the ADF in a way that applies equally well to deterministic and stochastic signals. This basic definition is extended to include the concepts of a joint ADF and the notion of amplitude independence, a notion analogous to independence in probability. By way of a sequence of lemmas and theorems, two significant results are

established. The first is the already promised result that the ADF of signal plus noise is the convolution of the signal ADF with the noise ADF. This is proved under the very general constraint that the second derivative of the noise autocorrelation exists and is finite at time difference zero. The remaining result of significance is a linkage between the ADF and the instantaneous probability distributions of the signal plus noise. We begin with a precise definition of the ADF.

Definition 3 *The Amplitude Distribution Function (ADF), written $F_X(a)$ for a stochastic process $X(t)$, is*

$$F_X(a) = \lim_{T \to \infty} \frac{1}{T} \mathcal{E}\left\{ \mathcal{L}\left[t : X(t) \le a,\ 0 \le t < T\right] \right\} \tag{4.3}$$

where $\mathcal{L}$ is a set function giving length. Additionally, the limit must exist for all a.

If the signal is deterministic, then the definition of the ADF reduces to

$$F_X(a) = \lim_{T \to \infty} \frac{1}{T}\left\{ \mathcal{L}\left[t : X(t) \le a,\ 0 \le t < T\right] \right\}. \tag{4.4}$$

The ADF is not a distribution function in the strict sense because its extreme values are not necessarily one or zero and it may not be right-continuous. For example, the function

$$S(t) = \begin{cases} t & \text{for } 0 < t \bmod 1 \le 1/2 \\ -t & \text{for other } t \end{cases} \tag{4.5}$$

137

has ADF identically equal to 1/2. Consider also the function

$$S(t) = \begin{cases} 1 & \text{for } 0 < t \bmod 1 \leq 1/2 \\ -1 & \text{for other } t. \end{cases} \qquad (4.6)$$

Its ADF is

$$F_S(a) = \begin{cases} 0 & \text{for } \infty < a < -1 \\ 1/2 & \text{for } -1 \leq a \leq 1 \\ 1 & \text{for } 1 < a \end{cases} \qquad (4.7)$$

which is right-continuous at -1 and left-continuous at 1. The ADF, not necessarily being a true distribution, creates problems in situations which require an ADF-induced measure, for instance, the Lebesgue-Stieltjes integral. In these cases, we use the right-continuous extension of the ADF, defined as

$$F^*(a) = \lim_{x \to a+} F(x). \qquad (4.8)$$

Analogous to the joint probability function of random variables, there exists a joint ADF of two different stochastic processes, defined as follows:

Definition 4 *The joint ADF, written $F_{X,Y}(a,b)$ for stochastic processes $X(t)$ and $Y(t)$, is*

$$F_{X,Y}(a,b) = \lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \mathcal{L} \left[t : X(t) \leq a \text{ and } Y(t) \leq b, \, 0 \leq t < T \right] \right\}. \qquad (4.9)$$

This definition will be used to define the following concept of amplitude independence, analogous to that of independence between random variables.

Definition 5 *Two stochastic processes $X(t)$ and $Y(t)$ are* amplitude indepen- dent, *if their joint ADF is the product of the ADF for each process. In other*

words,

$$F_{X,Y}(a, b) = F_X(a)F_Y(b). \tag{4.10}$$

We now will show a relationship between the ADF of stationary Gaussian noise and its instantaneous probability distribution, thus enabling the establishment of more directly applicable results.

Lemma 3 *Let $Y(t)$ be a stationary, zero-mean, Gaussian process with auto-correlation $R(t)$, such that $-R''(0) < \infty$; then, for any measurable set A,*

$$\mathcal{E}\{\mathcal{L}[t : Y(t) \leq a, \ t \in A]\} = \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A \tag{4.11}$$

where Φ is the distribution function of a standard Gaussian random variable and $\sigma_0 = \sqrt{R(0)}$.

The proof for Lemma 3 is included in Appendix 4.A.

In words, the above lemma means that the average time that the noise process is below the threshold a on the set A is equal to the percent of time that the noise process is below a at any single point times the length of the set A. This result implies that the ADF of stationary Gaussian noise is identical to its instantaneous distribution.

With the help of the previous results, we now can prove amplitude independence between a deterministic signal and stationary Gaussian noise, whose

autocorrelation has finite second derivative at time difference zero. This result will be necessary to prove the convolutional relationship between signal and additive noise.

Theorem 2 *Let $S(t)$ be a deterministic signal and let $N(t)$ be a stationary, zero-mean Gaussian process with autocorrelation $R(t)$ such that $-R''(0) < \infty$; then $S(t)$ and $N(t)$ are amplitude independent. Stated symbolically,*

$$F_{S,N}(a,b) = F_S(a)F_N(b). \tag{4.12}$$

Proof

By Lemma 3,

$$\mathcal{E}\left\{\mathcal{L}\left[t : S(t) \le a \text{ and } N(t) \le b,\ 0 \le t < T\right]\right\}$$

$$= \mathcal{E}\left\{\mathcal{L}\left[t : N(t) \le b,\ t \in S_a^{-1} \cap [0, T)\right]\right\} \tag{4.13}$$

$$= F_N(b)\mathcal{L}\left[S_a^{-1} \cap [0, T)\right] \tag{4.14}$$

where $S_a^{-1} = [t : S(t) \le a]$. Hence the joint ADF of $S(t)$ and $N(t)$ is

$$F_{S,N}(a,b) = F_N(b)\lim_{T \to \infty} \frac{1}{T}\mathcal{L}\left[S_a^{-1} \cap [0, T)\right] \tag{4.15}$$

$$= F_S(a)F_N(b) \tag{4.16}$$

implying the amplitude independence of $S(t)$ and $N(t)$.

We now can prove the most important result of this section, that of convolution between the ADFs of signal and additive noise. The idea of deconvolution and the deconvolution statistic rest firmly on this result.

Theorem 3 *Let $S(t)$ and $N(t)$ be amplitude independent and let either F_N or F_S be continuous; then the ADF of $Y(t) = S(t) + N(t)$ is*

$$F_Y(a) = \int_{-\infty}^{\infty} F_N(a - s) \, dF_S^*(s) \tag{4.17}$$

$$= \int_{-\infty}^{\infty} F_S(a - n) \, dF_N^*(n) \tag{4.18}$$

where F_S^* and F_N^* are the right-continuous extensions of F_S and F_N.

The proof for Theorem 3 is given in Appendix 4.B.

The next theorem is a general statement about the ADF of signal with additive noise and its instantaneous distribution. Alternatively, this theorem could have served as the definition of the ADF, but then there would be technical difficulties in determining the ADFs of purely deterministic signals.

Theorem 4 *Let $Y(t) = S(t) + N(t)$ and let F_S be right-continuous; then, if either F_N or F_S is continuous, the ADF of $Y(t)$ is*

$$F_Y(a) = \lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \int_0^T F_N[a - S(t)] \right\} dt. \tag{4.19}$$

The proof for Theorem 4 is included in Appendix 4.C.

4.3   Deconvolution Statistic

As shown in the previous section, the ADF of signal plus noise is the ADF of the signal convolved with the ADF of the noise. What was exactly shown is that

for $X(t) = S(t) + N(t)$ where $N(t)$ is a stationary Gaussian process with autocorrelation $R(t)$ satisfying $-R''(0) < \infty$ and where $S(t)$ is a deterministic signal with defined ADF; then $F_X = F_S * F_N$ where F_X, F_S, and F_N are the respective ADFs of $X(t)$, $S(t)$, and $N(t)$.

