Fundamental limits on detection in low SNR

by

Rahul Tandra

B.tech. (Indian Institute of Technology, Bombay) 2003

A thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science
in

Engineering - Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Anant Sahai, Chair
Professor Micheal Gastpar

Spring 2005

The dissertation of Rahul Tandra is approved.

Chair   Date

                                                                      Date

University of California, Berkeley

Spring 2005

Fundamental limits on detection in low SNR

Abstract

Fundamental limits on detection in low SNR

by

Rahul Tandra

Master of Science in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Anant Sahai, Chair

In this thesis we consider the generic problem of detecting the presence or absence of a weak signal in a noisy environment and derive fundamental bounds on the sample complexity of detection. Our primary motivation for considering the problem of low SNR detection is the idea of cognitive radios, where the central problem of the secondary user is to detect whether a frequency band is being used by a known primary user.

By means of examples, we show that the general problem of detecting the presence or absence of a weak signal is very hard, especially when the signal has no deterministic component to it. In particular, we show that the sample complexity of the optimal detector is $O(1/SNR^2)$, even when we assume that the noise statistics are completely known.

More importantly we explore the situation when the noise statistics are no longer completely known to the receiver, i.e., the noise is assumed to be white, but we know its distribution only to within a particular class of distributions. We then propose a minimalist model for the uncertainty in the noise statistics and derive fundamental bounds on detection performance in low SNR in the presence of *noise uncertainty*. For clarity of analysis, we focus on detection of signals that are BPSK-modulated random data without any pilot tones or training sequences. The results should all generalize to more general signal constellations as long as no deterministic component is present.

Specifically, we show that for every 'moment detector' there exists an SNR below which detection becomes impossible in the presence of noise uncertainty. In the neighborhood of that SNR wall, we show how the sample complexity of detection approaches infinity. We also show that if our radio has a *finite dynamic range* (upper and lower limits to the voltages we can quantize), then at low enough SNR, *any detector* can be rendered useless even under moderate noise uncertainty.

Professor Anant Sahai
Dissertation Committee Chair

Dedicated to my parents.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my advisor, Prof. Anant Sahai for his encouragement and inspirational guidance. The time spent in the several meetings with my advisor has been a great learning experience for me. I would also like to thank Prof. Micheal Gastpar, Prof. David Tse, Prof. Kannan Ramchandran and the students in the spectrum sharing group, from whom I have learned a lot during the 3R meetings.

Chapter 1

Introduction

The subject of signal detection and estimation deals with the processing of information-bearing signals in order to make inferences about the information that they contain. This field traces its origin from the classical work of Bayes [1], Gauss [10], Fisher [7], and Neyman and Pearson [21], on statistical inference. However, the theory of detection took hold as a recognizable discipline only from the early 1940's. Presumably, signal detection has a wide range of applications including communications. The work of Price and Abramson [22], is one of the early efforts highlighting the importance of signal detection in communications and a comprehensive survey of the results and techniques of detection theory developed in the past decades can be found in [15].

In this thesis we are interested in the problem of *detecting the presence or absence* of a signal in a noisy environment. Throughout this thesis, whenever we refer to 'signal detection', we mean detecting the presence of a signal in background noise. We mainly concentrate on two types of signal detection problems, i) detecting known signals in additive noise (coherent detection), and ii) random signals in additive noise (non-coherent detection). For example, receivers which try to pickup a known pilot signal/training data sent by a transmitter use coherent detection. In this case, the receiver knows the exact signal and hence the optimal detector is just a *matched filter* [19]. On the other hand, one of the most common example of a non-coherent detector is an *energy detector (radiometer)*. It is known that the radiometer is the optimal detector when the receiver knows only the power of the signal [25], [29].

As the title of the thesis suggests, we are mainly interested in the problem of detecting very low SNR signals in additive noise. Our interest in the very low SNR regime is motivated by the possibility for cognitive radios [6], [20]. *Cognitive radio* refers to wireless architectures in which a communication system does not operate in a fixed assigned band, but rather searches and finds an appropriate band in which to operate. This represents a new paradigm for spectrum utilization in which new devices can opportunistically scavenge bands that are not being used at their current time and location for their primary purpose [5]. This is inspired by actual measurements showing that most of the allocated spectrum is vastly underutilized [2].

One of the most important requirement for any cognitive radio system is to provide a guarantee that it would not interfere with the primary transmission. In order to provide such a guarantee it is obvious that a cognitive radio system should be able to detect the presence of the primary signal, to which it might be severely shadowed. This is a version of the hidden terminal problem in which the primary system might have a receiver vulnerable to secondary interference while simultaneously the primary transmissions are shadowed en route to the secondary user (see fig. 1.1)

However, fundamental theoretical questions remain as to the exact requirements for engineering a practical cognitive radio system so that they do not interfere with the primary users. An introduction to the tradeoffs and challenges faced by cognitive radios can be found in [23]. In particular, in order to guarantee non-interference with primary users without being restricted to very low transmit powers, the cognitive radio system needs to be able to detect the presence of very weak primary signals, [11], [12].

Detecting very weak signals introduces many new challenges. In this thesis we will highlight some of the fundamental limitations on detection in low SNR. In case of coherent detection, it is well known that the sample complexity[1] of the matched filter behaves like

[1]The variation of the number of samples required as a function of the signal to noise ratio in order to guarantee a target probability of false alarm and missed detection

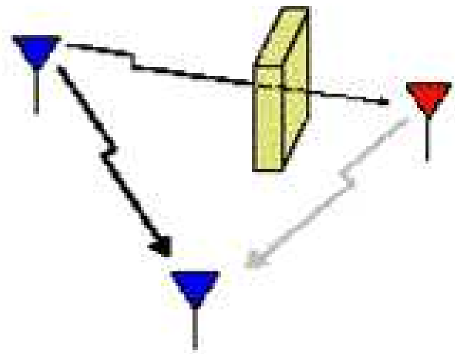


Figure 1.1. Instance of the hidden terminal problem. The blue antennas correspond to the primary system and the red antenna corresponds to the cognitive radio that is shadowed with respect to the primary transmitter.

$O(1/SNR)$, when the SNR is low. Whereas, detection becomes considerably harder when the signal has no deterministic component to it (non-coherent detection). For example, we consider the problem of detecting a random unknown BPSK modulated signal with very low power in additive Gaussian noise. It is know that if we use a radiometer to detect this signal then the sample complexity of the radiometer is $O(1/SNR^2)$, [17]. Furthermore, we show that the optimal detector also behaves as badly as the radiometer. Thus, non-coherent detection increases the sample complexity by a factor of $(1/SNR)$ which can be a serious computational cost on the design of the detector. Furthermore, additional number of samples implies that the receiver has to wait even longer to detect the signal. However, most stationarity assumptions (channel modeling assumptions) are valid only within small durations. Hence, this restricts the utility of the detector to very slowly varying environments. Further, in the context of cognitive radios, this additional delay might make the time required for detection more than the idle time of the primary transmitter, due to which the radiometer might be rendered useless for cognitive radios.

More importantly, there is another fundamental issue that needs to be considered while

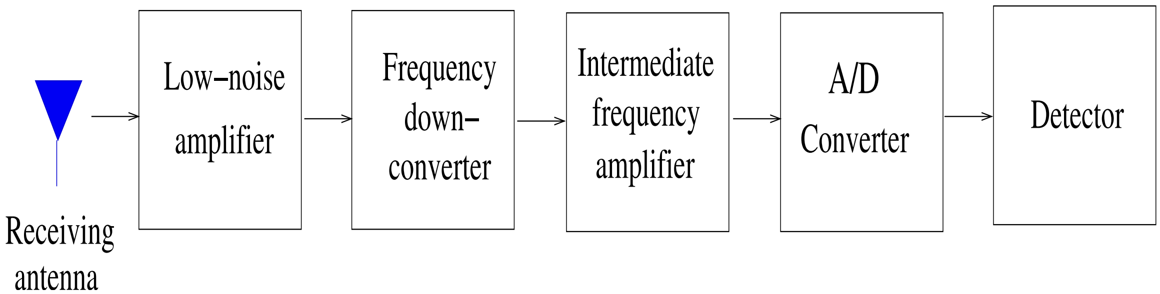


Figure 1.2. Abstraction of the detection problem

detecting signals in low SNR's. The classical Bayes, minimax and Neyman-Pearson design strategies for optimum signal detection and other decision strategies require a complete statistical description of the data in order to specify the optimum decision rule structure. However, there is frequently some uncertainty concerning the statistical structure of the data, for example the additive Gaussian noise assumption in our problem is only an approximation. It has been demonstrated that procedures designed around a particular model may perform poorly when actual data statistics differ from those assumed. For example, the performance of optimum procedures designed under the assumption of Gaussian distributions are often considerably degraded for slight deviations from the Gaussian assumption [28]. Furthermore, we claim that the degradation in performance is worsened in the low SNR regime.

To deal with situations mentioned above it is often of interest to find decision procedures which are *robust*, that is, which perform well despite small variations from the assumed statistical model. There were several studies which considered the design of robust detectors. The statistical work of Huber in estimation [13] and hypothesis testing [14] are generally regarded as the starting point of the area of minimax robustness. Huber considered the problem of robustly testing a simple hypothesis $\mathcal{H}_0$ against a simple alternative $\mathcal{H}_1$, assuming that the true underlying densities q_0 and q_1 lie in some neighborhood of the idealized model densities p_0 and p_1, respectively, and that they are related by

$$q_i = (1 - \epsilon_i)p_i(x) + \epsilon_i h_i(x), \qquad i = 0, 1 \quad h_i \in H \tag{1.1}$$

where $0 \leq \epsilon_i < 1$ are fixed numbers and H denotes the class of all density functions.

As Huber showed, this problem has a solution (existence of a robust detection procedure), when the ϵ_i are small enough, so that it is impossible for the distribution classes defined by (1.1) to overlap.

This work has been successfully applied to a long sequel of problems in detection and estimation, e.g., the works of Martin and Schwartz [18], Kassam and Thomas [16], and El-Sawy and Vandelinde [4]. Specifically, Martin and Schwartz consider the problem of detecting a signal of known form in additive, nearly Gaussian noise, and they derive a robust detector when the signal amplitude is known and the nearly Gaussian noise is specified by Huber's mixture model.

It is important to note that as the previous works mentioned above have considered the problem of detecting the presence or absence of known signals under noise uncertainty, i.e., they try to derive a *robust coherent detector*. However, the model considered by Huber is more general and can be applied to non-coherent detection also. In this thesis, we propose a similar model for the noise uncertainty, which can be shown to be a special case of the model given in (1.1). However, since we are interested in detection in low SNR, it is not reasonable to assume that the distribution classes under both the hypotheses don't overlap. In fact, we show that under sufficiently low SNR, the distribution classes must overlap, which in turn leads to serious limitations on the prospect of robust detection. Sonnenschein and Fishman [26], analyze the affect of noise power uncertainty in radiometric detection of spread-spectrum signals and identify a fundamental limit on the SNR of the signals we can detect. We extend their results beyond the energy detector to detectors that examine other moments.

Furthermore, we show that the fundamental limits on moment detectors become hard limits on *any possible detector* if the radio has a finite dynamic range on its input. This result says that under the finite dynamic range assumption it is impossible to find any robust detector below a certain SNR threshold. Intuitively, this result arises because, under the finite dynamic range assumption and in low SNR, the distribution classes under both hypothesis must overlap for any possible detector, i.e., it is possible for a noise distribution in each hypothesis to mimic the nominal distribution in the other hypothesis, thus leading to impossibility in signal detection. Since all physical radios have such a limit, we feel that the bounds here represent important constraints on practical systems.

In the next chapter, we formulate the detection problem as a hypothesis testing problem and discuss the sample complexity of various detectors, under the additive Gaussian noise assumption. In this chapter we assume that the noise statistics are completely known to the receiver. Chapter 3 relaxes this constraint and we build a model for the noise uncertainty. Under our model we show that for every moment detector there exists a SNR below which detection becomes impossible. We extend this impossibility result for a generic detector in chapter 4, where we make the practical assumption that our receivers are limited to a finite dynamic range of operation. Finally, we conclude by giving a brief discussion of the signal detection problem when the signal is no longer white. We give some arguments which suggest that even robust detectors (like feature detectors) might have to face the practical limitations of noise uncertainty. However, we feel that the limits in this case arise due to the uncertainty in the whiteness assumption of noise (in practice noise is also never white!). Therefore, we feel that these limitations are very fundamental in nature and systems which work under the very low SNR regime (like cognitive radios) cannot escape such limitations.

Chapter 2

Detection in Gaussian noise

2.1 Introduction

We start off by formulating the general problem of detecting the presence or absence of a signal in additive noise as a binary hypothesis testing problem. In the most general formulation we allow for the noise process to lie in a class of distributions around the nominal distribution.

However, as a preliminary discussion to the general problem we first assume that the noise statistics is completely known to the receiver. We assume that the noise is white Gaussian, and is independent of the signal. Under these assumptions we derive the sample complexities of optimal detectors[1] for various classes of signals and compare them with the sample complexity of the corresponding matched filter and radiometer.

Before we begin our discussion of signal detection, we first point out the critical differences between detection and demodulation. By demodulation, we mean the process of decoding a message based on the received signal. Shannon in his landmark paper [24], showed that it is possible to reliably demodulate the signal only if the rate of transmission is below a certain threshold, what we now call the Shannon capacity of the channel. The

[1]Note that the notion of optimality is in the classical Neyman-Pearson sense, i.e., a detector is optimal if it minimizes the number of samples required for detection subject to given target probabilities of false alarm and missed detection constraints

main problem in demodulation is to decode the transmitted message from the received (possibly corrupted) signal. However, we are already assuming the presence of a transmission in demodulation, i.e., the receiver knows that there is a primary signal present and all it has to do is to figure out what it is.

On the other hand, detecting signals at very low SNR is the main focus of this thesis. In particular we assume that the SNR is so low that we are operating below the capacity of the channel. Thus demodulating the signal is not possible, and we can only hope to detect the signal. In our problem the important case is detecting the absence of the primary signal. This is a hard problem, since there is no easy way to prove the absence of a signal. The only possible way to do this is to contradict the fact that there is a primary transmission present. However, contradicting the presence of a signal is hard when the signal power is very low. Therefore, the problem of detection is considerably harder than demodulation.