To apply this result in the construction of a detector, we will assume from herein that the above restrictions are met and that the noise, signal, and observations have densities defined as $f_X \triangleq dF_X(x)/dx$, $f_S \triangleq df_S/da$, and $f_N \triangleq dF_N/da$. These densities will be called the amplitude density functions (adf). We will, for reasons explained later, make the restriction that the noise autocorrelation is zero after some duration (i.e. $R(t) = 0$, for t greater than some T_1). This is a sufficient but probably not necessary condition for the asymptotic statistical characterization of the detection statistic.

If the adf of the observed signal could somehow be measured or estimated, then by deconvolution the signal component could be separated from the noise

142

component and hence detected. An understanding of deconvolution is prereq-uisite to understanding of how the deconvolution statistic effects this separa-tion.

In general to deconvolve, take a function that is the convolution of two different functions and convolve it again with a kernel function whose net effect is to undo the first convolution. As such, convolution of the kernel function with the original convolving function should result in a delta function. The construction of a kernel function with this property is usually not possible exactly and so some approximation must be made.

To use the idea of deconvolution in a detection scheme, the adf could be measured using standard density estimation techniques and then convolved with a kernel function to separate signal from noise. But there is a more direct approach, that of the deconvolution statistic

$$\hat{f}_X(x) = \frac{1}{T} \int_0^T k[x - X(t)] \, dt \tag{4.20}$$

where $k(x)$ is the deconvolution kernel.

The usefulness of this statistic is that, as T gets large, it converges uni-formly in probability to the desired convolution of the kernel function with the adf of the observations. In a sense, the statistic maps the observations into another domain, the amplitude domain. The problem of detection is now one on this new domain and from herein, it is approached in a classical manner.

4.3.1 Deconvolution Kernel

We want to construct a kernel with the property that when it is convolved with the noise adf, a function closely approximating the delta function results. The approach is to solve the equivalent problem after Fourier transformation, that is, to find a function that when multiplied by the Fourier transform of the noise adf yields a constant. It is shown below that this can be done only approximately.

Proceeding by Theorems 3 and 4, the noise has adf

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{1}{2}\frac{x^2}{\sigma_0^2}} \tag{4.21}$$

where $\sigma_0^2 = R(0)$. The Fourier transform of the noise adf is

$$C_N(\omega) \triangleq \int_{-\infty}^{\infty} f_N(x) e^{j2\pi\omega x}\, dx = e^{-\frac{1}{2}\sigma_0^2\omega^2}. \tag{4.22}$$

From the convolutional relationship between noise and signal, the transform of the observation adf is $C_X(\omega) = C_N(\omega)C_S(\omega)$. From this relationship, we might be led to construct a kernel with characteristic function $1/C_N(\omega)$, but this proves fruitless since the inverse Fourier transform does not exist. Instead, if before Fourier inversion, we multiply $1/C_N(\omega)$ with window function

$$\cos^2\left(\frac{d\omega}{2}\right), \quad -\frac{\pi}{d} \le \omega \le \frac{\pi}{d}, \quad d > 0 \tag{4.23}$$

we then get a family of kernels $k_d(x)$ indexed by d with the property that

$k_d * f_N(x) \to \delta(x)$ as $d \to 0$, hence convolution with $k_d(x)$ can approximate, to an arbitrary accuracy, perfect deconvolution.

Application of the above strategy with equations (4.22) and (4.23) produces

$$k_d(x) = \frac{1}{\pi} \int_0^{\pi} e^{\frac{1}{2}\left(\frac{\theta}{d}\right)} \cos^2 \left(\frac{d}{2}\right) \cos \left(\frac{x\theta}{d}\right) \, d\theta. \qquad (4.24)$$

The quantity d, which controls the amount of deconvolution, is called the deconvolution index.

4.3.2 Statistical Characterization

In order to set up the detection problem, we need at least an asymptotic statistical characterization of the deconvolution statistic. Specifically, we want to show that samples of the deconvolution statistic are asymptotically jointly Gaussian for large T, and we want its asymptotic mean and variance.

As for the jointly Gaussian property, consider samples of the deconvolution statistic

$$z_i = \frac{1}{T} \int_0^T k\left[x_i - X(t)\right] \, dt \qquad (4.25)$$

for some finite sequence $\{x_i\}_{i=1}^n$. In order to prove that the z_i's are jointly Gaussian, it is sufficient to prove that $\sum_i^n c_i z_i$ is Gaussian for arbitrary constants c_i. Rewrite

$$\sum_i^n c_i z_i = \sum_{j=1}^l I_{2j} + \sum_{j=1}^m I_{2j-1} \qquad (4.26)$$

145

where

$$I_k = \frac{1}{T} \int_{(k-1)T_1}^{\max(T, kT_1)} \sum_{i=1}^{n} c_i k \left[x_i - X(t) \right] \, dt \qquad (4.27)$$

$$l = \left[\frac{T}{2T_1} \right] \qquad (4.28)$$

$$m = \left[\frac{T}{2T_1} + \frac{1}{2} \right]. \qquad (4.29)$$

It is easily seen that the I_{2j}'s are independent and the I_{2j-1}'s are independent. Hence, for large T, each sum, $\sum_{j=1}^{l} I_{2j}$ and $\sum_{j=1}^{m} I_{2j-1}$, is individually asymptotically Gaussian. Furthermore, even though the sums are correlated, the overall sum is approximately Gaussian, since each component sum is Gaussian. Hence, samples of the deconvolution statistic are jointly Gaussian.

As for the mean,

$$\mu(x) \triangleq \mathcal{E} \left[\hat{f}_X(x) \right] \qquad (4.30)$$

$$= \frac{1}{T} \mathcal{E} \left\{ \int_0^T k \left[x - X(t) \right] \right\} \qquad (4.31)$$

$$\approx k * f_X(x) \quad \text{for large } T \qquad (4.32)$$

with the last step following from Theorems 3 and 4.

As for the variance,

$$\sigma^2(x, y) \triangleq \mathcal{E} \{ [\hat{f}_X(x) - \mu(x)][\hat{f}_X(y) - \mu(y)] \} \qquad (4.33)$$

146

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(x-v)k(y-w)h(v,w,T)\, dvdw - \mu(x)\mu(y)$$

$$(4.34)$$

where

$$h(v,w,T) = \frac{1}{T^2} \int_0^T \int_0^T \frac{1}{2\pi\sigma_0^2\Delta} e^{-\frac{1}{2\sigma_0^2\Delta} \boldsymbol{M}^T \boldsymbol{\Sigma} \boldsymbol{M}}\, dsdt \qquad (4.35)$$

$$(4.36)$$

$$\Delta = [1 - \rho^2(s-t)]$$

$$\boldsymbol{M} = \begin{bmatrix} v - S(s) \\ w - S(t) \end{bmatrix} \qquad (4.37)$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -\rho(s-t) \\ -\rho(s-t) & 1 \end{bmatrix} \qquad (4.38)$$

$$(4.39)$$

$$\rho(s-t) = \frac{R(s-t)}{\sigma_0^2}$$

We want to show that the variance goes to zero as the integration time T goes to infinity from whence it follows that the detection statistic converges, in probability, to its mean. To this end, assume $T > T_1$ and rewrite

$$h(v,w,T) = \frac{1}{T^2} \int\int_{D_1} \frac{1}{2\pi\sigma_0^2\Delta} e^{-\frac{1}{2\sigma_0^2\Delta} \boldsymbol{M}^T \boldsymbol{\Sigma} \boldsymbol{M}}\, dsdt \qquad (4.40)$$

$$+ \frac{1}{T^2} \int\int_{D_2} \frac{1}{2\pi\sigma_0^2\Delta} e^{-\frac{1}{2\sigma_0^2\Delta} \boldsymbol{M}^T \boldsymbol{\Sigma} \boldsymbol{M}}\, dsdt \qquad (4.41)$$

where the regions within the square, $Q = [s, t : 0 \le s \le T, 0 \le t \le T]$, are

$$D_1 = [s, t : s - t \le T_1] \cap Q \qquad (4.42)$$

147

$$D_2 = [s, t : s - t > T_1] \cap Q. \tag{4.43}$$

The area of D_1 is less than $2TT_1$ and, for $\sigma_0^2 > 0$, the integrand is bounded by $1/2\pi\sigma_0^2$. Furthermore $\Delta = 1$ in region D_2. Hence,

$$h(v, w, T) \le \frac{2T_1}{T} + \frac{1}{2\pi\sigma_0^2 T^2} \int_0^T e^{-\frac{1}{2\sigma_0^2}[v - S(s)]^2} ds \int_0^T e^{-\frac{1}{2\sigma_0^2}[w - S(t)]^2} dt. \tag{4.44}$$

Now assuming $k(x)$ is integrable, we have from (4.44) and (4.32) that $\sigma^2(x, y) \to 0$ for large T, as conjectured.

The fact that samples of the deconvolution statistic are asymptotically jointly Gaussian, and the expressions of its mean and variance will be necessary for the detector development, below.

4.4   ADF-Based Detector

Using the fact that samples of detection statistic are asymptotically jointly Gaussian, a classical detector can be constructed that observes the amplitude domain. We assume the original observed waveform to be of the form $X(t) = \sqrt{2S'} \sin \omega_0 t + n(t)$, for $0 \le t \le T$, where $n(t)$ is white noise of spectral density $N_0/2$, S' is the average signal energy, and $\omega_0/2\pi$ is unknown frequency in the band $[f_c - W/2, f_c + W/2]$. In this setting, the detection problem is one of choosing between H_0 (signal absent) and H_1 (signal present with SNR, $\gamma' = S'T/N_0 \ge \gamma = ST/N_0$). Before we apply our detector, we filter the observations with a bandpass filter to produce the waveform $Y(t)$ (see
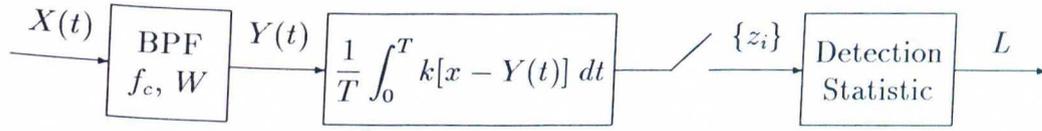
Figure 4.5: ADF-Based Detector

Figure 4.5). This filter has unity gain within the bandwidth W and has center frequency f_c. It further has response such the noise has autocorrelation $R(t)$ with $R(t) = 0$ for t greater than some T_1.[1] We next transform the output of the filter into the amplitude domain via the deconvolution statistic. The new detection problem becomes, after sampling, one of deciding between the presence and absence of a signal given $\{z_i = \hat{f}_Y(x_i)\}_{i=1}^n$ As shown earlier, the z_i's are jointly Gaussian and have means (4.32),

$$\mu_{\gamma'}(x_i) = \begin{cases} p(x_i), & \text{signal absent} \\ p * r(x_i), & \text{signal present} \end{cases} \tag{4.45}$$

where

$$p(x) = \frac{1}{\sqrt{2\pi N_0 W}} e^{-\frac{1}{2}\frac{x^2}{N_0 W}} \tag{4.46}$$

$$r(x) = \frac{1}{\pi}\sqrt{2S' - x^2}, \quad x^2 < 2S' \tag{4.47}$$

and covariances $\sigma_{\gamma'}(x_i, x_j)$ defined by (4.34) with

$$R(t) = N_0 W (1 - Wt), \quad -1 < Wt < 1 \tag{4.48}$$

$$S(t) = \sqrt{2S'} \sin \omega t. \tag{4.49}$$

---

[1]This filter is not realizable exactly but can be approximated to an arbitrary degree.

The $R(t)$ used here is a first approximation to the one described earlier. Because of this, there will be a small amplitude variation in the signal with respect to frequency that will be ignored in the analysis to follow. The above analysis completely characterizes the asymptotic joint statistics of samples of the deconvolution statistic.

We now use classical methods to design a detector on the amplitude domain. The philosophy assumes that both the noise and signal amplitudes are known. Then under this condition, an optimum detector is synthesized. To relax this unrealistic assumption, we will assume that the noise level is known or measured and that the signal level is above that used in the detector's synthesis. This will be suboptimum in general but in the important low-signal-level case, performance will approach the optimum.

We can now, via the likelihood ratio [47], define the optimal test statistic as

$$L = \boldsymbol{E}_0^T \boldsymbol{R}_0^{-1} \boldsymbol{E}_0 - \boldsymbol{E}_\gamma^T \boldsymbol{R}_\gamma^{-1} \boldsymbol{E}_\gamma \qquad (4.50)$$

where

$$\boldsymbol{R}_\gamma = \begin{bmatrix} \sigma_\gamma(x_1, x_1) & \cdots & \sigma_\gamma(x_1, x_n) \\ \vdots & & \vdots \\ \sigma_\gamma(x_n, x_1) & \cdots & \sigma_\gamma(x_n, x_n) \end{bmatrix} \qquad (4.51)$$

and

$$\boldsymbol{E}_\gamma = \begin{bmatrix} z_i - \mu_\gamma(x_1) \\ \vdots \\ z_i - \mu_\gamma(x_n) \end{bmatrix}. \qquad (4.52)$$

In order to evaluate the performance of this detector, we note that, since R_γ is nonnegative-definite and symmetric, there exists a matrix $R_\gamma^{\frac{1}{2}}$ such that $R_\gamma^{\frac{T}{2}} R_\gamma^{\frac{1}{2}} = R_\gamma$. Use this fact and the diagonalization of $R_\gamma^{-\frac{T}{2}} R^{-1} R_\gamma^{-\frac{1}{2}} = T_{\gamma',\gamma}^T \Lambda_{\gamma',\gamma} T_{\gamma',\gamma}$ to rewrite the test statistic as

$$L = (G + M_{\gamma',0})^T \Lambda_{\gamma',0} (G + M_{\gamma',0}) - (G + M_{\gamma',\gamma})^T \Lambda_{\gamma',\gamma} (G + M_{\gamma',\gamma}) \quad (4.53)$$

where

$$G = \begin{bmatrix} g_1 \\ \vdots \\ g_2 \end{bmatrix} \quad (4.54)$$

with $\{g_i\}$ independent, zero mean, and unity variance and where

$$M_{\gamma',\gamma} \triangleq \begin{bmatrix} m_{1,\gamma',\gamma} \\ \vdots \\ m_{n,\gamma',\gamma} \end{bmatrix} = T_{\gamma',\gamma}^T R_{\gamma'}^{-\frac{1}{2}} \begin{bmatrix} \mu_{\gamma'}(x_1) - \mu_\gamma(x_1) \\ \vdots \\ \mu_{\gamma'}(x_n) - \mu_\gamma(x_n) \end{bmatrix}. \quad (4.55)$$

We note that $(G + M_{\gamma',\gamma})^T \Lambda_{\gamma',\gamma} (G + M_{\gamma',\gamma})$ is a sum of squares of independent Gaussian variables. We now have, through application of the Berry-Esseén Theorem (see [25]), that this term is approximately Gaussian distributed with an error of no more than $4c/\sigma$ where

$$c = \max_i \lambda_{i,\gamma',\gamma} \frac{24 m_{i,\gamma',\gamma}^2 + 7}{4 m_{i,\gamma',\gamma}^2 + 2} \quad (4.56)$$

$$\sigma^2 = \sum_{i=1}^n \lambda_{i,\gamma',\gamma}^2 (4 m_{i,\gamma',\gamma}^2 + 2). \quad (4.57)$$

For each particular detection problem and for each value of SNR, this error bound determines the validity of the CLT argument. Assuming the bound is small, then L itself must be approximately Gaussian, being the sum of two