2.2 Problem formulation

The problem of signal detection in additive noise can be formulated as a binary hypothesis testing problem with the following hypotheses:

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] && n = 1, \ldots, N \\ \mathcal{H}_1 : Y[n] &= X[n] + W[n] && n = 1, \ldots, N \end{aligned} \tag{2.1}$$

where $X[n]$ and $W[n]$ are the signal and noise samples respectively. We assume that both the signal and noise processes are white and are independent of each other. However, we allow for the noise process to have any distribution from a class of noise distributions $\mathcal{W}$. Without loss of generality we can assume that $\mathcal{W}$ is centered around a nominal noise distribution W_n.

Given a particular target probability of false alarm P_{FA} and probability of missed detection P_{MD}, our aim is to derive the sample complexities for various possible detectors, i.e., we are interested in calculating the number of samples required N, as a function of the signal to noise ratio, which is defined as

$$snr = \frac{\mathbb{E}X_1^2}{\mathbb{E}W_n^2} \qquad\qquad SNR = 10\log_{10}\left(\frac{\mathbb{E}X_1^2}{\mathbb{E}W_n^2}\right) \tag{2.2}$$

For notational convenience, an upper case SNR is used to denote the signal to noise ratio in decibels. Also, since the noise distribution can lie within the class $\mathcal{W}$, we define snr with respect to the nominal noise distribution W_n as given in (2.2). In particular we are interested in the sample complexity of the optimal detector, i.e., the detector which minimizes N, for a given P_{FA} and P_{MD}.

2.3 Sample complexity of signal detection under WGN

Now, we consider the following hypothesis testing problem:

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] \qquad\qquad n = 0, 1, \ldots, N-1 \\ \mathcal{H}_1 : Y[n] &= X[n] + W[n] \qquad n = 0, 1, \ldots, N-1 \end{aligned} \tag{2.3}$$

Here we assume that $W[n]$'s are samples from a white Gaussian noise process with power spectral density σ^2, i.e., $W[n] \sim N(0, \sigma^2)$. This is a special case of the general problem in (2.1) with no uncertainty in the noise process and its statistics are completely know to the receiver.

We restrict ourselves to this model throughout this section. We start with the most simplest version of the problem: the case when the signal $X[n]$ is completely known to the receiver. Then, we make the problem progressively harder by dropping some of the simplifying assumptions made above and derive the sample complexity of the optimal detector in each case.

2.3.1 Matched filter

Here we assume that the signal samples $X[n]$ in (2.3) are completely known to the receiver. In this case, the optimal detector is

$$T(\mathbf{Y}) = \sum_{n=0}^{N-1} Y[n]X[n] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

This detector is referred to as a *correlator* or the *matched filter*. It can easily be shown that [17],

$$\begin{aligned} N &= [Q^{-1}(P_D) - Q^{-1}(P_{FA})]^2 (snr)^{-1} \\ &= O(snr^{-1}) \end{aligned} \tag{2.4}$$

Thus we see that the number of samples required varies as $O(1/snr)$.

2.3.2 Radiometer

Now, we consider a completely opposite problem to the one in section 2.3.1. Here we assume absolutely no deterministic knowledge about the signal $X[n]$, i.e., we assume that we know only the average power in the signal.

In this case the optimal detector is

$$T(\mathbf{Y}) = \sum_{n=0}^{N-1} Y^2[n] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

that is, we computes the energy in the signal and compare it to a threshold. Hence it is known as an *energy detector* or *radiometer*. In this case, it is easy to verify that [17],

$$\begin{aligned} N &= 2\left[(Q^{-1}(P_{FA}) - Q^{-1}(P_D))snr^{-1} - Q^{-1}(P_D)\right]^2 \\ &= O(snr^{-2}) \end{aligned} \tag{2.5}$$

Thus the number of samples, N varies as $O(1/snr^2)$ for the energy detector.

2.3.3 Sinusoidal signals

Now we assume that the signal $X[n] = A\cos(2\pi f_0 n + \phi)$ in (2.3), i.e., we consider the following detection problem:

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] & n = 0, 1, \ldots, N-1 \\ \mathcal{H}_1 : Y[n] &= A\cos(2\pi f_0 n + \phi) + W[n] & n = 0, 1, \ldots, N-1 \end{aligned} \tag{2.6}$$

Here we consider the cases when any subset of the parameter set $\{A, f_0, \phi\}$ may be unknown. We assume that these parameters are fixed but unknown in (2.6) and hence use the Generalized Likelihood Ratio Test (GLRT) (see [17]) for the following cases.

1. A unknown

2. A, ϕ unknown

3. A, ϕ, f_0 unknown

We solve each of the following cases in order.

Amplitude unknown

Here we assume that A is unknown (2.6). The GLRT decides $\mathcal{H}_1$ if,

$$\frac{f(\mathbf{Y}, \hat{A}|\mathcal{H}_1)}{f(\mathbf{Y}|\mathcal{H}_0)} > \gamma$$

where $\hat{A}$ is the Maximum Likelihood Estimate (MLE) of A under $\mathcal{H}_1$, i.e.,

$$\begin{aligned}
\hat{A} &= \arg\max_A f(\mathbf{Y}, A|\mathcal{H}_1) \\
&= \arg\max_A \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(Y[n] - AX[n])^2\right] \\
&= \arg\min_A \sum_{n=0}^{N-1}(Y[n] - AX[n])^2 \\
&= \frac{\sum_{n=0}^{N-1} Y[n]X[n]}{\sum_{n=0}^{N-1} X^2[n]}
\end{aligned}$$

Using this $\hat{A}$ to calculate the likelihood ratio we have the following detection rule

$$T(\mathbf{Y}) = \left(\sum_{n=0}^{N-1} Y[n]X[n]\right)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

The performance of this detector can be easily calculated and we get

$$P_D = Q(Q^{-1}\left(\frac{P_{FA}}{2}\right) - \sqrt{N \cdot snr}) + Q(Q^{-1}\left(\frac{P_{FA}}{2}\right) + \sqrt{N \cdot snr}) \qquad (2.7)$$

It can be observed from the above equation that for a fixed P_{FA} and P_D, the number of samples N varies as snr^{-1}. Thus even when the amplitude is unknown the sample complexity varies as $O(1/snr)$.

Unknown Amplitude and phase

Now, we assume that both A and ϕ are unknown in (2.6). As before we use the GLRT and decide upon $\mathcal{H}_1$ if

$$L_G(\mathbf{Y}) = \frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}|\mathcal{H}_1)}{f(\mathbf{Y}|\mathcal{H}_0)} > \gamma \tag{2.8}$$

where $\hat{A}, \hat{\phi}$ are the MLE's of A and ϕ under $\mathcal{H}_1$ respectively, i.e.,

$$\begin{aligned}
(\hat{A}, \hat{\phi}) &= \arg\max_{(A,\phi)} f(\mathbf{Y}, A, \phi|\mathcal{H}_1) \\
&= \arg\max_{(A,\phi)} \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(Y[n] - A\cos(2\pi f_0 n + \phi))^2\right] \\
&= \arg\min_{(A,\phi)} \sum_{n=0}^{N-1}(Y[n] - A\cos(2\pi f_0 n + \phi))^2 \\
&= \arg\min_{(\alpha_1,\alpha_2)} \sum_{n=0}^{N-1}(Y[n] - \alpha_1\cos(2\pi f_0 n) - \alpha_2\sin(2\pi f_0 n))^2
\end{aligned}$$

Where $\alpha_1 = A\cos(\phi)$ and $\alpha_2 = -A\sin(\phi)$. We can easily optimize the above expression and obtain the MLE's of A and ϕ to be

$$\begin{aligned}
\hat{A} &= \sqrt{\hat{\alpha_1} + \hat{\alpha_2}} \\
\hat{\phi} &= \arctan(-\frac{\hat{\alpha_2}}{\hat{\alpha_1}})
\end{aligned}$$

where,

$$\begin{aligned}
\hat{\alpha_1} &\approx \frac{2}{N}\sum_{n=0}^{N-1} Y[n]\cos(2\pi f_0 n) \\
\hat{\alpha_2} &\approx \frac{2}{N}\sum_{n=0}^{N-1} Y[n]\sin(2\pi f_0 n)
\end{aligned}$$

for f_0 not near 0 or $\frac{1}{2}$. Note that we can choose any $f_0 \in (0, \frac{1}{2})$ by sampling the original continuous time sinusoid at a suitable frequency greater than the Nyquist frequency.

Substituting the MLE's in (2.8) and taking logarithm we get

$$\ln L_G(\mathbf{Y}) = -\frac{1}{2\sigma^2}\left[\sum_{n=0}^{N-1} -2Y[n]\hat{A}\cos(2\pi f_0 n + \hat{\phi}) + \sum_{n=0}^{N-1} \hat{A}^2\cos^2(2\pi f_0 n + \hat{\phi})\right]$$

Using the parameter transformation we have $\hat{\alpha_1} = \hat{A}\cos(\hat{\phi})$, $\hat{\alpha_2} = \hat{A}\sin(\hat{\phi})$ so that

$$\begin{aligned}
& \sum_{n=0}^{N-1} Y[n]\hat{A}\cos(2\pi f_0 n + \hat{\phi}) \\
= & \sum_{n=0}^{N-1} Y[n]\cos(2\pi f_0 n)\hat{A}\cos\hat{\alpha_1} - \sum_{n=0}^{N-1} Y[n]\sin(2\pi f_0 n)\hat{A}\sin\hat{\alpha_1} \\
= & \frac{N}{2}(\hat{\alpha_1}^2 + \hat{\alpha_2}^2)
\end{aligned}$$

Also, making use of

$$\sum_{n=0}^{N-1} \cos^2(2\pi f_0 n + \hat{\phi}) \approx \frac{N}{2}$$

results in

$$\begin{aligned}
\ln L_G(\mathbf{x}) &= -\frac{1}{2\sigma^2}\left[-2\frac{N}{2}(\hat{\alpha_1}^2 + \hat{\alpha_2}^2) + \frac{N}{2}\hat{A}^2\right] \\
&= -\frac{1}{2\sigma^2}\left[-\frac{N}{2}(\hat{\alpha_1}^2 + \hat{\alpha_2}^2)\right] \\
&= \frac{N}{4\sigma^2}(\hat{\alpha_1}^2 + \hat{\alpha_2}^2) \qquad (2.9)
\end{aligned}$$

or we decide $\mathcal{H}_1$ if

$$\begin{aligned}
& \frac{N}{4\sigma^2}(\hat{\alpha_1}^2 + \hat{\alpha_2}^2) > \ln\gamma \\
\Rightarrow \quad & \frac{N}{4\sigma^2}\left(\frac{2}{N}\right)^2\left[\left(\sum_{n=0}^{N-1} Y[n]\cos(2\pi f_0 n)\right)^2 + \left(\sum_{n=0}^{N-1} Y[n]\sin(2\pi f_0 n)\right)^2\right] > \ln\gamma \\
\Rightarrow \quad & \frac{1}{\sigma^2}\frac{1}{N}\left|\sum_{n=0}^{N-1} Y[n]\exp\left(-j2\pi f_0 n\right)\right|^2 > \ln\gamma \\
\Rightarrow \quad & \frac{1}{\sigma^2}I(f_0) > \ln\gamma \qquad (2.10)
\end{aligned}$$

where $I(f_0)$ is the *periodogram* evaluated at $f = f_0$. Thus, the sample complexity in this case turns out to be

$$P_D = Q_{\mathcal{X}'^2_2(\lambda)}\left(2\ln\frac{1}{P_{FA}}\right) \qquad (2.11)$$

where $\lambda = N \cdot snr$ and $Q_{\mathcal{X}'^2_2(\lambda)}$ is the tail probability of the standard chi-squared random variable with 2 degrees of freedom and noncentrality parameter λ. It is clear that $N \cdot snr$ is the only variable in the above equation for fixed P_{FA} and P_D. Thus, in this case also the sample complexity varies as $O(1/snr)$

Amplitude, phase and frequency unknown

Now, we assume that the frequency f_0 is also unknown in (2.6). Here we assume that the continuous time frequency of the signal f can lie anywhere from $(0, f_{max})$. However, by sampling at a greater rate than $2f_{max}$ we can also make sure that the discrete frequency f_0 lies in $(0, \frac{1}{2})$. Therefore there is only a bounded set of uncertainty in the frequency of the discrete time signal. Similarly assuming that the original continuous time signal is within a narrow band $(f - W, f + W)$ does not help too much, because the frequency of the discrete time signal will again lie in $(0, \frac{1}{2})$. To summarize, the amount of uncertainty in the frequency is not a real issue while detection discrete time signals. Now, we try to solve the detection problem assuming that $f_0 \in (0, \frac{1}{2})$.

When the frequency is also unknown in (2.6), the GLRT decides $\mathcal{H}_1$ if

$$\frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}, \hat{f}_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} > \gamma$$

or

$$\frac{\max_{f_0} f(\mathbf{Y}, \hat{A}, \hat{\phi}, f_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} > \gamma$$

Since the PDF under $\mathcal{H}_0$ does not depend on f_0 and is nonnegative, we have

$$\max_{f_0} \frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}, f_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} > \gamma$$

taking the logarithm in the above inequality, we have

$$\ln \max_{f_0} \frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}, f_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} > \ln \gamma$$

and by monotonicity of log we can write this as

$$\max_{f_0} \ln \frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}, f_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} > \ln \gamma$$

but from (2.10) we have

$$\ln \frac{f(\mathbf{Y}, \hat{A}, \hat{\phi}, f_0|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} = \frac{I(f_0)}{\sigma^2}$$

so finally we decide $\mathcal{H}_1$ if

$$\max_{f_0} I(f_0) > \sigma^2 \ln \gamma = \gamma'$$

Thus, this detector detects that the signal is present if the peak of the periodogram exceeds a threshold. Assuming that an N-point FFT is used to evaluate $I(f)$ and that the maximum is found over the frequencies $f_k = \frac{k}{N}$ for $k = 1, 2, \ldots, \frac{N}{2} - 1$, we have that

$$P_D = Q_{\mathcal{X}'^2_2(\lambda)}\left(2 \ln \frac{\frac{N}{2} - 1}{P_{FA}}\right) \tag{2.12}$$

Here we assume that the actual frequency f_0 is a multiple of $\frac{1}{N}$. Clearly there will be some loss if f_0 is not near a multiple of $\frac{1}{N}$.