Gaussian distributed random variables. The distribution and hence performance of L is determined by its mean and variance as computed below. The mean is

$$\mathcal{M}_{\gamma',\gamma} = \sum_{i=1}^{n} \left[\lambda_{i,\gamma',0}(m_{i,\gamma',0}^2 + 1) - \lambda_{i,\gamma',\gamma}(m_{i,\gamma',\gamma}^2 + 1) \right] \quad (4.58)$$

while the variance is

$$\mathcal{V}_{\gamma',\gamma} = \sum_{i=1}^{n} \left[2 \left(\lambda_{i,\gamma',0}^2 - \lambda_{i,\gamma',\gamma}^2 \right)^2 + 4 \left(m_{i,\gamma',0}\lambda_{i,\gamma',0} - m_{i,\gamma',\gamma}\lambda_{i,\gamma',\gamma} \right)^2 \right].$$

$$(4.59)$$

Since the test statistic L has an approximately Gaussian distribution, the threshold v and probability of detection P_D, for a given probability of false alarm, follow as

$$v = \sqrt{\mathcal{V}_{0,\gamma}}\,\Phi^{-1}(1 - P_F) + \mathcal{M}_{0,\gamma} \quad (4.60)$$

and

$$P_D = 1 - \Phi \left(\frac{\sqrt{\mathcal{V}_{0,\gamma}}\,\Phi^{-1}(1 - P_F) - \mathcal{M}_{\gamma',\gamma} + \mathcal{M}_{0,\gamma}}{\sqrt{\mathcal{V}_{\gamma',\gamma}}} \right) \quad (4.61)$$

where $\Phi(x)$ is the distribution function of the standard Gaussian.

4.5 Conclusions

The ADF was introduced and shown to be roughly the average probability distribution of a random process. Because the ADF is original, a mathematical

foundation was laid consisting of a sequence of definitions, lemmas, and theorems, the most significant of which was the fact that the ADF of signal plus noise is the convolution of the ADF of signal and the ADF of noise taken separately. The technique of deconvolution used in image processing was the germ from whence emerged the deconvolution statistic, a statistical transform that mapped a stochastic process into the amplitude domain. It was shown that by proper choice of a kernel any degree of separation between the noise and signal components could be obtained. For the particular problem of detecting a modulated sinusoid in stationary Gaussian noise, a detector was developed around the detection statistic. The detector's performance was analyzed and compared with that of a radiometer.

The use of the ADF in detection has room for many new developments in just its mathematical development, not to mention specific applications in detection theory. For instance, the basic results possibly could be extended to non-Gaussian and nonstationary noise. In terms of the deconvolution statistic, the class of kernels used was very narrow. The investigation of various kernels, especially those used in image processing, would be in itself a worthwhile undertaking. Within the class of kernels presented, the choice of an optimal kernel for various classes of detection problems would be a possible topic to pursue. Finally, the most fertile ground for the application of the ADF to detection could be the detection of noise in noise.

4.A Proof of Lemma 3

We first consider the sets A, which are finite half open intervals. Of these, it is necessary to consider only the interval $[0, T)$, since $Y(t)$ is stationary. Partition $A = [0, T)$ into n subintervals

$$B_i = \left[(i-1)\frac{T}{n}, \ i\frac{T}{n} \right) \ \text{ for } i = 1, \ldots, n \tag{4.62}$$

and define the set $Y_a^{-1} = [t : Y(t) \leq a]$. Observing that the length of the set $Y_a^{-1} \cap A$ is the sum of the lengths of the sets $Y_a^{-1} \cap B_i$, we can write

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] = \sum_{i=1}^{n} \mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_i)\right] \tag{4.63}$$

$$= n\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_1)\right] \tag{4.64}$$

since $Y(t)$ is stationary. Now define the following three events:

C is the event that $Y(t) < a$, for some $t \in B_1$
D is the event that $Y(t) \leq a$, for all $t \in B_1$
E is the event that $Y(t)$ crosses a on B_1.

Notice that

$$\mathcal{L}(Y_a^{-1} \cap B_1) \geq \frac{T}{n} I_D \tag{4.65}$$

where I_D is the indicator of the event D. To understand this relationship, consider the case that the sample path $Y(t)$ is in D, meaning that it is not above a during the entire interval B_1. It follows that the amount of time that it is not above a, i.e. $\mathcal{L}(Y_a^{-1} \cap B_1)$, equals the length of B_1, which is T/n.

Upon taking expectations of this relationship,

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_1)\right] \geq \frac{T}{n}\mathcal{E}(I_D) \tag{4.66}$$

$$= \frac{T}{n}\Pr(D) \tag{4.67}$$

$$= \frac{T}{n}\{\Pr[Y(0) \leq a] - \Pr[E \text{ and } (Y(0) \leq a)]\} \tag{4.68}$$

since the probability that $Y(t)$ is not above level a over the interval B_1 is exactly the probability that $Y(0)$ is not above a and, under this condition, $Y(t)$ does not cross a. Now, since $\Pr[E \text{ and } (Y(0) \leq a)]$ is less than or equal to $\Pr(E)$, we have

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_1)\right] \geq \frac{T}{n}\{\Pr[Y(0) \leq a] - \Pr(E)\}. \tag{4.69}$$

In an analogous manner, we can produce a complementary inequality by observing that

$$\mathcal{L}(Y_a^{-1} \cap B_1) \leq \frac{T}{n}I_C. \tag{4.70}$$

This inequality follows by considering two cases. When the sample path $Y(t)$ is in C, meaning that it is below a sometime during the interval B_1, the amount of time that it is not above a is not greater than the length of B_1 or T/n. Alternatively, whenever $Y(t)$ is not in C, it is not above a for zero time. Taking expectations of this inequality yields

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_1)\right] \leq \frac{T}{n}\mathcal{E}(I_C) \tag{4.71}$$

155

$$= \frac{T}{n} \Pr(C) \tag{4.72}$$

$$= \frac{T}{n} \{\Pr[Y(0) \leq a] - \Pr[E \text{ and } (Y(0) \geq a)]\} \tag{4.73}$$

since the probability that $Y(t)$ is below level a for some time during the interval B_1 is exactly the probability that $Y(0)$ is not below a but crosses a during the interval B_1 plus the probability that $Y(0)$ is below a initially. Now, since $\Pr[E \text{ and } Y(0) \geq a]$ is less than or equal to $\Pr(E)$, we have

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap B_1)\right] \leq \frac{T}{n} \{\Pr[Y(0) \leq a] - \Pr(E)\}. \tag{4.74}$$

Equation (4.69) and (4.74) applied to (4.64) imply

$$T\{\Pr[Y(0) \leq a] - \Pr(E)\} \leq \mathcal{E}\left[\mathcal{L}\left(Y_a^{-1} \cap B_1\right)\right] \leq T\{\Pr[Y(0) \leq a] + \Pr(E)\} \tag{4.75}$$

thus

$$\mathcal{E}\left[\mathcal{L}\left(Y^{-1} \cap B_1\right)\right] = \Pr[Y(0) \leq a]T \tag{4.76}$$

$$= \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A \tag{4.77}$$

if $\lim_{n \to \infty} \Pr(E) \to 0$.

In order to prove that $\lim_{n \to \infty} \Pr(E) \to 0$, we define the counting process $N_a(t)$ as the number of crossings of the threshold a by the process $Y(t)$ on the interval $[0, t)$. By Chebyshev's inequality,

$$\Pr(E) = \Pr\left[N_a\left(\frac{T}{n}\right) \geq 1\right] \tag{4.78}$$

156

$$\leq \mathcal{E}\left[N_a\left(\frac{T}{n}\right)\right] \tag{4.79}$$

but from [16],

$$\mathcal{E}\left[N_a\left(\frac{T}{n}\right)\right] = \frac{T}{n}\frac{\alpha}{\sigma_0}e^{-\frac{a^2}{2\sigma_0}} \tag{4.80}$$

where $\alpha^2 = -R(0)''$. Upon letting $n \to \infty$, the last two equations imply $\Pr(E) \to 0$. Now that the result

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] = \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A \tag{4.81}$$

has been proven for A, an interval, it can be extended to any finite set as follows. Let A be a set of finite length; then for any $\epsilon > 0$ there exists a finite set of intervals $\{L_i\}_{i=1}^{l}$ such that

$$\mathcal{L}\left[A - \bigcup_{i=1}^{l} L_i\right] \leq \frac{\epsilon}{2} \tag{4.82}$$

and

$$\mathcal{L}\left[\bigcup_{i=1}^{l} L_i - A\right] \leq \frac{\epsilon}{2} \tag{4.83}$$

(see [35]). Hence,

$$\sum_{i=1}^{l}\mathcal{L}(Y_a^{-1} \cap L_i) - \frac{\epsilon}{2} \leq \mathcal{L}(Y_a^{-1} \cap A) \leq \sum_{i=1}^{l}\mathcal{L}(Y_a^{-1} \cap L_i) + \frac{\epsilon}{2}. \tag{4.84}$$

Taking expectations and applying the result for intervals

$$\Phi\left(\frac{a}{\sigma_0}\right)\sum_{i=1}^{l}\mathcal{L}L_i - \frac{\epsilon}{2} \leq \mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] \leq \Phi\left(\frac{a}{\sigma_0}\right)\sum_{i=1}^{l}\mathcal{L}L_i + \frac{\epsilon}{2}. \tag{4.85}$$

But

$$\mathcal{L}A - \frac{\epsilon}{2} \leq \sum_{i=1}^{l}\mathcal{L}L_i \leq \mathcal{L}A + \frac{\epsilon}{2} \tag{4.86}$$

157

implying

$$\Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A - \epsilon \leq \mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] \leq \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A + \epsilon. \tag{4.87}$$

But, since ϵ is an arbitrary positive number, we have for a set A with finite length,

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] = \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A. \tag{4.88}$$

To extend to the case in which A is not of finite length, write $A = \cup_{i=0}^{\infty} A_i$, where the A_i's are disjoint and of finite length; then

$$\mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A)\right] = \sum_{i=0}^{\infty} \mathcal{E}\left[\mathcal{L}(Y_a^{-1} \cap A_i)\right] \tag{4.89}$$

$$= \sum_{i=0}^{\infty} \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A_i \tag{4.90}$$

$$= \Phi\left(\frac{a}{\sigma_0}\right)\mathcal{L}A. \tag{4.91}$$

4.B   Proof of Theorem 3

Without loss of generality, assume F_S is continuous. To prove the result, we wish to compute

$$F_Y(a) = \lim_{T\to\infty} \frac{1}{T}\left\{\mathcal{L}\left[Y_a^{-1} \cap [0,T)\right]\right\} \tag{4.92}$$

where

$$Y_a^{-1} = [t : \ S(t) + N(t) \leq a]. \tag{4.93}$$

158

Begin by selecting an integer m and constructing a partition of the real line,

$(s_{-m^2} < s_{-m^2+1} < \cdots < s_{m^2})$, where

$$s_i = \begin{cases} -\infty & \text{for } i = -m^2 \\ i/m & \text{for } -m^2 < i < m^2 \\ \infty & \text{for } i = m^2. \end{cases} \qquad (4.94)$$

Continue by defining the sets

$$(4.95)$$

$$A_i^- = [t : N(t) \le a - s_i \text{ and } s_{i-1} < S(t) \le s_i]$$

$$(4.96)$$

$$A_i^+ = [t : N(t) \le a - s_{i-1} \text{ and } s_{i-1} \le S(t) < s_i]$$

for $i = -m^2 + 1, \cdots, m^2$. Observe that $A_i^- \subset Y_a^{-1}$, because $S(t) + N(t) \le a$

whenever $N(t) \le a - s_i$ and $s_{i-1} < S(t) \le s_i$. Hence,

$$(4.97)$$

$$\bigcup_{i=1}^{n} A_i^- \subset Y_a^{-1}.$$

Furthermore,

$$(4.98)$$

$$Y_a^{-1} \subset \bigcup_{i=1}^{n} A_i^+$$

because, for any t where $N(t) + S(t) \le a$, there exists an i in the range $-m^2 \le i \le m^2$, such that $s_{i-1} \le S(t) < s_i$ implying $N(t) \le a - s_i$. Because

$\{A_i^-\}_{i=-m^2+1}^{m^2}$ and $\{A_i^+\}_{i=-m^2+1}^{m^2}$ are disjoint, (4.97) and (4.98) imply that

$$\frac{1}{T}\mathcal{E}\left\{\mathcal{L}\left[Y_a^{-1} \cap [0, T)\right]\right\} \ge \sum_{i=-m^2+1}^{m^2} \frac{1}{T}\mathcal{E}\left\{\mathcal{L}\left[A_i^- \cap [0, T)\right]\right\} \qquad (4.99)$$

$$\frac{1}{T}\mathcal{E}\left\{\mathcal{L}\left[Y_a^{-1} \cap [0, T)\right]\right\} \le \sum_{i=-m^2+1}^{m^2} \frac{1}{T}\mathcal{E}\left\{\mathcal{L}\left[A_i^+ \cap [0, T)\right]\right\}. \qquad (4.100)$$

Now notice that

$$(4.101)$$

$$\lim_{T \to \infty} \frac{1}{T}\mathcal{E}\left\{\mathcal{L}\left[Y_a^{-1} \cap [0, T)\right]\right\} \triangleq F_Y(a).$$

159

Since $S(t)$ and $N(t)$ are amplitude independent, F_S is continuous, and $F_S(+\infty)$ $\overset{\Delta}{=} \lim_{s \to \infty} F_S(s)$ and $F_S(-\infty) \overset{\Delta}{=} \lim_{s \to -\infty} F_S(s)$, it follows that

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \mathcal{L} \left[A_i^- \cap [0, T) \right] \right\} = F_N(a - s_i) \left[F_S(s_i) - F_S(s_{i-1}) \right] \qquad (4.102)$$

Similarly,

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \mathcal{L} \left[A_i^+ \cap [0, T) \right] \right\} = F_N(a - s_{i-1}) \left[F_S(s_i) - F_S(s_{i-1}) \right]. \qquad (4.103)$$

By defining two particular step functions, the above results with equation (4.100) and (4.99) form a relationship between the ADF of $S(t) + N(t)$ and the integrals of the two step functions. The step functions are

$$F_N^-(a - s) \quad \overset{\Delta}{=} \quad F_N(a - s_i) \qquad (4.104)$$

$$F_N^+(a - s) \quad \overset{\Delta}{=} \quad F_N(a - s_{i-1}) \qquad (4.105)$$

whenever $s_{i-1} < s \le s_i$. Upon passing T to ∞, (4.100) and (4.99) become with the aid of the above definitions and (4.101), (4.102), and (4.103),

$$\int_{-\infty}^{\infty} F_N^-(a - s) \, dF_S(s) \le F_Y(a) \le \int_{-\infty}^{\infty} F_N^+(a - s) \, dF_S(s). \qquad (4.106)$$

Proceed by enlarging m and find that F_N^- and F_N^+ converge weakly to F_N which coupled with (4.106), implies that

$$F_Y(a) = \int_{-\infty}^{\infty} F_N(a - s) \, dF_S(s) \qquad (4.107)$$

$$= \int_{-\infty}^{\infty} F_N(a - s) \, dF_S^*(s) \qquad (4.108)$$

160

where the last equation, following from the continuity of F_S, proves the first convolution. Note that the interchange between limit and integration is justified in the last operation, because the integrals have essentially all mass within a bounded domain, upon which the integrand is bounded between zero and one. Note also that, if F_N is not continuous, then F_N^+ or F_N^- may converge to some function that differs from F_N at a countable number of points, but luckily the hypothesis that F_S is continuous makes (4.106) invariant to the limit value of F_N^- and F_N^+ at these problem points.