Fig. 2.1 plots the sample complexities of all the examples discussed in this section. Observe that the radiometer and the matched filter have the best and the worst sample complexities respectively, while the rest of the cases lie strictly between them. Also, observe that the lack of knowledge of frequency of a sinusoid does not affect the sample complexity by much. The sample complexity is still quite close to $O(1/snr)$ as seen in the figure.

2.3.4 BPSK detection

We have shown that the radiometer is the worst case possible. This is expected because the radiometer assumes minimal knowledge about the signal, namely only the average power in the signal. The cost paid is an increase in the the sample complexity, which is $O(1/snr^2)$.

Now, we assume that the in addition to knowing the power of the signal, we also know that structure of the modulation scheme used by the transmitter. For concreteness, we assume that the signal is random BPSK modulated data. However, all these results continue to hold for any general zero-mean signal constellation with low signal amplitude.

Hence our detection problem is to distinguish between these hypotheses:

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] \qquad\qquad n = 1, \ldots, N \\ \mathcal{H}_1 : Y[n] &= X[n] + W[n] \quad n = 1, \ldots, N \end{aligned} \tag{2.13}$$

where $W[n]$ are independent, identically distributed $N(0, \sigma^2)$ random variables and $X[n]$ are independent, identically distributed Bernoulli$(\frac{1}{2})$ random variables taking values in

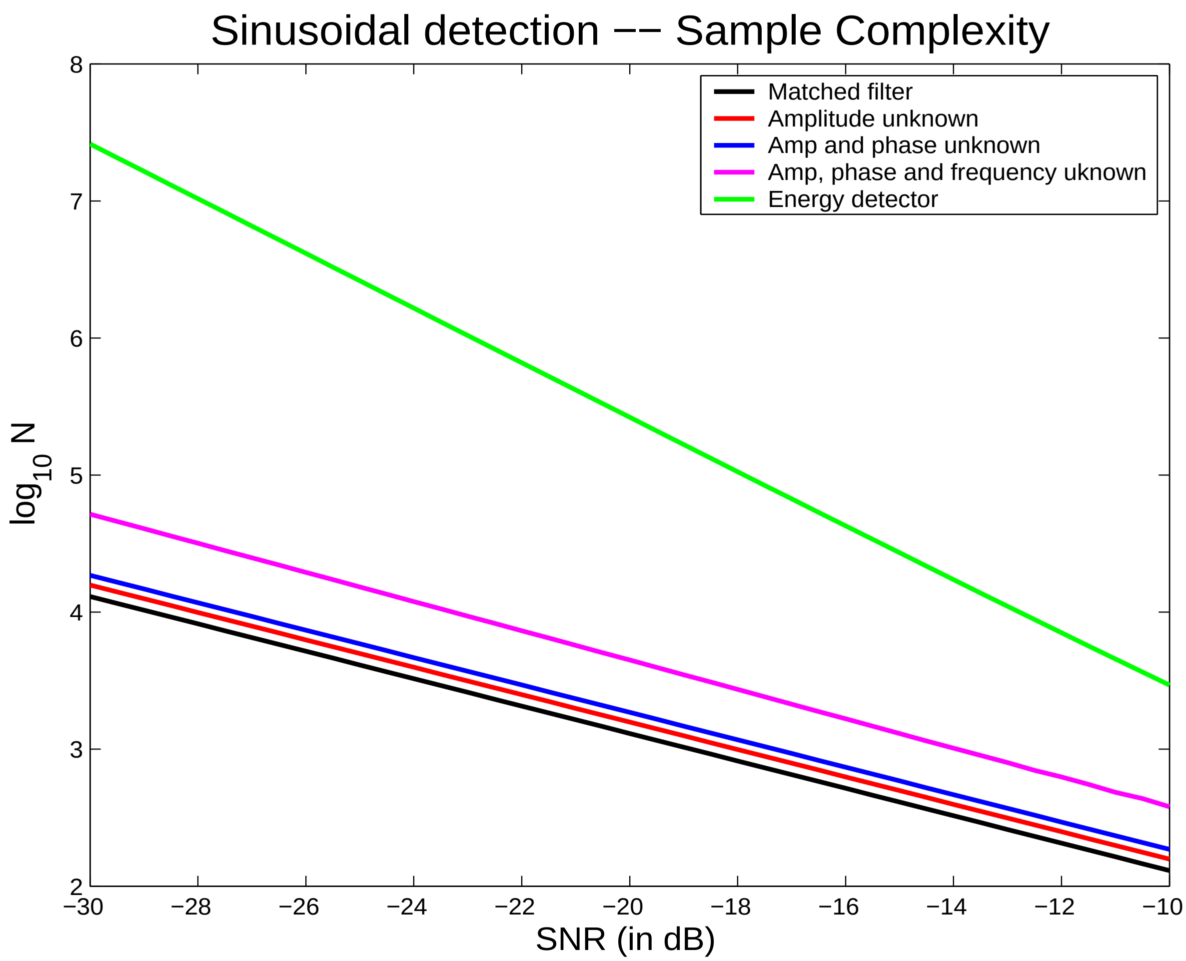


Figure 2.1. The sample complexities of the various detectors discussed in section 2.3 are plotted in this figure. These curves (read from top to bottom in the figure) were obtained by using the sample complexities derived in the equations (2.5), (2.12), (2.11), (2.7) and (2.4) respectively.

$\{-\sqrt{P}, \sqrt{P}\}$. Finding the log-likelihood test statistic is straightforward and yields:

$$T(\mathbf{Y}) = \sum_{n=1}^{N} \ln \left[\cosh \left(\frac{\sqrt{P}}{\sigma^2} Y[n] \right) \right] \tag{2.14}$$

Choosing a suitable threshold γ such that the target error probabilities are met, gives us the following detection rule: $T(\mathbf{Y}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$.

Since we are interested in the low SNR regime, the number of samples required is large. This implies that the test-statistic $T(\mathbf{Y})$ is a sum of a large number (N) of independent random variables. Thus, we can use the central limit theorem[2] (see [3]) to approximate the log-likelihood test statistic, $T(\mathbf{Y})$ as a Gaussian. This approximation helps us to get a closed form expression for the sample complexity of the optimal detector,

$$N = \left[\frac{Q^{-1}(P_{FA})\sqrt{\mathbb{V}((T(\mathbf{Y}))|\mathcal{H}_0)} - Q^{-1}(P_D)\sqrt{\mathbb{V}((T(\mathbf{Y}))|\mathcal{H}_1)}}{\mathbb{E}(T(\mathbf{Y})|\mathcal{H}_1) - \mathbb{E}(T(\mathbf{Y})|\mathcal{H}_0)} \right]^2 \tag{2.15}$$

where $\mathbb{V}(.)$ stands for the variance operator. In order to compute the sample complexity we need to evaluate, $\mathbb{E}(T(\mathbf{Y})|\mathcal{H}_i)$ and $\mathbb{V}((T(\mathbf{Y}))|\mathcal{H}_i)$ for $i = 0, 1$. Since, the snr is low, we can use Taylor series approximations and show that the sample complexity varies as $O(1/snr^2)$ (see (A.13)). The derivation of this result is not essential for the rest of the thesis, and hence it is postponed to appendix A.

Our shows that knowing the signal constellation does not improve the sample complexity. This result generalizes to any signal constellation with zero-mean and low signal amplitude.

Fig. 2.2 shows the sample complexity of the optimal detector derived in (2.14). This figure has been obtained from (2.15) by numerically evaluating the terms on the right hand side for different values of the snr. The green curve (labeled as 'Energy detector' in Fig. 2.2) is the sample complexity of the radiometer and the black curve (labeled as 'Deterministic BPSK' in Fig. 2.2) is the sample complexity of the matched filter, which is the optimal detector when the BPSK signal is completely known to the detector. Note that the optimal detector performs as badly as the energy detector.

[2]Since we are assuming that P_{FA} and P_{MD} are moderate, i.e., they are not dependent on N, the central limit theorem continues to apply. The situation is not like the probability of decoding error in a code where extremely tiny probabilities are targeted.

2.4 Pilot signals

In section 2.3.4 we considered the detection of signals which are zero-mean BPSK modulated, with no deterministic component. In this case we showed that the sample complexity behaves like $O(1/snr^2)$. This is bad news in the context of cognitive radios for reasons discussed in chapter 1. However, many communication schemes have training sequences which can act as weak pilot signals if their structure is perfectly known. "How does the sample complexity change in the presence of a weak pilot signal?".

2.4.1 BPSK signal with a weak pilot

In this section we consider the problem of detecting a random BPSK signal in the presence of a weak pilot tone. The detection problem in this case is to differentiate between the following hypotheses:

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] & n = 0, 1, \ldots, N-1 \\ \mathcal{H}_1 : Y[n] &= X[n] + W[n] & n = 0, 1, \ldots, N-1 \end{aligned}$$

where $X[n]$ are sample from a random BPSK modulated signal taking values in $\{d(1 + \theta), -d(1-\theta\}), d = \sqrt{\frac{P}{1+\theta^2}}$ and $W[n]$ is WGN with variance σ^2. As before, we assume that the noise is independent of the signal and $X[n] \sim$ Bernoulli(1/2) are independent, identically distributed random variables. Note that the signal $X[n]$ in this case has a non-zero mean given by $\mathbb{E}X[n] = d\theta$.

As in the pure BPSK case we can derive the sample complexity of the optimal detector. But we face the same problem here as before, i.e., the sample complexity does not have a closed form solution and hence we will have to resort to approximate methods again. Instead, we first propose a suboptimal detector and derive its sample complexity. The sample complexity of a sub-optimal detector will always give us an upper bound to the sample complexity of the optimal detector. Hence, if the performance of this sub-optimal detector is reasonable then we need not worry about solving for the sample complexity of the optimal detector. Thus, in this section we first derive the sample complexity of a sub-optimal detector and then numerically show that its sample complexity is close to that of the optimal detector in the low SNR regime.

Sub-optimal Detector

This detector decides $\mathcal{H}_1$ if

$$T(\mathbf{Y}) = \sum_{n=0}^{N-1} Y[n] > \gamma \tag{2.16}$$

Thus, this detector calculates the empirical mean of the samples and compares it to a threshold. The performance of this detector can easily be derived and is given by

$$N = \left[Q^{-1}(P_{FA}) - Q^{-1}(P_D)\right]^2 \left(\frac{\theta^2}{P}\right)^{-1} snr^{-1} \tag{2.17}$$

where P is the total signal power. Thus $\left(\frac{\theta^2}{P}\right)$ is the fraction of the total power that is allocated to the pilot signal. From(2.17) we observe that if,

$$\left(\frac{\theta^2}{P}\right) \approx snr$$

then this sub-optimal detector behaves like the energy detector and hence its sample complexity is a bad upper bound for the sample complexity of the optimal detector. But, if

$$\left(\frac{\theta^2}{P}\right) \gg snr$$

then $N = O(1/snr)$ and hence the sample complexity of this suboptimal detector has the same order as that of the matched filter in section 2.3.1. Thus, in this case the sub-optimal detector performance gives us a very tight bound to the optimal detector sample complexity. The plot of the sample complexity of this sub-optimal detector is compared to the sample complexities of other detectors in fig. 2.2. In particular note that this sub-optimal detector performs as well as the optimal detector in the low snr regime.

Thus we have shown that, by designing a suboptimal detector that just searches for known pilot signals, we can substantially reduce the number of samples required to detect when the snr is low (see fig. 2.2). Furthermore, these observations hold for general zero-mean signal constellations also, as long as the signal amplitude is low.

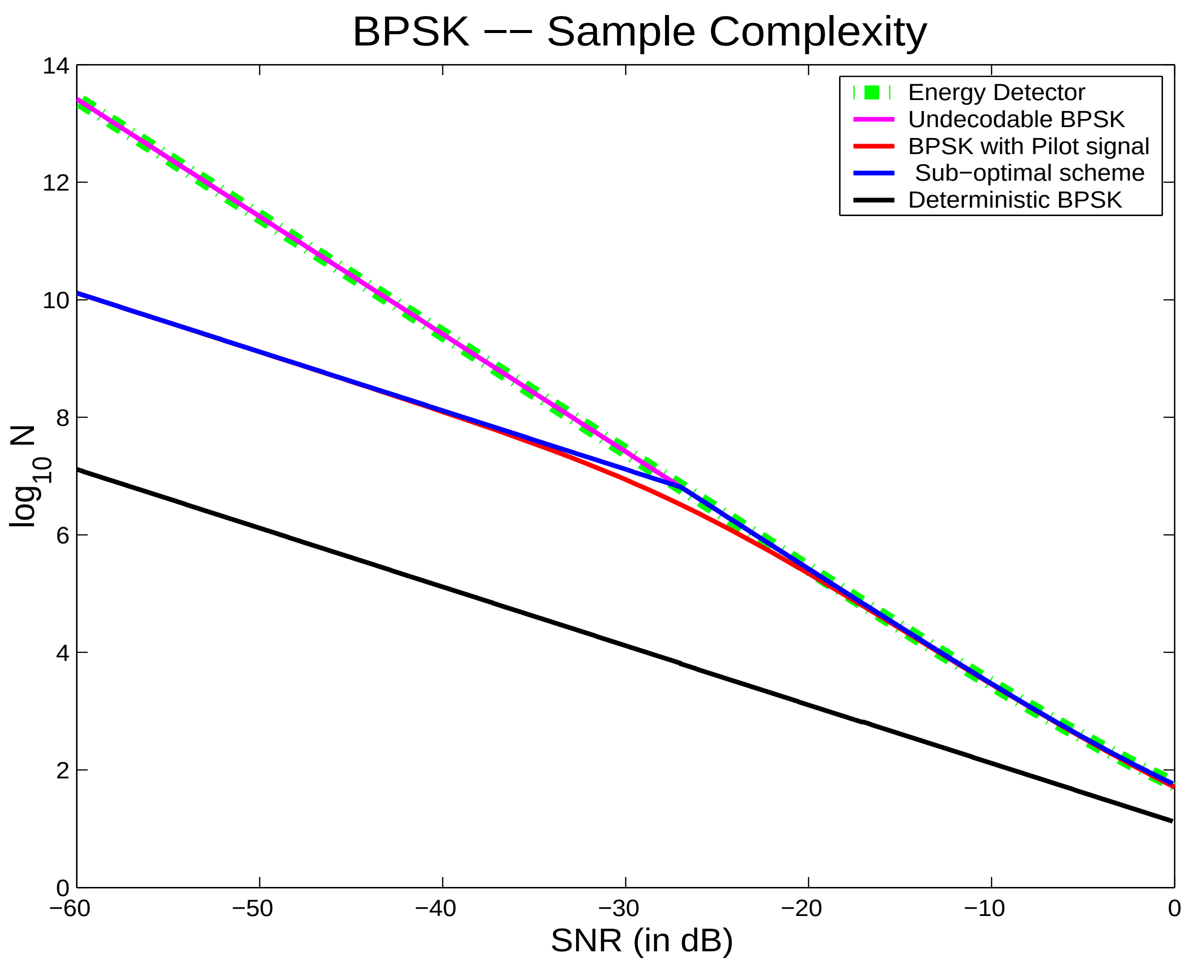


Figure 2.2. The figure compares the sample complexity curves for an undecodable BPSK signal without a pilot and the sample complexity curves of an undecodable BPSK signal with a known pilot signal. The dashed green curve shows the performance of the energy detector, the pink curve corresponds to the performance of the optimal detector. Both these curves are for the case without a pilot signal. These curves show that the energy detector performance is same as that of the optimal detector. The red curve gives the performance of the optimal detector in the presence of a known weak pilot signal. Note, that there is a significant decrease in sample complexity due to the known pilot signal, especially at very low SNR's. Specifically, at low SNR, the sample complexity changes from $O(1/SNR^2)$ to $O(1/SNR)$ due the presence of the pilot.