The complementary convolution is obtained by integrating the product measure $dF_S^*(s) \times dF_N^*(n)$ over the half plane $H \triangleq [s, n : s + n \leq a]$. Proceeding, we obtain with the help of Fubini's theorem

$$\int_H dF_S^*(s) \times dF_N^*(n) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{a-s} dF_N^*(n) \right] dF_S^*(s) \quad (4.109)$$

$$= \int_{-\infty}^{\infty} F_N^*(a - s) dF_S^*(s) \quad (4.110)$$

or, alternatively,

$$\int_H dF_S^*(s) \times dF_N^*(n) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{a-n} dF_S^*(s) \right] dF_N^*(n) \quad (4.111)$$

$$= \int_{-\infty}^{\infty} F_S^*(a - n) dF_N^*(n) \quad (4.112)$$

implying

$$\int_{-\infty}^{\infty} F_N^*(a - s) dF_S^*(s) = \int_{-\infty}^{\infty} F_S^*(a - n) dF_N^*(n). \quad (4.113)$$

The continuity of F_S and (4.108) yields

$$F_Y(a) = \int_{-\infty}^{\infty} F_N^*(a-s)\, dF_S^*(s) \tag{4.114}$$

which with (4.113)

$$F_Y(a) = \int_{-\infty}^{\infty} F_S^*(a-n)\, dF_N^*(n) \tag{4.115}$$

$$= \int_{-\infty}^{\infty} F_S(a-n)\, dF_N^*(n) \tag{4.116}$$

by the continuity of F_S. Equation (4.116), the remaining convolutional relationship, is now proven.

4.C   Proof of Theorem 4

Begin by selecting an integer m and constructing a partition of the real line $(s_{-m^2} < s_{-m^2+1} < \cdots < s_{m^2})$, where

$$s_i = \begin{cases} -\infty & \text{for } i = -m^2 \\ i/m & \text{for } -m^2 < i < m^2 \\ \infty & \text{for } i = m^2. \end{cases} \tag{4.117}$$

Continue by defining the sets

$$\Xi_i = [t : s_{i-1} < S(t) \le s_i] \tag{4.118}$$

for $i = -m^2 + 1, \cdots, m^2$. The sets $\{\Xi_i\}$ partition the real line into disjoint subsets, therefore,

$$\int_{(0,T]} F_N[a - S(t)]\, dt = \sum_{i=-m^2+1}^{m^2} \int_{\Xi_i \cap (0,T]} F_N[a - S(t)]\, dt. \tag{4.119}$$

162

Furthermore, since F_N is increasing,

$$\int_{(0,T]} F_N[a - S(t)] \, dt \geq \sum_{i=-m^2+1}^{m^2} \int_{\Xi_i \cap (0,T]} F_N(a - s_i) \, dt \qquad (4.120)$$

and

$$\int_{(0,T]} F_N[a - S(t)] \, dt \leq \sum_{i=-m^2+1}^{m^2} \int_{\Xi_i \cap (0,T]} F_N(a - s_{i-1}) \, dt. \qquad (4.121)$$

Divide (4.120) by T, take expectations, and enlarge T; then, noting that $F_S(+\infty) \triangleq \lim_{s \to \infty} F_S(s)$ and $F_S(-\infty) \triangleq \lim_{s \to -\infty} F_S(s)$ implies

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \int_0^T F_N[a - S(t)] \, dt \right\} \geq \sum_{i=-m^2+1}^{m^2} F_N(a - s_i) \left[F_S(s_i) - F_S(s_{i-1}) \right]. \qquad (4.122)$$

Operate similarly on (4.121) and

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \int_0^T F_N[a - S(t)] \, dt \right\} \leq \sum_{i=-m^2+1}^{m^2} F_N(a - s_{i-1}) \left[F_S(s_i) - F_S(s_{i-1}) \right]. \qquad (4.123)$$

In order to express (4.122) and (4.123) in convolutional form, define the two step functions

$$F_N^-(a - s) \triangleq F_N(a - s_i) \qquad (4.124)$$

$$F_N^+(a - s) \triangleq F_N(a - s_{i-1}) \qquad (4.125)$$

whenever $s_{i-1} < s \leq s_i$. Since F_S is assumed to be at least right-continuous, it follows from (4.122) and (4.123) that

$$\int_{-\infty}^{\infty} F_N^-(a - s) \, dF_S^*(s) \leq \lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \int_0^T F_N[a - S(t)] \, dt \right\} \qquad (4.126)$$

163

and hence

$$\int_{-\infty}^{\infty} F_N^-(a - s) \, dF_S^*(s) \leq \int_{-\infty}^{\infty} F_N^+(a - s) \, dF_S^*(s). \qquad (4.127)$$

Finish the proof by enlarging m and find that F_N^- and F_N^+ converge weakly to F_N which, coupled with (4.127), implies

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{E} \left\{ \int_0^T F_N[a - S(t)] \, dt \right\} = \int_{-\infty}^{\infty} F_N(a - s) \, dF_S^*(s) \qquad (4.128)$$

$$= F_Y(a) \qquad (4.129)$$

where the final step follows from Theorem 3. Note that the interchange between limit and integration is justified in the last operation, because the integrals have essentially all mass within a bounded domain upon which the integrand is bounded between zero and one. Note also that, if F_N is not continuous, then F_N^+ or F_N^- may converge to some function that differs from F_N at a countable number of points; but for this case, F_S is hypothesized to be continuous, making (4.128) invariant to the limit value of F_N^- and F_N^+ at these problem points.

Chapter 5

Concluding Remarks and Extensions

5.1   Summary and Remarks

In the first chapter, the general setting for the interception problem was described. A brief exposition on the function of the intercept receiver was given, emphasizing the functions of initial detection and feature detection that are pertinent to this work. A brief description of existing detectors was presented with a qualitative, and in some cases quantitative, performance evaluation. These were the radiometer, the optimal channelized detector with various suboptimal simplifications, and the autocorrelation detector. Of existing feature detectors of hop frequency, a maximum likelihood receiver was described along with a related receiver employing course subband selection. The detectors presented in this work were then briefly described and related to existing detectors. They were a sequential detector, two detectors and a hop frequency

estimator based on the compressive receiver, and another detector based on the new concept of the amplitude distribution function.

In the second chapter, methods for the sequential detection of noncoherent fast FH waveforms were developed. In the process, the FH waveform was modeled to have an information component, which consisted of a series of chips with a known constant epoch where each chip frequency was one of a known ensemble of frequencies. In the model, a particular chip frequency was independently determined by a uniform random variable on the frequency ensemble. The FH waveform also was assumed to have an additive white-noise component. By assuming the modeled FH waveform was of a known SNR, the optimal detector based on a single-epoch observation (SELF) was developed using likelihood-function theory. SELF was the sum of many nearly identical and nearly independent random variables and thus had nearly Gaussian statistics. This central-limit argument allowed a multi-epoch collection of SELFs to be considered an equivalent set of Gaussian i.i.d. variables. From these simplified observations, a log-likelihood function (ALLF) was computed that was asymptotic to the exact log-likelihood function, as the number of possible hop frequencies became large. The ALLF became the test statistic on which three detection tests were based. The tests were the FSS test, the SPRT, and the TST. These were defined to ensure that detection errors were below desired levels. By modeling the ALLF as a Wiener process, diffusion theory yielded

the performance of the three tests not only for an FH waveform of the assumed SNR, but also the test performance for all SNRs below the one assumed. This analysis compared favorably with a computer simulation of the detector and thus validated the analysis. The analysis also became a tool used to optimize numerically the performance of the TST when the actual FH SNR deviated from the one assumed. In order to study the performance of tests synthesized by assuming an extremely small FH SNR, expressions for the asymptotic test efficiencies were computed. This asymptotic analysis also yielded simplified test parameter expressions applicable to the small-SNR case.

A significant feature of the SPRT exposed by the analysis is that, with the same error probabilities, an FH waveform with a given SNR can be detected in less than half the time of the corresponding FSS test. This reduction in detection time is especially significant for Low Probability of Intercept (LPI) applications, where the transmissions are purposely short. For the pure SPRT, detection time increased whenever the observed SNR differed from that assumed in the test's synthesis. And for SNRs midway between zero and the assumed value, it was even comparable to the corresponding FSS test. The TST significantly improves this anomaly, while sacrificing little performance over that of the purely sequential test; whatever little performance is lost, the optimal TST largely regains. The decrease in the detection time of the sequential tests can be used to robustify the test with respect to the input SNR

while maintaining better performance than that of the non-robust FSS test. The simplified test parameter expressions derived by asymptotic methods may be useful for any schemes to adapt these tests for varying FH SNRs. The three tests and their corresponding design and performance analysis also apply to the slow-FH case. The detector structure is suboptimal for slow FH, but it is believed that the performance loss is small, especially for detection times that include a large number of hops.

In the third chapter, two detectors of frequency-hopped waveforms based on the compressive receiver were presented. The first was developed by applying likelihood ratio theory to the observed compressive-receiver output and yielded a locally optimal (low-SNR) detector. The second, motivated by simplicity of implementation, was a time-multiplexed version of the first that, through the choice of a parameter, could either, at the expense of a low duty cycle, achieve the detectability of the first or could, at the expense of degraded performance, achieve higher duty cycles. The second detector was modified into a maximum likelihood estimator of hop frequency. Both detectors and the hop frequency estimator were performance analyzed and compared.

The compressive receiver fulfilled its promise as a simple, yet high-performing interceptor. The performance of the locally optimal detector shows that relatively little detectability is lost by the processing of the compressive receiver. Most of the discrepancy is due to the difference in coherent integration time

(one half for the parameters used). Furthermore, for a small performance cost, the simplicity of the compressive-receiver approach can be retained by the time-multiplexed detector. Also, the hop frequency estimator again compares favorably with the corresponding device that used raw input instead of compressive-receiver output.

In the fourth chapter, a new idea in detection, that of the Amplitude Distribution Function (ADF) was introduced. The ADF is roughly the average probability distribution of a random process. Because the ADF is original, a mathematical foundation was laid consisting of a sequence of definitions, lemmas, and theorems, the most significant of which was the fact that the ADF of signal plus noise is the convolution of the ADF of signal and the ADF of noise taken separately. The technique of deconvolution used in image processing was the germ from which emerged the Deconvolution Statistic, a statistical transform that mapped a stochastic process into the amplitude domain. It was shown that by proper choice of a kernel any degree of separation between the noise and signal components could be obtained. For the particular problem of detecting a modulated sinusoid in stationary Gaussian noise, a detector was developed around the Detection Statistic. The detector's performance was analyzed and compared with a radiometer.

5.2   Extensions—Directions for Future Work

In applying sequential detection to interception, other simplifications and extensions are possible. For instance, it was assumed that the starting time and duration of the chip epoch were known. This first restriction might be relaxed by redefining the SELF to perform sliding window integration instead of the integrate-and-dump operation now performed. This, of course, would degrade the detector's performance for some values of epoch starting time, but it would probably exhibit a better average performance. There are also possible simplifications to the SELF to improve its implementability. Among these could be the removal of the emphasizing function, which would make the detector structure suboptimal but it probably still would be asymptotically optimal for small assumed SNRs. Another simplification could be coarse subband preselection, where the total spread-spectrum bandwidth is subdivided into subbands, each containing a large number of chip frequencies. An algorithm could be used to select a subset of the subbands most likely to contain the intercepted signal. Detailed processing on these preselected bands then could be done with the methods described in this work.

There are many remaining avenues to be traveled in the use of the compressive receiver to interception. We provided many results applying to frequency-hopped waveforms, but the essence of these ideas can apply to other spread-

spectrum modulations as well. But even in realm of frequency-hopped waveforms, much work remains to be done. For instance, a locally optimal detector was derived consisting of a bank of filters operating on the compressive-receiver output. From the filter responses in this configuration was formed the filter response for the time-multiplexed detector. The direct derivation of this filter response using some optimality criterion over the class of expected signals might be a promising endeavor. The extension of the detectors to a multihop observation period was largely ignored. We simply assumed that results of the individual detectors would be combined, as if they were independent. Here the issue of performance versus overlap between data windows is one to be potentially explored. We showed how the compressive receiver could be used to estimate hop frequency. Other types of feature detectors such as the hop-rate detectors or the carrier frequency detectors also could be pursued.

The use of the ADF in detection, being a new idea, has room for many new developments in just its mathematical development, not to mention specific applications in detection theory. For instance, the basic results could be extended to non-Gaussian and nonstationary noise. In terms of the Deconvolution Statistic, the class of kernels used was very narrow. The investigation of various kernels, especially those used in image processing, would be, by itself, a worthwhile undertaking. Within the class of kernels presented, the choice of an optimal kernel for various classes of detection problems is a possible area

to pursue. Finally, the most fertile ground for the application of the ADF to detection could be the detection of noise in noise.

5.3 Final Remarks

This work represents a significant step in the development of new detection and interception techniques for spread-spectrum waveforms and frequency-hopped waveforms in particular. With the three new detectors and the hop frequency estimator presented, many new trails were blazed. Armed with the idea of sequential detection, an existing optimal detector using a fixed number of samples was extended, with improved performance resulting. This useful and eclectic endeavor brought previous theoretical and practical results on sequential detection into the context of interception. The second detector and hop frequency estimator solved the problem of how to apply effectively the compressive receiver to interception. The firm mathematical development starkly contrasted with previous ad hoc attacks on the problem. Out of this work also came some mathematical results of general interest; among these were bounding distributions on the sum of squares of Gaussian random variables and an extension of the Riemann-Lebesgue Lemma to integrals of linearly frequency-modulated sinusoids. A brand new idea in detection was conceived, yielding yet another new detector. The original idea melded the image process-

ing technique of deconvolution to those of density estimation. The detector thus developed was quite independent of the details of signal modulation.

While there is much presented here, there is still much to be done. There are other interceptors to be analyzed and new detection techniques to be developed. As such, the area of spread-spectrum interception will yield new results for many years to come.

Bibliography

[1] T.W. Anderson, "A Modification of the Sequential Probability Ratio Test to Reduce the Sample Size", *Ann. Math. Stat.*, vol. 31, pp. 165–197, 1960.

[2] H.C. Andrews and B.R. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[3] N.C. Beaulieu, W.L. Hopkins, and P.J. McLane, "Interception of Frequency Hopped Spread Spectrum Signals", to be published in *IEEE Trans. Commun.*

[4] K.D. Breuer, J.J Whelehan, and K. Ross, "Compressive Receivers Applied to ESM System Design", *MSN Microwave Syst. News & Commun. Techno.*, vol. 16, no. 11, pp. 66–68, 70, 72, 74–75, Oct. 1986.

[5] T.W. Bristol, "Review of Spectrum Analysis with SAW Chirp Transforms and Filter Banks", *Proceedings of International Specialist Seminar on Case Studies in Advanced Signal Processing*, 1979, pp. 226–231.

[6] D.A. Darling and A.J.F. Siegert, "The First Passage Problem for a Continuous Markov Process", *Ann. Math. Stat.*, vol. 24, pp. 624–639, 1953.