Chapter 3

Noise Uncertainty

3.1 Introduction

In section 2.2 we formulated the problem of signal detection as a hypothesis testing problem and derived the sample complexity of the optimal detector for the special case of Gaussian noise in section 2.3.4. In this chapter we relax the Gaussian noise assumption, and assume that the noise lies in some neighborhood of a nominal Gaussian. We then derive the sample complexity of moment detectors. For clarity of analysis we continue to assume the the signal is random BPSK-modulated with low signal power. However, all the results derived here, can easily be extended to any general zero-mean signal constellation as long as the signal amplitude is low.

We begin by motivating the need for relaxing the Gaussian noise assumption and then propose a model for the class of distributions that noise is allowed to choose. Under this model, we show that the sample complexity of moment detectors is infinite for SNR's below a certain threshold, i.e., detection is impossible using these moment detectors below this threshold. We close this chapter by arguing that our noise uncertainty model is minimalistic and show that our results continue to hold under any other reasonable uncertainty model.

3.2 Need for an uncertainty model

Noise in most communication systems is an aggregation of various independent sources including not only thermal noise at the receiver, but also interference due to nearby unintended emissions, weak signals from transmitters very far away, etc. By appealing to the Central Limit Theorem (CLT), one usually assumes that the noise at the receiver is a Gaussian random variable. But we know that the error term in the CLT goes to zero only as $\frac{1}{\sqrt{N}}$, where N is the number of independent random variables being summed up to constitute noise (see theorem 2.4.9 in [3] for details). In real life N is usually moderate and the error term in the CLT should not be neglected, especially while detecting low SNR signals.

For example, consider $X_1, X_2, \cdots$ be i.i.d. with $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^2 = \sigma^2$. Let $F_N(x)$ be the distribution of $W_N := \frac{X_1+X_2+\cdots+X_n}{\sqrt{N}}$ and $\mathcal{N}(x)$ is the standard normal distribution, then from theorem 2.4.9 in [3], we have

$$|F_N(x) - \mathcal{N}(x)| \leq K\frac{1}{\sqrt{N}}$$

for some constant K. This means that the actual distribution function $F_N(x)$ is within a certain band of the Gaussian distribution $\mathcal{N}(x)$. Now, consider a very weak zero-mean signal X that is added to W_N. If the signal power is weak then the distribution of the signal plus noise will also lie within a band of the Gaussian distribution. In this case, it is possible that the receiver detects the presence of a signal even when it is actually absent. This is possible because, the error term in the CLT might start looking like a weak additive signal.

Therefore, it is clear that in this simple example, detection can become very hard if the receiver makes the assumption that the noise is perfectly Gaussian. Further, as the CLT is a statement about convergence in distribution, we cannot tell anything about the behavior of the moments of the noise.

The above discussion leads us to the conclusion that the actual noise process is only approximately Gaussian. However, most receivers work under the Gaussian assumption of noise and try to measure the noise variance σ_n^2 (assume that 'n' stands for nominal and 'a' stands for actual in all the subscripts hereafter). Therefore it is important to understand how the detection procedures designed under the assumption of Gaussian distributions work when there are deviations from the Gaussian assumption.

3.3 Noise uncertainty model

We know that the receiver can try to estimate the noise distribution by taking a large number of samples. But, there will always be some *residual uncertainty* in estimating the noise distribution. We model this residual uncertainty by assuming that the receiver can narrow down the noise process only to within a class of distributions denoted by $\mathcal{W}_x$, where x parameterizes the amount of residual uncertainty. We call this set, the *uncertainty class of noise* for a given receiver. In general, the uncertainty class of noise includes a nominal noise distribution, based on which receivers design their detection procedures. From the discussion in the previous section it is reasonable to assume that the nominal noise distribution is Gaussian. This is a very practical assumption and most of the receivers built are designed for operating under this nominal Gaussian assumption.

To begin with we need to make the following basic assumptions on this set $\mathcal{W}_x$:

- The noise processes in this set must all be 'white', i.e., samples are i.i.d.
- The noise processes must be symmetric, i.e., if $W_a \in \mathcal{W}_x$ then we assume that $\mathbb{E}W_a^{2k-1} = 0$. We make this assumption for physical reasons.

Finally, we need to model the fact that there is at most x dB of uncertainty in the noise processes in $\mathcal{W}_x$. Therefore, for every noise random variable $W_a \in \mathcal{W}_x$ we assume that

- All the moments of the noise process must be close to the nominal noise moments, i.e., we assume that $\mathbb{E}W_a^{2k} \in [\frac{1}{\alpha}\mathbb{E}W_n^{2k}, \alpha\ \mathbb{E}W_n^{2k}]$, where $W_n \sim N(0, \sigma_n^2)$ is the estimated nominal Gaussian random variable and $\alpha = 10^{x/10} > 1$. That is,

$$\mathcal{W}_x := \left\{ W_a : \frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \in \left[\frac{1}{\alpha}, \alpha\right], W_n \sim N(0, \sigma_n^2) \right\} \tag{3.1}$$

Having proposed a model for noise uncertainty, we rewrite the detection problem incorporating our model.

3.3.1 The problem of detection under uncertainty

For concreteness we rewrite the detection problem in (2.1) for our specific uncertainty model. As before, we formulate it as a binary hypothesis test between

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= W[n] \qquad\qquad n = 1, \ldots, N \\ \mathcal{H}_1 : Y[n] &= X[n] + W[n] \qquad n = 1, \ldots, N \end{aligned} \tag{3.2}$$

where $W \in \mathcal{W}_x$ and X can be any general zero-mean digitally modulated signal with low signal amplitude, and is independent of the noise processes. However, as mentioned before, we focus on the case when X is random BPSK modulated data, i.e., $X \sim \text{Bernoulli}(\frac{1}{2})$ taking values in $\pm\sqrt{P}$.

We now use the following definition of a *robust detector* in our formulation.

***Definition** 1:* A statistical procedure is called *robust*, if its performance is insensitive to small deviations of the actual situation from the idealized theoretical model.

In particular, we require that a robust detector meets *both* the target[1] probability of false alarm P_{FA}, and probability of missed detection P_{MD}, regardless of which possible noise process actually occurs. Otherwise, we say that detection is impossible with this particular detector under our uncertainty set $\mathcal{W}_x$. This makes sense because the uncertain set $\mathcal{W}_x$ models post-noise-estimation uncertainty.

Given the above formulation, our ultimate goal is to find out if there exists a *robust detector* for our problem. It turns out that the analysis of this case is hard and hence we consider a particular class of detectors, namely moment detectors and show that they are non-robust under noise uncertainty.

[1]The target P_{FA} and P_{MD} are parameters that certainly will impact detection performance in terms of sample complexity. However, our primary interest in this paper is in showing detection-impossibility results as a function of SNR. Thus, we will not dwell on the quantitative impact of the chosen P_{FA}, P_{MD} on our results.

Also, for calculations in the next section, we recall that the moments of $W_n \sim N(0, \sigma_n^2)$ random variable are given by

$$\mathbb{E}W_n^k = \begin{cases} 0 & \text{if } k \text{ is odd;} \\ 1 \cdot 3 \cdot 5 \cdots (k-1)\sigma_n^k & \text{if } k \text{ is even.} \end{cases}$$

For convenience, we use the following notation

$$(2k-1)!! := 1 \cdot 3 \cdot 5 \cdots (2k-1)$$

therefore, we denote the even moments of a Gaussian random variable by $\mathbb{E}W_n^{2k} = (2k-1)!!\ \sigma_n^{2k}$.

3.4 Implications of noise uncertainty

Analysis of (3.2) for the optimal detector is mathematically intractable. Therefore, we focus on a class of sub-optimal detectors (moment detectors) and show that detection is impossible for these detectors at sufficiently low SNR's. Before we try to answer this general question, lets review the radiometer case, [26] (second moment detector).

Clearly, if $\sigma_a^2 = \sigma_n^2 + P$ the radiometer can be made to believe that the signal is present even when the signal is actually absent. On the flip side, if $\sigma_a^2 + P = \sigma_n^2$ the radiometer thinks that the signal is absent even when it is actually present. Thus, this example clearly illustrates that if the SNR is sufficiently low, there is enough uncertainty in the noise to render the radiometer useless.

Recall that we use lower case snr to denote the signal to noise ratio and reserve SNR to denote the signal to noise ratio in decibels. We now give the following important definition.

Definition *2:* Define snr_{wall} (SNR_{wall}) to be the maximal snr (SNR) such that for every $snr \leq snr_{wall}$ ($SNR \leq SNR_{wall}$) detection is impossible for the particular detector[2]. In particular we denote $snr_{wall}^{(2k)}$ ($SNR_{wall}^{(2k)}$) for the special case of $2k$-th moment detectors.

Now, we try to see if the same threshold behavior as in the case of the radiometer is observed for more sophisticated detectors that look at higher moments of the received signal.

***Theorem** 1:* Consider the detection problem in (3.2), under the $2k$-th moment detector. Here we assume that the actual noise distribution lies in the uncertainty class $\mathcal{W}_x$, which is centered around a nominal Gaussian $W_n \sim N(0, \sigma_n^2)$. Under this model, we prove that there exists an SNR wall ($snr_{wall}^{(2k)}$) for each $2k$-th moment detector. Further, we show that the $snr_{wall}^{(2k)}$ satisfies

$$\frac{\alpha - 1}{k}\left[1 - \frac{\alpha - 1}{2 - \alpha}\right] \leq snr_{wall}^{(2k)} \leq \frac{\alpha - 1}{k} \tag{3.3}$$

where $\alpha = 10^{x/10}$.

Proof: The test statistic for the $2k$-th moment detector is given by: $T(\mathbf{Y}) = \frac{1}{N}\sum_{i=1}^{N} Y[i]^{2k}$, where $Y[i]$ is the i-th received sample. It is easy to see that the detector in consideration fails if the mean of this test statistic under both the hypothesis are equal. This can happen in the following two ways.

Case I: $\mathbb{E}W_a^{2k} = \mathbb{E}[X + W_n]^{2k}$, i.e., the actual noise moments are high enough to make the receiver wrongly believe that there is a signal present. This makes the false alarm probability, P_{FA} go to 1. We begin by expanding $\mathbb{E}[X + W_n]^{2k}$ and re-writing the condition as

$$\mathbb{E}W_a^{2k} = \sum_{i=0}^{k} \binom{2k}{2i} \mathbb{E}W_n^{2k-2i}\, \mathbb{E}X^{2i} \tag{3.4}$$

Dividing by $\mathbb{E}W_n^{2k}$ on both sides of (3.4) we get,

[2]This definition of course depends on the detection problem of interest. We are assuming that the detection problem under consideration is the one given in (3.2). Specifically we are assuming that we are detect the presence or absence of a zero-mean randomly modulated BPSK signal under the noise uncertainty model in (3.1).

$$
\begin{aligned}
\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} &= \sum_{i=0}^{k} \binom{2k}{2i} \left(\frac{\mathbb{E}W_n^{2k-2i}}{\mathbb{E}W_n^{2k}} \right) \mathbb{E}X^{2i} \\
&= \sum_{i=0}^{k} \binom{2k}{2i} \left(\frac{1 \cdot 3 \cdots (2k-2i-1)}{1 \cdot 3 \cdots (2k-1)} \right) \frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}} \\
&= \sum_{i=0}^{k} \frac{\binom{2k}{2i}}{(2k-2i+1) \cdot \cdots \cdot (2k-1)} \, snr^i \\
&= 1 + k \cdot snr + \cdots + \frac{1}{(2k-1)!!} \, snr^k \qquad (3.5)
\end{aligned}
$$

Until now we have not used the fact that the actual noise random variable $W_a \in \mathcal{W}_x$, i.e., $\frac{1}{\alpha} \leq \frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \leq \alpha$. Since, the right hand side of (3.5) is greater than 1, and is also monotonic in snr, $snr_{wall}^{(2k)}$ must be a solution to

$$
\begin{aligned}
\alpha &= 1 + k \cdot snr + \cdots + \frac{1}{(2k-1)!!} \, snr^k \qquad (3.6) \\
&= \sum_{i=0}^{k} \frac{\binom{2k}{2i}}{(2k-2i+1) \cdot \cdots \cdot (2k-1)} \, snr^i \qquad (3.7)
\end{aligned}
$$

From (3.7) it is obvious that $1 + k \cdot snr \leq \alpha$. This simplifies to give $snr \leq \frac{\alpha-1}{k}$, which is the desired upper bound in (3.3). To get a lower bound, we substitute this upper bound into (3.7)

$$
\begin{aligned}
\alpha &\leq \sum_{i=0}^{k} \frac{\binom{2k}{2i}}{(2k-2i+1) \cdot \cdots \cdot (2k-1)} \left(\frac{\alpha-1}{k} \right)^i \\
&< 1 + k \cdot snr + \sum_{i=2}^{k} (\alpha-1)^i \\
&< 1 + k \cdot snr + \sum_{i=2}^{\infty} (\alpha-1)^i \\
&= 1 + k \cdot snr + \frac{(\alpha-1)^2}{2-\alpha} \qquad (3.8)
\end{aligned}
$$

for $1 \leq \alpha \leq 2$.

In the second inequality above we have used the fact that $\frac{\binom{2k}{2i}}{(2k-2i+1) \cdot \cdots \cdot (2k-1)} \frac{1}{k^i} < 1$ for $k \geq 2$. Re-writing (3.8) as

$$k \cdot snr \geq (\alpha - 1)\left[1 - \frac{\alpha - 1}{2 - \alpha}\right]$$

we get the lower bound in (3.3).

Case II: $\mathbb{E}[W_n]^{2k} = \mathbb{E}[X + W_a]^{2k}$, i.e., the actual noise moments are lower than the receiver's estimate by enough so that the receiver fails to detect the signal. In this case the missed detection probability, P_{MD} goes to 1. Expanding $\mathbb{E}[X + W_a]^{2k}$, we must have

$$\mathbb{E}W_n^{2k} = \sum_{i=0}^{k} \binom{2k}{2i} \mathbb{E}W_a^{2k-2i}\ \mathbb{E}X^{2i}$$

As in the previous case, dividing by $\mathbb{E}W_n^{2k}$, we get

$$1 = \sum_{i=0}^{k} \binom{2k}{2i} \left(\frac{\mathbb{E}W_a^{2k-2i}}{\mathbb{E}W_n^{2k}}\right) \mathbb{E}X^{2i}$$

which simplifies to

$$\begin{aligned} 1 &= \sum_{i=0}^{k} \binom{2k}{2i} \left(\frac{1 \cdot 3 \cdots (2k-2i-1)}{1 \cdot 3 \cdots (2k-1)}\right) \frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}} \frac{\mathbb{E}W_a^{2k-2i}}{\mathbb{E}W_n^{2k-2i}} \\ &= \sum_{i=0}^{k} \frac{\binom{2k}{2i}}{(2k-2i+1) \cdot \cdots (2k-1)}\ snr^i\ \frac{\mathbb{E}W_a^{2k-2i}}{\mathbb{E}W_n^{2k-2i}} \\ &= \frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} + k \cdot snr \frac{\mathbb{E}W_a^{2k-2}}{\mathbb{E}W_n^{2k-2}} + \ldots + \frac{1}{(2k-1)!!} snr^k \end{aligned} \quad (3.9)$$

Now, using the condition that $W_a \in \mathcal{W}_x$, we must have $\frac{\mathbb{E}W_a^{2l}}{\mathbb{E}W_n^{2l}} \geq \frac{1}{\alpha}$, for all $l \leq k$. Substituting this in (3.9) we can easily see that $snr_{wall}^{(2k)}$ must be a solution to

$$\begin{aligned} 1 &= \frac{1}{\alpha} + k \cdot snr \frac{1}{\alpha} + \ldots + \frac{1}{(2k-1)!!} snr^k \\ \Rightarrow \alpha &= 1 + k \cdot snr + \ldots + \frac{1}{(2k-1)!!} snr^k\ \alpha \end{aligned} \quad (3.10)$$

Equation (3.10) looks very similar to (3.7). In fact it is exactly identical to it except the 'α' in the last term on the right hand side of (3.10). As $\alpha > 1$, the snr solution to (3.10) is strictly smaller than the solution to (3.7). Therefore, the final $snr_{wall}^{(2k)}$ is given by the solution to equation (3.10). However, the difference between the solution to equations (3.7) and (3.10) is tiny in the interesting case when $\alpha \approx 1$. Thus, the lower bound in (3.3) continues to hold even for this case (see fig. 3.1). ■

***Remarks** 1:*

- Note that the difference between the derived upper and lower bound (in dB) is $-10\log_{10}\left[1 - \frac{\alpha-1}{2-\alpha}\right] \approx -10(\alpha-1)\log_{10} 2$, for $\alpha \approx 1$. This shows that the gap between our bounds is negligible compared to the value of the $SNR_{wall}^{(2k)}$ (Refer to fig. 3.1).
- Also, note that the position of $SNR_{wall}^{(2k)}$ varies as a logarithm of k and hence checking for large moments buys us very little in terms of detectability.
- Finally, from fig. 3.1 it is clear that the upper bound in (3.3) is very tight compared to the lower bound. This makes sense because the lower bound was obtained using very weak bounding techniques. So, for all practical purposes we can assume that the location of the $snr_{wall}^{(2k)}$ is given by the upper bound in (3.3), i.e., $snr_{wall}^{(2k)} \approx \frac{\alpha-1}{k}$

3.5 Discussion of results

Theorem 1 shows that when the signal to be detected is random, *non-detectability* is always a problem at low enough SNR's regardless of how many samples we take or which moment we look at. A perfectly known signal, on the other hand, can be detected even under such noise uncertainty since we get coherent processing gain using the matched filter. The signal is isolated, while the noise is averaged. The moments of the averaged noise decrease to zero as the number of samples increases regardless of where it lies in the uncertainty set.

In retrospect, the result of theorem 1 is not very surprising. In any binary hypothesis testing problem, with *simple hypotheses*, the distribution under each hypothesis determines how well it is possible to differentiate between the two hypothesis. For example, if both the hypotheses are the same then there is nothing to differentiate between them. This observation is trivial and relatively unimportant. Now, lets consider the case when both

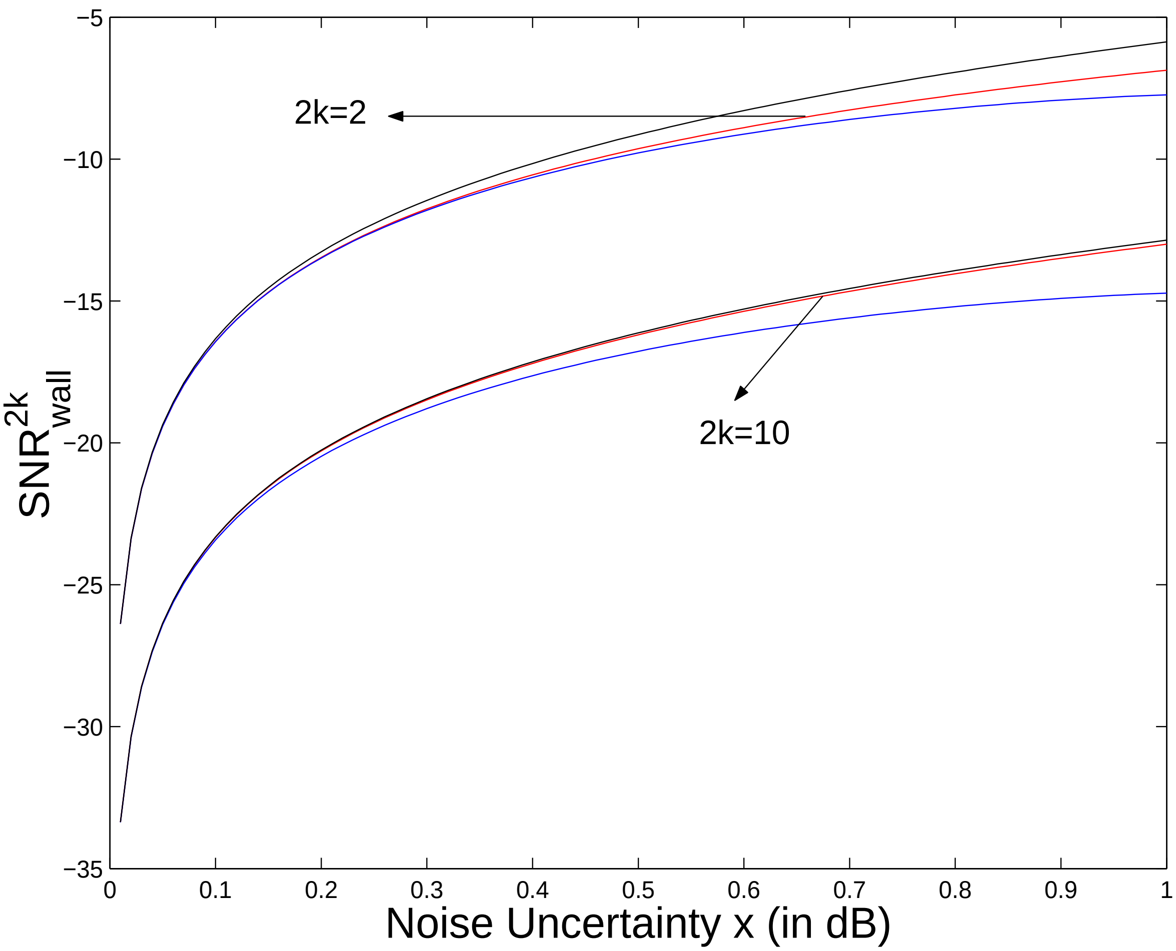


Figure 3.1. The location of the $SNR_{wall}^{(2k)}$ (in dB) for $2k = 2, 10$, as a function of the noise uncertainty x (in dB) is plotted in this figure. The red curves in the plot have been computed numerically by calculating the root of the equation (3.7), the blue curves are computed using the lower bound in Theorem 1, and the black curves are plotted using the upper bound in Theorem 1.

the hypotheses are *composite*, i.e., both the hypotheses can represent a class of distributions and not a single distribution as before. What happens if these distributional classes overlap?

On some thought it should be clear that one cannot guarantee reliable detection, i.e., for any possible detector, the error probability can never be made arbitrarily small. In other words there does not exist any robust detector in this case. When thought carefully, this is the essence of our result.

The minor difference between the argument given above and our result is that, we restrict ourself to a specific class of detectors and we show that the distribution classes under both hypotheses overlap. Note that the overlap of distributional classes is conditioned on the fact that we are constrained to use a particular class of detectors, i.e., we are using only a subspace of the information by restricting to these sub-optimal detectors.

The same argument is pictorially presented in Fig. 3.2a. The horizontal line represents the possible locations of the test statistic $T(\mathbf{Y})$, and the $2k$-th moments under both hypotheses are marked as two points on this line. The dotted vertical line represents the threshold γ which divides the whole space into two decision regions corresponding to the two hypotheses.

Now, let us see how the picture changes when we consider uncertainty in noise. According to our model for uncertainty, noise can have a set of possible $2k$-th moments under both hypotheses, which are denoted by intervals in Fig. 3.2b. For low signal powers, these two intervals must overlap, as shown by the shaded region. Note that whatever threshold γ (vertical line in the figure) we choose, the detector cannot guarantee that both P_{FA} and P_{MD} are low enough for all the noise processes in the set $\mathcal{W}_x$. Due to the shaded region, at least one of P_{FA} and P_{MD} can be made to go to 1. Therefore, the moment detectors are useless for detecting signals below the SNR wall.

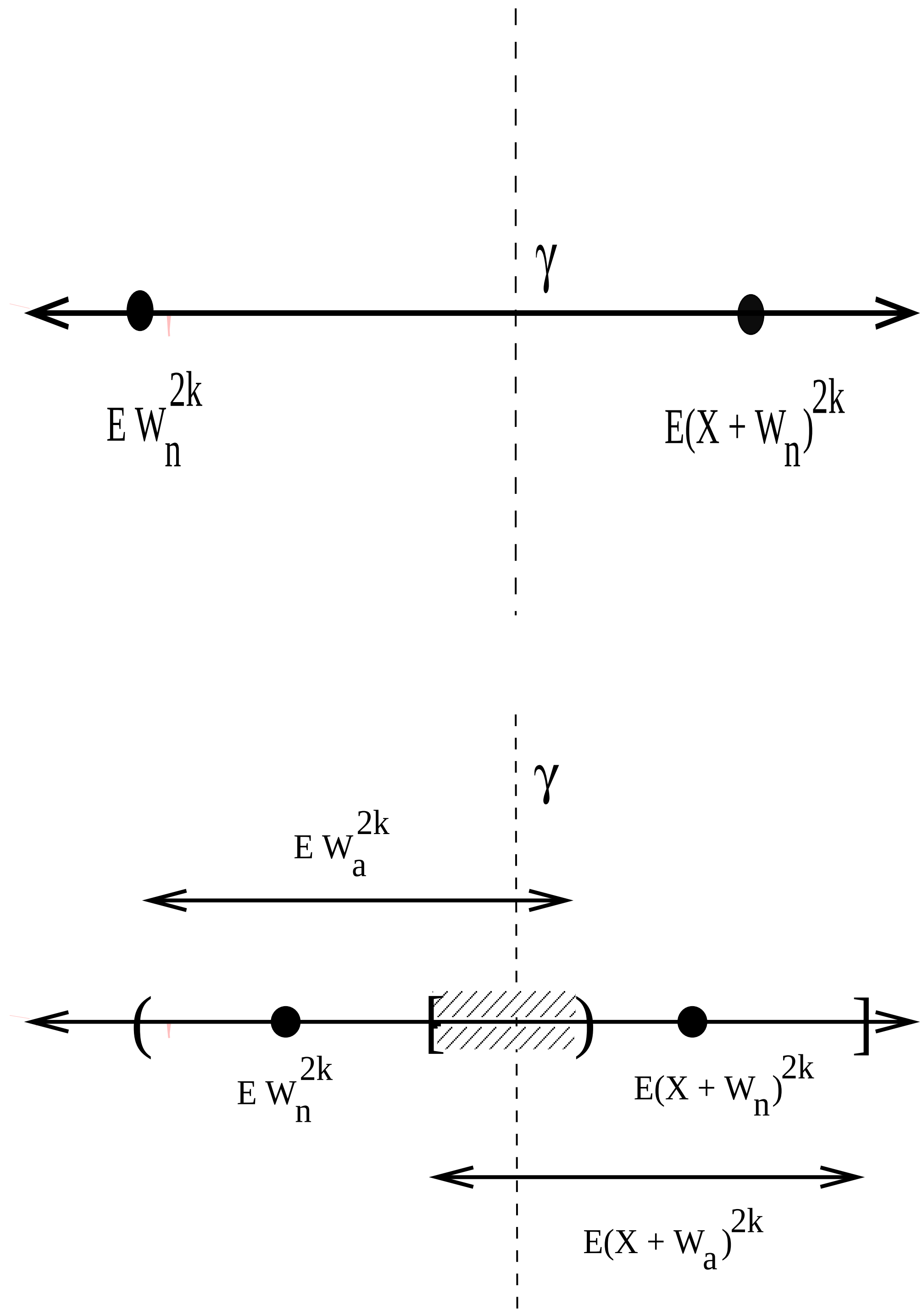


Figure 3.2. This figure pictorially depicts the two hypothesis and the threshold in a $2k$-th moment detector. The figure on the top shows the case when there is no noise uncertainty and the one in the bottom is the case when there is noise uncertainty. Note that under noise uncertainty and sufficiently weak signal power, there is an overlap region under both hypotheses, which renders the detector useless.