[7] P. Das and L.B. Milstein, "Adaptive Spread Spectrum Receiver Using Acoustic Surface Wave Technology", Report No. ARO 17570-8-EL, May 31 1984, U.S. Army Research Office, Post Office Box 12211, Research Triangle Park, N.C. 27709.

[8] G.M. Dillard, "A Moving-Window Detector for Binary Integration", *IEEE Trans. Inform. Theory*, IT-13, no. 1, pp. 2–6, January 1967.

[9] R.A. Dillard, "Detectability of Spread Spectrum Signals", *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-15, no. 5, pp. 526–537, July 1979.

[10] J.D. Edell, "Wideband, Noncoherent, Frequency-Hopped Waveforms and their Hybrids in Low-Probability-of-Intercept Communications", Naval Research Laboratory, Washington, D.C., Technical Report No. 8025, November 1976.

[11] R. Fry and P. Yansouni, "Detection and Direction Finding of Frequency Hoppers by Amplitude Comparison Techniques", Defense Research Establishment, Ottawa, Ontario Canada, Report No. 940, December 1986.

[12] W.A. Gardner, "Signal Interception: A Unifying Theoretical Framework for Feature Detection", *IEEE Trans. Commun.*, vol. COM-36, pp. 897–906, 1988.

[13] A.B. Glenn, "Low Probability of Intercept", *IEEE Commun. Mag.*, vol. 21, no. 4, pp. 26–33, July 1983.

[14] B.K. Harms and D.R. Hummels, "Calculation of Detection Probability for Frequency Compressive Receivers", *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-21, pp. 106–116, January 1985.

[15] J.H. Higbie, "Adaptive Nonlinear Suppression of Interference", *MILCOM '88 Conference Proceedings*, 23.3.1–23.3.9, October 1988.

[16] S. Karlin and H.M Taylor, *A First Course in Stochastic Processes*, Second Edition, Academic Press, San Diego, Calif., 1975.

[17] L. Kleinrock, "Detection of Energy Peak of an Arbitrary Signal", MIT Lincoln Laboratory, Lexington, Mass., Technical Report No. 325, August 23, 1963.

[18] N.F. Krasner, "Optimal Detection of Digitally Modulated Signals", *IEEE Trans. Commun.*, vol. COM-30, no. 5, pp. 885–895, May 1982.

[19] N.F. Krasner, "Maximum Likelihood Parameter Estimation for LPI Signals", *MILCOM '82 Conference Proceedings*, Boston, Mass., pp. 2.3-1–2.3-4, October 17–20, 1982.

[20] K.H. Li and L.B. Milstein, "On the Use of a Compressive Receiver for Signal Detection", to be published in *IEEE Trans. Commun.*

[21] R.J. Mammone, "Frequency Resolution Enhancement of a Compressive Receiver by Spectral Estimation", *Proceedings of the IEEE Military Communications Conference*, vol. 3, pp. 713–19, 1983.

[22] G.L. Moule, "SAW Compressive Receivers for Radar Intercept", *IEE Proceedings, Part F — Communications, Radar and Signal Processing*, vol. 129, pt. F, no. 3, pp. 180–186, June 1982.

[23] D.L. Nicholson, "Spread Spectrum Signal Design", Computer Science, Rockville, Md., 1988.

[24] J.E. Ohlson, "Efficiency of Radiometers Using Digital Integration", *Radio Science*, vol. 6, no. 3, pp. 341–345, March 1971.

[25] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Second Edition, McGraw-Hill, New York, 1984.

[26] R.F. Pawula, S.O. Rice, and J.H. Roberts, "Distribution of the Phase Angle Between Two Vectors Perturbed by Gaussian Noise", *IEEE Trans. Commun.*, vol. COM-30, pp. 1828–1840, August 1982.

[27] M. Peligrad, "Recent Advances in the Central Limit Theorem and Its Weak Invariance Principle for Mixing Sequences of Random Variables (A Survey)", *Dependence in Probability and Statistics*, Birkhauser, p.193, 1985.

[28] A. Polydoros and C.L. Weber, "Optimal Detection Considerations for Low Probability of Intercept", *MILCOM '82 Conference Proceedings*, Boston, Mass., pp. 2.1-1–2.1-5, October 17–20, 1982.

[29] A. Polydoros and J.K. Holmes, "Autocorrelation Techniques for Wide-band Detection of FH/DS Waveforms in Random Tone Interference", *MILCOM '83 Conference Proceedings*, Boston, Mass., pp. 781–785, October 17–20, 1983.

[30] A. Polydoros, J.K. Holmes, and K.T. Woo, "Advanced LPI Intercept Detector Research", Technical Report No. R8403-1, Axiomatix, Los Angeles, Calif., March 26, 1984.

[31] A. Polydoros and K.T. Woo, "LPI Detection of Frequency-Hopping Signals using Autocorrelation Techniques", *IEEE J. Select. Areas Commun.*, vol. SAC-3, no. 5, September 1985.

[32] A. Polydoros, J.K. Holmes and K.T. Woo, "Advanced LPI Intercept Detector Research", Technical Report No. R8511-3, Axiomatix, Los Angeles, Calif., November 13, 1985.

[33] A. Polydoros and C.L. Nikias, "Detection of Unknown Frequency Sinusoids in Noise: Spectral versus Correlation Domain", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 6, June 1987.

[34] S.O. Rice, "Statistical Properties of a Sine Wave Plus Random Noise," *Bell Syst. Tech. J.*, pp. 109–157, January 1948.

[35] H.L. Royden, *Real Analysis*, MacMillan, New York, 1968.

[36] J. Salz, and S. Stein, "Distribution of Instantaneous Frequency for Signal Plus Noise", *IEEE Trans. Inform. Theory*, vol. IT-10, pp. 272–274, October 1964.

[37] D.C. Schleher, *Introduction to Electronic Warfare*, Artech House, Inc., Dedham, Mass., 1986.

[38] N. Shiryayev, *Optimal Stopping Rules*, Springer-Verlag, New York, 1977.

[39] W.W. Short and R.D. Chapman, "Adaptively Configured Channelized Receiver for Frequency Hopped Signal Detection and Tracking", *IEEE International Conference on Communications, 1985*, pp. 832–838, 1985.

[40] M.K. Simon, J.K. Omura, R.A. Scholtz, and B.K. Levitt, *Spread Spectrum Communications*, vol III, Computer Sciences Press, Rockville, Md., 1985.

[41] W.E. Snelling and E.A. Geraniotis, "Sequential Detection of Unknown, Fast Frequency-Hopped Waveforms", *IEEE J. Select. Areas Commun.*, vol. 7, no. 4, pp. 602–617, May 1989.

[42] W.E. Snelling and E.A. Geraniotis, "The Optimal Interception of Frequency-Hopped Waveforms via a Compressive Receiver", to be published in *IEEE Trans. Commun.*.

[43] S. Tantaratana and H.V. Poor, "Asymptotic Efficiencies of Truncated Sequential Tests", *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 911–923, November 1982.

[44] S. Tantaratana and J.B. Thomas, "Truncated Sequential Probability Ratio Test", *Inform. Sci.*, vol. 13, pp. 283–300, 1977.

[45] D. Torrieri, *Principles of Secure Communication Systems*. Artech House, Dedham, Mass., 1985.

[46] J.B.Y. Tsui, "Dual Detection Scheme for Compressive Receivers", U.S. Patent 4,200,840, April 29 1980.

[47] H.L. van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.

[48] A. Wald, *Sequential Analysis*, Wiley, New York, 1947.

[49] G.N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, New York, 1980.

[50] D.G. Woodring, "Performance of Optimum and Suboptimum Detectors for Spread Spectrum Waveforms", Naval Research Laboratory, Washington, D.C., Technical Report No. 8432, December 1980.