3.6 Approaching the SNR wall

Figure 3.3 shows the sample complexity of moment detectors near the SNR wall. The green curve shows the sample complexity of the radiometer without noise uncertainty ($N = O(1/SNR^2)$, slope =2). The remaining curves in the figure show the sample complexities of various moment detectors with noise uncertainty. Recall that the test statistic for a $2k$-th moment detector is given by $T(\mathbf{Y}) = \frac{1}{N}\sum_{i=1}^{N} Y[i]^{2k}$, where $Y[i]$ is the i-th received sample. The values of N we are considering here are very large, in fact larger than the number of samples required for the radiometer without noise uncertainty. Also, our target probability of missed detection and false alarm are moderate, i.e., not changing with N. Therefore the Central Limit theorem is a good approximation here (Note that the error in the central limit theorem decays as $1/\sqrt{N}$. [3]). Hence, we assume that $\frac{T(\mathbf{Y}) - N\mathbb{E}Y_i^{2k}}{\sqrt{N}var(Y_i)} \sim \mathcal{N}(0,1)$. This reduces the problem into a standard binary hypothesis testing problem with different mean under both hypotheses. Therefore, N is given by:

$$N = \left[\frac{Q^{-1}(P_{FA})\sqrt{(\mathbb{V}(Y_i^{2k})|H_0)} - Q^{-1}(P_D)\sqrt{(\mathbb{V}(Y_i^{2k})|H_1)}}{\mathbb{E}(Y_i^{2k}|H_1) - \mathbb{E}(Y_i^{2k}|H_0)} \right]^2 \tag{3.11}$$

where $\mathbb{V}(.)$ stands for the variance operator. Recall that the detector must hit the target error probabilities uniformly over the whole uncertainty set $\mathcal{W}_x$. Therefore, the sample complexity is dominated by the case when the difference in means (denominator term in (3.11)) in the above equation is minimized.

Thus, the sample complexity required to meet our performance targets uniformly over the uncertain noise tends to infinity as the SNR tends to $SNR_{wall}^{(2k)}$. Also note that these performance curves shift to the left by $10\log k$, which verifies that the upper bound obtained in theorem 1 is very tight.

3.7 Other possible noise uncertainty models

One might argue that the results in section 3.4 arise due to the specific model we used for noise uncertainty, and that they are not fundamental. In this section, we try to show

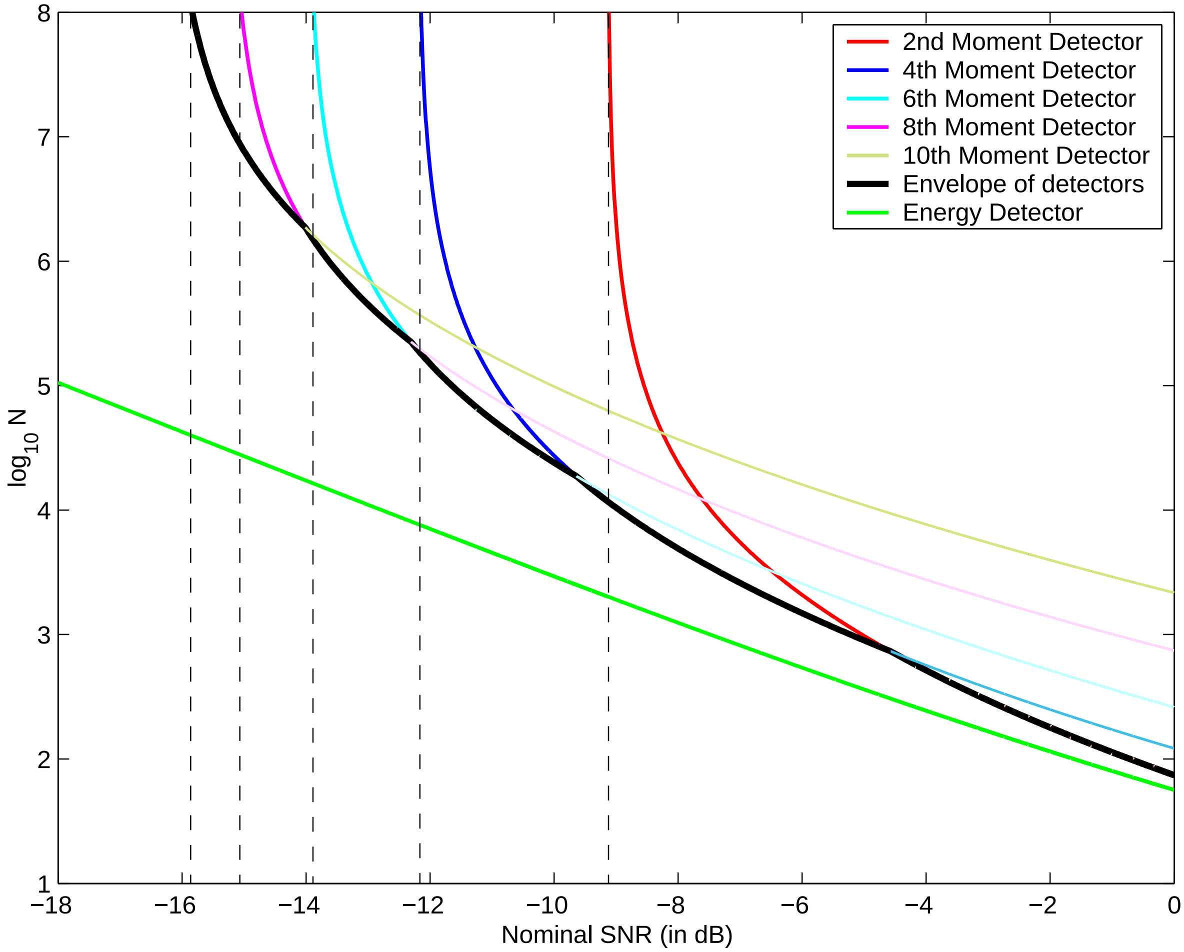


Figure 3.3. This figure shows the sample complexity of the moment detectors (number of samples as a function of SNR) at a moderate noise uncertainty of $x = 0.5$ dB. The curves in this figure have been computed using the equation (3.11)

that our model is a minimalist model and any other reasonable noise uncertainty model will lead to an uncertainty class which includes our uncertainty class. Thus in practice, the problem will only get worse.

For example, consider the simple case in which the receiver assumes that the noise is Gaussian, but its estimate of the noise variance is off by some factor, i.e., the receiver estimates the noise variance to be σ_n^2, whereas the actual noise variance is σ_a^2 ($\sigma_n \neq \sigma_a$). If $\sigma_n > \sigma_a$ then the ratio of their $2k$-th moments is $\left(\frac{\sigma_n}{\sigma_a}\right)^{2k}$, which goes to infinity as $k \to \infty$. Conversely, if $\sigma_n < \sigma_a$ then $\left(\frac{\sigma_n}{\sigma_a}\right)^{2k}$ goes to zero as $k \to \infty$. Therefore we have shown that even such a simple case is not included in our noise model. In fact, the above argument also shows that our uncertainty class $\mathcal{W}_x$ includes just one Gaussian, W_n and contains no other Gaussian random variable.

Motivated by the above example, we propose an alternative noise uncertainty model. As before, the basic assumptions on the noise remain the same, i.e.,

- We assume that all the noise processes in the uncertainty set to be white.
- Also, we assume that the noise distribution is symmetric, and hence all the odd moments of the noise random variables are zero.
- Now, suppose that the user estimates the noise to be Gaussian with variance σ_n^2, i.e., $\widetilde{W}_n \sim N(0, \sigma_n^2)$. Define the new noise uncertainty set $\widetilde{\mathcal{W}}_x$ to be the set of all noise random variables, $\widetilde{W}_a$ such that the moments of $\widetilde{W}_a$ are all sandwiched between the corresponding moments of $\frac{1}{\alpha}\,\widetilde{W}_n$ and $\alpha\widetilde{W}_n$, i.e., $\mathbb{E}\widetilde{W}_a^{2k} \in [\frac{1}{\alpha^{2k}}\mathbb{E}\widetilde{W}_n^{2k}, \alpha^{2k}\,\mathbb{E}\widetilde{W}_n^{2k}]$ where $\alpha = 10^{(x/10)}$. In other words,

$$\widetilde{\mathcal{W}}_x := \left\{ \widetilde{W}_a : \frac{\mathbb{E}\widetilde{W}_a^{2k}}{\mathbb{E}\widetilde{W}_n^{2k}} \in \left[\frac{1}{\alpha^{2k}}, \alpha^{2k}\right], \widetilde{W}_n \sim N(0, \sigma_n^2) \right\} \tag{3.12}$$

This model is very practical, since in real life, most receivers are not sensitive enough to be able to differentiate between two Gaussian random variables with very close variances. Under this model, we now show that there exists a single threshold below which *every* *possible detector* fails to detect the signal. Thus, under this model robust detection becomes *absolutely impossible*.

For our convenience we rewrite the detection problem under the new noise uncertainty model

$$\begin{aligned}\mathcal{H}_0 : Y[n] &= \widetilde{W}[n] \qquad\qquad n = 1, \ldots, N \\ \mathcal{H}_1 : Y[n] &= X[n] + \widetilde{W}[n] \qquad n = 1, \ldots, N\end{aligned} \tag{3.13}$$

where $\widetilde{W} \in \widetilde{\mathcal{W}}_x$ and the remaining setup of the problem is same as in (3.2).

***Theorem** 2:* Consider the detection problem in (3.13). Here we assume that the actual noise distribution lies in the alternate uncertainty class $\widetilde{\mathcal{W}}_x$ described in (3.12). Under this model, we show that there exists an *absolute SNR wall* (snr^*_{wall}) for any possible robust detector[3]. Further, the snr^*_{wall} (the $*$ is used to refer to the fact that it is an absolute wall) satisfies

$$snr^*_{wall} = \min_{k>0} snr^{(2k)}_{wall} \leq \alpha^2 - 1 \tag{3.14}$$

where $snr^{(2k)}_{wall}$ is the snr wall for the $2k$-th moment detector and $\alpha = 10^{x/10}$.

Proof: It is clear that, detection is *absolutely impossible* if there exists a noise distribution $\widetilde{W}_a \in \widetilde{\mathcal{W}}_x$, such that $\widetilde{W}_a = \widetilde{W}_n + X$ or $\widetilde{W}_n = \widetilde{W}_a + X$, i.e., either P_{FA} goes to 1 or P_{MD} goes to 1 respectively. We show that the first condition is satisfied, i.e., $\widetilde{W}_a := \widetilde{W}_n + X$.

Observe that

$$\widetilde{W}_a = \widetilde{W}_n + X \tag{3.15}$$

$$\Leftrightarrow \quad \mathbb{E}\widetilde{W}_a^k = \mathbb{E}[\widetilde{W}_n + X]^k \qquad \text{for every } k > 0 \tag{3.16}$$

Since (3.16) is trivially satisfied for k odd, we consider only the case when k is even. Now, fix $k = 2k_0$. Therefore we must have

$$\mathbb{E}\widetilde{W}_a^{2k_0} = \mathbb{E}[\widetilde{W}_n + X]^{2k_0} \tag{3.17}$$

[3]Whenever we refer to snr_{wall} in this theorem, we are referring to the wall with respect to the detection problem in (3.13) and with the noise uncertainty model given by (3.12)

Expanding the above equation, we have

$$\begin{aligned}
\frac{\mathbb{E}\widetilde{W}_a^{2k_0}}{\mathbb{E}\widetilde{W}_n^{2k_0}} &= \sum_{i=0}^{k_0} \binom{2k_0}{2i} \left(\frac{\mathbb{E}\widetilde{W}_n^{2k_0-2i}}{\mathbb{E}\widetilde{W}_n^{2k}} \right) \mathbb{E}X^{2i} \\
&= \sum_{i=0}^{k_0} \binom{2k_0}{2i} \left(\frac{1 \cdot 3 \cdots (2k_0 - 2i - 1)}{1 \cdot 3 \cdots (2k_0 - 1)} \right) \frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}} \\
&= \sum_{i=0}^{k_0} \frac{\binom{2k_0}{2i}}{(2k_0 - 2i + 1) \cdot \cdots (2k_0 - 1)} \, snr^i \\
&= 1 + k_0 \cdot snr + \cdots + \frac{1}{(2k_0 - 1)!!} \, snr^{k_0} \qquad (3.18)
\end{aligned}$$

Since, $\widetilde{W}_a \in \widetilde{\mathcal{W}}_x$, we must have $\frac{\mathbb{E}\widetilde{W}_a^{2k_0}}{\mathbb{E}\widetilde{W}_n^{2k_0}} \leq \alpha^{2k_0}$. Using this bound for the left hand side of (3.18), we get

$$\alpha^{2k_0} \geq 1 + k_0 \cdot snr + \cdots + \frac{1}{(2k_0 - 1)!!} \, snr^{k_0} \qquad (3.19)$$

The above equation is exactly the same as in (3.7), and hence we must have

$$snr \leq snr_{wall}^{(2k_0)} \leq \frac{\alpha^{2k_0} - 1}{k_0} \qquad (3.20)$$

for all $k_0 \geq 0$. Here the second equation follows from theorem 1, when applied for the case when $\alpha = \alpha^{2k_0}$. Hence, detection is impossible iff

$$\begin{aligned}
\widetilde{W}_a &= \widetilde{W}_n + X \\
\Leftrightarrow \mathbb{E}\widetilde{W}_a^{2k_0} &= \mathbb{E}[\widetilde{W}_n + X]^{2k_0}, \qquad \forall\ k_0 > 0 \\
\Leftrightarrow snr &\leq \min_{k_0 > 0} snr_{wall}^{(2k_0)} \leq \min_{k_0 > 0} \frac{\alpha^{2k_0} - 1}{k_0} \\
\Leftrightarrow snr &\leq \min_{k_0 > 0} snr_{wall}^{(2k_0)} = \alpha^2 - 1
\end{aligned}$$

Here the last step is true because of the following fact. The function $f(t) := \frac{\alpha^{2t}-1}{t}$ is monotonically increasing in $t \in (0, \infty)$ for all $\alpha > 1$. This can be verified by differentiating $f(t)$ and showing that the first derivative is strictly positive for $t > 0$. Hence, we must have

$$\min_{k_0 > 0} \frac{\alpha^{2k_0} - 1}{k_0} = \alpha^2 - 1$$

Therefore, it follows from the definition of snr_{wall}^* that,

$$snr_{wall}^* = \min_{k_0 > 0} snr_{wall}^{(2k_0)} \leq \alpha^2 - 1 \qquad (3.21)$$

which is the required result. ■

***Remarks** 2:* In the proof of the above theorem, we have denoted $snr_{wall}^{(2k_0)}$ to be the solution[4] to the equation

$$\alpha^{2k_0} = 1 + k_0 \cdot snr + \cdots + \frac{1}{(2k_0-1)!!}\, snr^{k_0} \tag{3.22}$$

The name coincides with the fact that, this is indeed the location of the SNR wall for the $2k_0$-th moment detector for the detection problem in (3.2) with the noise uncertainty given by $\alpha = \alpha^{2k_0}$ (instead of α as before). Now, consider the solution to equation (3.22) as k_0 varies. We conjecture that the $snr_{wall}^{(2k_0)}$ is monotonically increasing in k_0 for any $\alpha > 1$. This have been verified numerically for large enough values of k_0 and some of the plots are given in fig. 3.5. If one believes in the above fact then, we must have

$$\begin{aligned} \min_{k_0>0} snr_{wall}^{(2k_0)} &= snr_{wall}^{(2)} \\ &= \alpha^2 - 1 \end{aligned}$$

Using this in (3.21), we get

$$snr_{wall}^{*} = \alpha^2 - 1 \tag{3.23}$$

This gives us an actual closed form expression for the location of the SNR wall under this alternate model.

In theorem 1 we derived bounds on the SNR walls for the various moment detectors, under our minimalist noise uncertainty model. There it was shown that the SNR walls decreases as we check for higher moments and hence the radiometer had the highest value for SNR wall, given by $\alpha - 1$. This is equivalent to $\alpha^2 - 1$ in our model, because in the minimalist model, the uncertainty in the second moment is bounded by α, whereas in this model the uncertainty in the second moment is bounded by $\alpha^2 - 1$. Comparing this to (3.23) we see that, the walls obtained under the alternative noise uncertainty model are greater than or equal to those obtained for moment detectors in theorem 1. The equality occurs specifically in the case of the radiometer (see fig. 3.4). This suggests that in real life the problem of non-detectability can be worse than suggested by the results in theorem 1.

[4] By solution, we mean the roots of the k_0-th degree polynomial given in (3.22). In general this equation will have k_0 roots, but we are interested in the non-negative root only, since $snr \geq 0$. However, it is easy to show that there is a unique non-negative solution to (3.22). Hence the term solution is unambiguous in this context.

3.7.1 Discussion

Again, observe that the result in theorem 2 is due to the fact that the distributional classes under both the hypotheses in (3.13) overlap. The only difference in this case is that there is overlap irrespective of the detector used, and hence detection is absolutely impossible below a certain threshold.

In fig. 3.4 the green curve is a plot of the snr^*_{wall} as a function of the uncertainty in the noise x. This curve has been obtained by numerically finding the minimum of $snr^{(2k_0)}_{wall}$ for a large range of values of k_0. The red curve in the plot is the upper bound for snr^*_{wall} obtained in (3.21). Both these curves are exactly lying over each other. This shows that (3.23) is indeed true. Finally, the blue curve in the plot is a plot of the snr_{wall} for the radiometer obtained in theorem 1. Observe the that the absolute snr^*_{wall}, is exactly equal to the snr_{wall} for the radiometer, as observed before.

Figure 3.5 plot the solution to (3.22) as a function of k_0, for three different values of x. These figures show that the solution to (3.22) is indeed monotonic in k_0 as conjectured. This shows that the minimum in (3.21) is indeed attained for $k_0 = 1$.

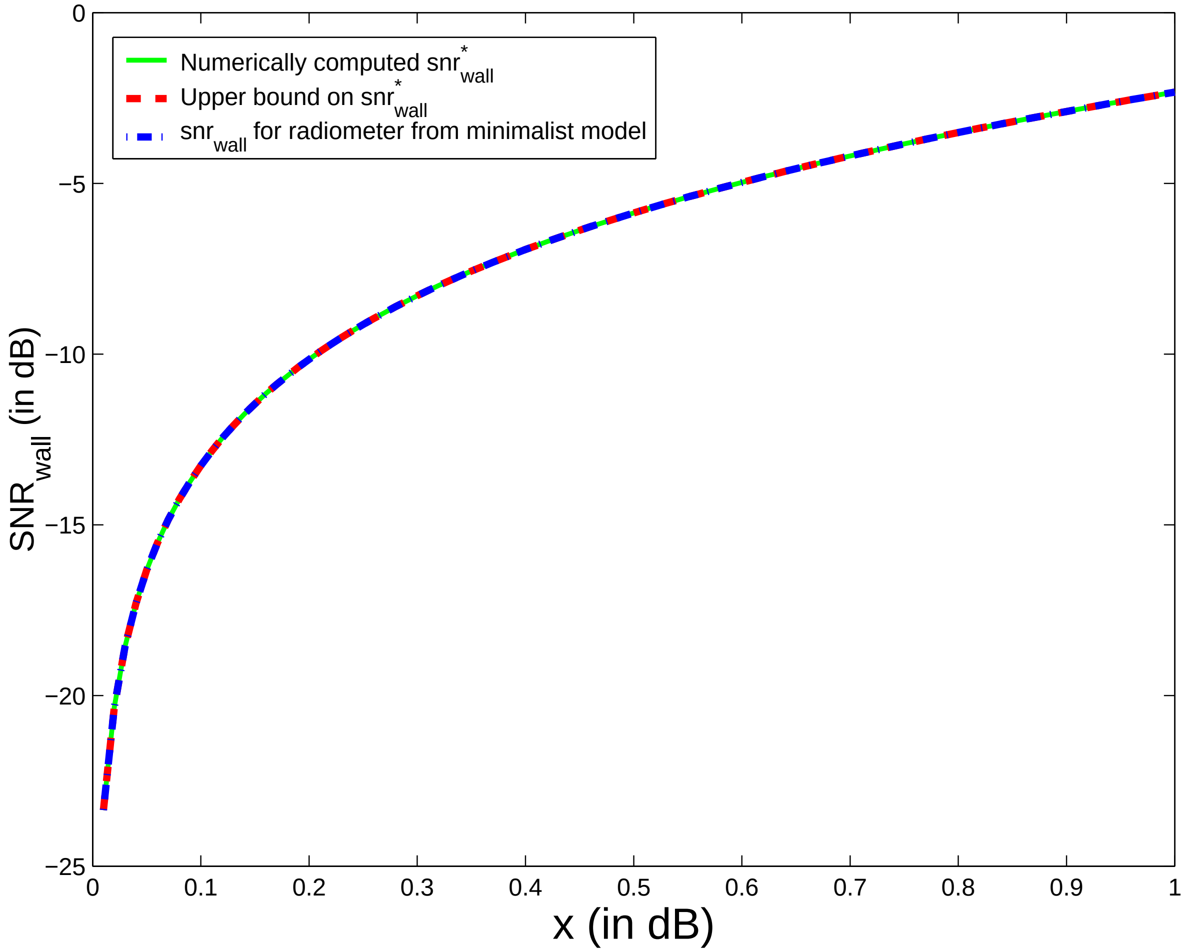


Figure 3.4. Variation of snr^*_{wall} is plotted as a function of the noise uncertainty x. The green curve is a plot of the snr^*_{wall}. This curve has been obtained by numerically finding the minimum of $snr^{(2k_0)}_{wall}$ for a large range of values of k_0. The red curve in the plot is the upper bound for snr^*_{wall} obtained in (3.21). These curves are compared with the blue curve which is the snr_{wall} for the radiometer under the minimalist uncertainty model in theorem 1. Note that all the curves lie on top of each other. This shows that (3.23) is indeed true.

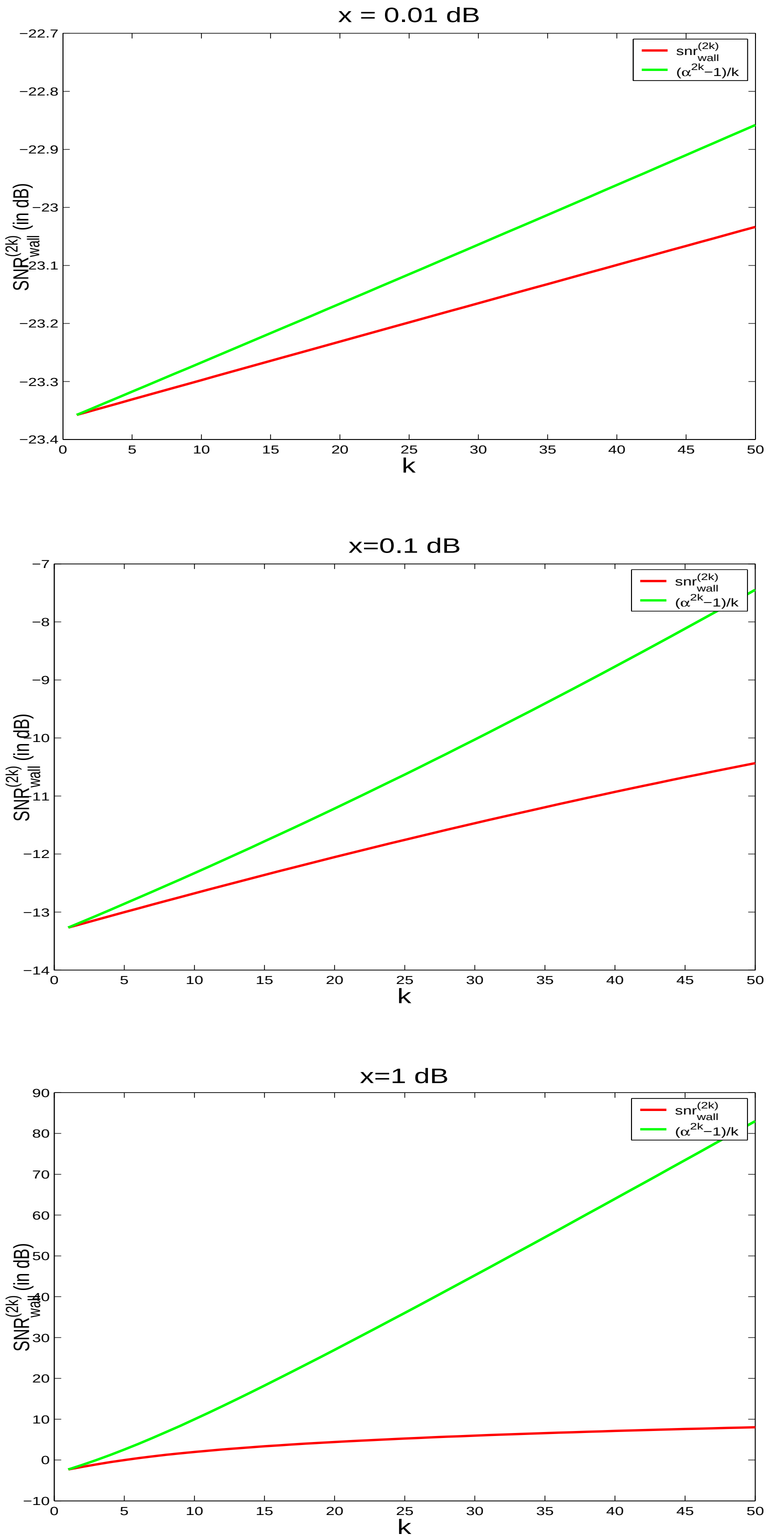


Figure 3.5. The red curves are plots of $snr_{wall}^{(2k)}$ as a function of k for three different values of x. These plots show that, $snr_{wall}^{(2k)}$ is monotonically increasing for $k > 0$ and hence its minimum is attained at $k = 1$. The green curves are an upper bound for $snr_{wall}^{(2k)}$ given by $\frac{\alpha^{2k}-1}{k}$.

Chapter 4

Finite Dynamic Range and Quantization

4.1 Introduction

In chapter 3 we motivated the need for taking the uncertainty in noise into consideration while detecting weak zero-mean signals. We proposed two models for the uncertainty in noise, where we assumed that the noise distribution can lie in a class of distributions $(\mathcal{W}_x, \widetilde{\mathcal{W}_x})$. For both these models we discussed the sample complexity of various detectors. In particular we showed that under the minimalist model discussed in section 3.3, detection is impossible for moment detectors, at low SNR. Under the alternate model for noise uncertainty in section 3.7, we showed that robust detection becomes *absolutely impossible* for *any detector*.

Another aspect of practical systems that has not been considered yet is: receivers having a *finite dynamic range* limitation. In practice the front ends of receivers have a *finite dynamic range* of operation, i.e., the signal coming into the processing unit is bounded in amplitude. In fact all the receivers have an internal A/D converter following the RF chain (see fig. 1.2). Therefore, all the detection that the receiver does is based on quantized versions of the received signal, not the received signal itself. In the previous chapters we ignored these two aspects and assumed that the receiver had infinite precision. It is intuitively clear that the finite dynamic range assumption and quantization should make detection harder.

In this chapter, we show that detection is impossible for any detector even under our minimalist model if the receiver has a finite dynamic range of operation. In particular this implies that all quantizers with finite number of quantization bins have to face a serious penalty, which is *non-detectability* below a certain SNR. However, we consider the 2-bit quantization case first to motivate the general impossibility result. Through out this chapter, we work with our minimalist model for noise uncertainty discussed in section 3.3[1].

4.2 Motivating example: 2-bit quantizer

In section 3.4, we have seen that noise uncertainty renders moment detectors useless if the SNR falls below a particular threshold. Does this same behavior hold if the signal is quantized? We analyze a 2-bit quantizer with noise uncertainty. The detection problem under quantization can be formulated as the following hypothesis testing problem,

$$\begin{aligned} \mathcal{H}_0 : Y[n] &= \mathcal{Q}(W[n]) \qquad\qquad n = 1, \ldots, N \\ \mathcal{H}_1 : Y[n] &= \mathcal{Q}(X[n] + W[n]) \quad n = 1, \ldots, N \end{aligned} \tag{4.1}$$

where $\mathcal{Q}(.)$ represents a 2-bit quantizer. Fig. 4.1 shows that bins in a 2-bit quantizer. Assume that the nominal noise variance is σ_n^2 and the actual noise variance is σ_a^2. In this case, it is easy to verify that the quantizer output is a binary random variable with the following pmf's

$$p_1^{(1)} = 1 - p_0^{(1)} = \left[Q\left(\frac{d_1 - \sqrt{P}}{\sigma_a}\right) + Q\left(\frac{d_1 + \sqrt{P}}{\sigma_a}\right) \right]; \qquad p_1^{(0)} = 1 - p_0^{(0)} = 2Q\left(\frac{d_1}{\sigma_a}\right)$$

[1]Considering the minimalist model for noise uncertainty is sufficient because of two reasons. Firstly, any impossibility result under this model, will directly imply the corresponding impossibility result in the other model. Secondly, we already showed that under the alternative noise uncertainty model in section 3.7, detection is absolutely impossible even without a finite dynamic range assumption.

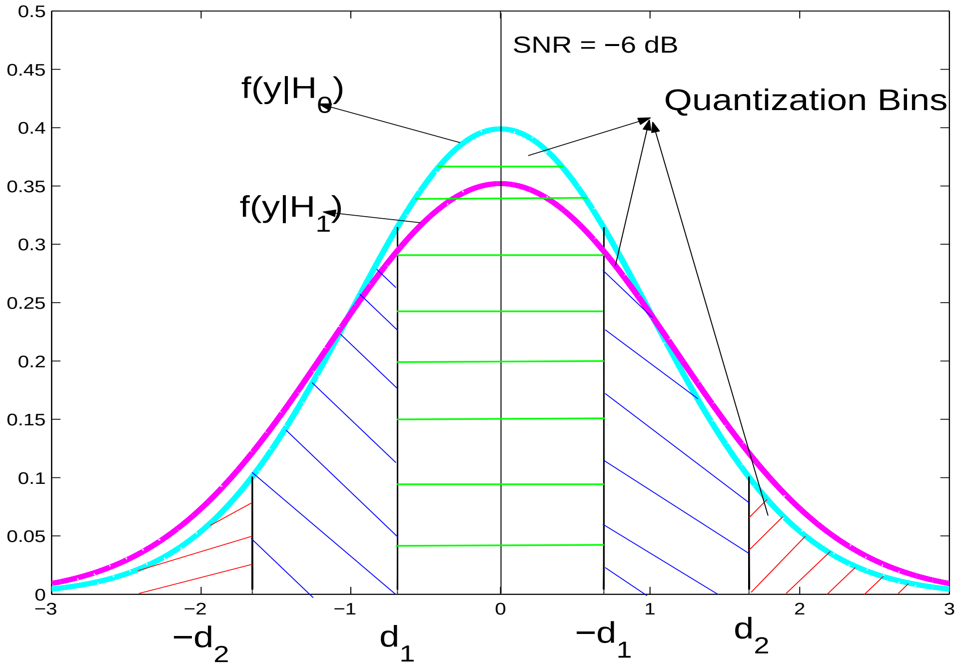


Figure 4.1. 2-bit quantizer

where $Q(.)$ is the standard Gaussian tail probability function, and $p_j^{(i)}$ is the probability mass within bin j under hypothesis $\mathcal{H}_i$, $i, j \in \{0, 1\}$. Since the receiver assumes that the noise variance is σ_n^2, it is clear that $p_1^{(0)} = p_1^{(1)}$, if

$$2Q\left(\frac{d_1}{\sigma_n}\right) = \left[Q\left(\frac{d_1 - \sqrt{P}}{\sigma_a}\right) + Q\left(\frac{d_1 + \sqrt{P}}{\sigma_a}\right)\right]$$

or if,

$$2Q\left(\frac{d_1}{\sigma_a}\right) = \left[Q\left(\frac{d_1 - \sqrt{P}}{\sigma_n}\right) + Q\left(\frac{d_1 + \sqrt{P}}{\sigma_n}\right)\right]$$

In the former case, the receiver thinks that there is no signal when there the signal is actually present and hence $P_{MD} = 1$. In the latter case, the receiver thinks that there is a signal present when it is actually absent and hence $P_{FA} = 1$.

Clearly, under both cases the two hypotheses become identical and hence the signal is undetectable by any detector. Fig. 4.2 shows the performance of a 2-bit quantizer under noise uncertainty and the variation of the location of the wall with the noise uncertainty x. Comparing fig. 4.2 with the performance curve for the radiometer in fig. 3.3 we see that the location of the SNR walls are almost the same. It is important to note that the walls depicted in Fig. 3.3 were specifically for radiometers, and hence other higher moment detectors could possibly overcome them. However, the walls for the quantized detector are absolute; the signal is completely undetectable with enough noise uncertainty.

We now consider the case when there are more than 2 quantization bins. However, as we discussed before, we assume that the receivers have a finite dynamic range of operation. We show that the impossibility result in the above example continues to hold even with more number of quantization bins. In fact, we show that this result holds even in the limit of infinite number of quantization bins, i.e., the receiver has access to the unquantized signal, but has access to the signal only within the finite dynamic range. We prove this result under our minimalist model for noise uncertainty.

4.3 Finite dynamic range: absolute walls

We now consider the detection problem in (3.2) with the added constraint that all the observations are limited to within some finite dynamic range, $(-M, M)$. Under this assumption, we prove the following result:

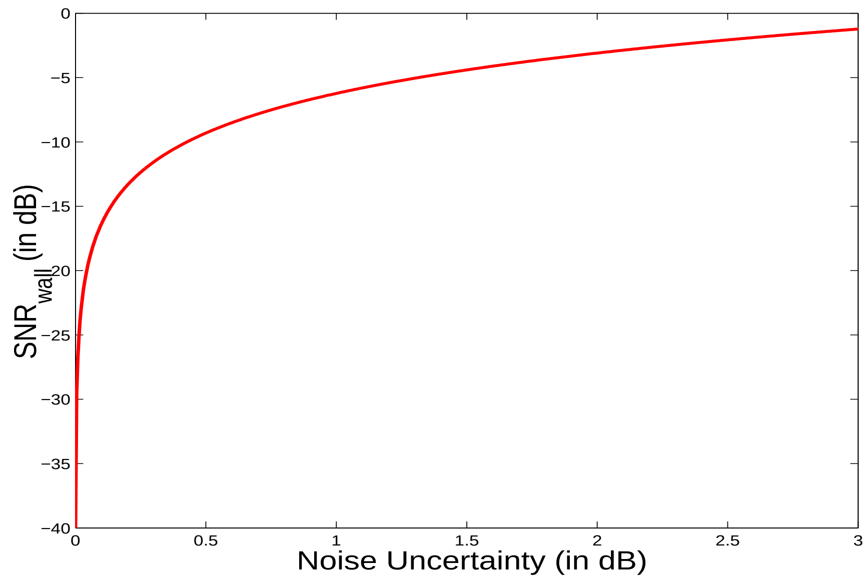
Theorem *3:* For detection of an unknown BPSK signal under the noise uncertainty model described in section 3.3 and with receivers constrained to operate within a finite dynamic range $(-M, M)$, there exists a single snr^*_{wall} for *any possible* detector.

Proof: Note that $W_n \sim N(0, \sigma_n^2)$ and is independent of X. Therefore, the random variable $W_n + X$ must have a density function. The detector is unable to detect signals iff there exits a pair of noise random variables W_a and W_n in the set $\mathcal{W}_x$ such that $f_{W_n}(w) = f_{W_a+X}(w)$ or $f_{W_a}(w) = f_{W_n+X}(w)$, for $w \in (-M, M)$. To prove this result we construct random variables $W_a \in \mathcal{W}_x$ satisfying the conditions above.

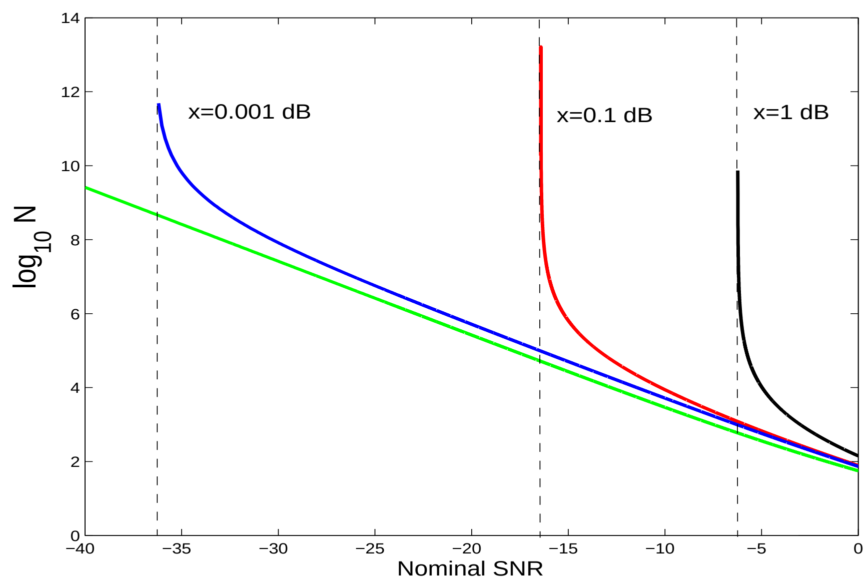
Case I:

$$f_{W_a}(w) = f_{W_n+X}(w), \quad \text{for} \;\; w \in (-M, M) \tag{4.2}$$

We need the noise to have the same density as f_{W_n+X} in $(-M, M)$. It is clear that if



(a)



(b)

Figure 4.2. The plot on the top shows the variation of the SNR_{wall} with noise uncertainty x and the plot on the bottom is the sample complexity of the optimal detector for a 2-bit quantizer, with three different values of the noise uncertainty x. Note, that even at a very low noise uncertainty of $x = 0.1$ dB, the SNR_{wall} is at a relatively high value of -17 dB.

we don't restrict this equality to within this finite dynamic range, then the ratio of the moments $\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \rightarrow \infty$ as $k \rightarrow \infty$ and hence the constructed distribution will fall outside the uncertainty class. Hence, we equate the densities exactly within the finite dynamic range and try to reduce the moments by adjusting the density outside the finite dynamic range. Also, since the signal is very weak, it is safe to assume that $|X| \ll M$ and hence, $\mathbb{P}(|W_a| > M) > \mathbb{P}(|W_n| > M)$. We complete the construction of W_a by defining

$$f_{W_a}(w) = f_{W_n}(w) + \frac{\lambda}{2}[\delta(w - M) + \delta(w + M)]$$

for $w \in (-M, M)^c$ where the constant $\lambda \triangleq \mathbb{P}(|W_a| > M) - \mathbb{P}(|W_n| > M)$. Fig. 4.3a illustrates the noise random variable W_a's density. Having constructed this random variable, all we have to check now is whether it actually falls within our uncertainty class, i.e., if $\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \leq \alpha$.

For this we start off by observing that

$$\begin{aligned}\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} &= \frac{1}{\mathbb{E}W_n^{2k}}[\mathbb{E}W_a^{2k}\mathrm{I}_{|W_a|<M} + M^{2k}\mathbb{P}(|W_a| = M) \\ &\quad + \mathbb{E}W_a^{2k}\mathrm{I}_{|W_a|>M}] \end{aligned} \tag{4.3}$$

Bound the first two terms in (4.3) by

$$\begin{aligned}\frac{M^{2k}\left[\mathbb{P}(|W_a| < M) + \mathbb{P}(|W_a| = M)\right]}{\mathbb{E}W_n^{2k}} &\leq \frac{M^{2k}}{\mathbb{E}W_n^{2k}} \\ &= \frac{M^{2k}}{(2k-1)!!\ \sigma_n^{2k}} \\ &\longrightarrow 0 \end{aligned} \tag{4.4}$$

as $k \rightarrow \infty$. Since the sequence above tends to zero, there exists some k_{min} such that

$$\frac{\mathbb{E}W_a^{2k}\mathrm{I}_{|W_a|<M} + M^{2k}\mathbb{P}(|W_a| = M)}{\mathbb{E}W_n^{2k}} < (\alpha - 1) \tag{4.5}$$

for all $k \geq k_{min}$. Table 3 lists a few typical values for k_{min} as a function of M. Now, we can write the third term in (4.3) as

$$\frac{\mathbb{E}W_a^{2k}\mathrm{I}_{|W_a|>M}}{\mathbb{E}W_n^{2k}} = \frac{\mathbb{E}W_n^{2k}\mathrm{I}_{|W_n|>M}}{\mathbb{E}W_n^{2k}} \leq 1 \tag{4.6}$$

where the first step is true because $W_a = W_n$ outside the finite dynamic range. Therefore, for $k \geq k_{min}$ we have

$$\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \leq (\alpha - 1) + 1 = \alpha \tag{4.7}$$

where the first term in the inequality follows from the definition of k_{min} in (4.5) and the second term follows from (4.6). This implies that $W_a \in \mathcal{W}_x$ for $k \geq k_{min}$. For $k < k_{min}$, we use the fact that,

$$\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \leq \frac{\mathbb{E}(X + W_n)^{2k}}{\mathbb{E}W_n^{2k}} \tag{4.8}$$

Note, that the above inequality is true for all k, but it gives meaningful results only for $k \leq k_{min}$. From theorem 1, it is clear that $\frac{\mathbb{E}(X+W_n)^{2k}}{\mathbb{E}W_n^{2k}} \leq \alpha$ iff $snr \leq snr_{wall}^{(2k)}$. Since we want this to hold true for all $k \leq k_{min}$, we must have $snr \leq snr_{wall}^{(2k_{min})}$. Using the lower bound in (3.3), we get

$$snr \leq \frac{\alpha - 1}{k_{min}}\left[1 - \frac{\alpha - 1}{2 - \alpha}\right] \tag{4.9}$$

Now, from (4.7) and (4.9) it is clear that if $snr < \frac{\alpha-1}{k_{min}}\left[1 - \frac{\alpha-1}{2-\alpha}\right]$, then $\frac{\mathbb{E}W_a^{2k}}{\mathbb{E}W_n^{2k}} \leq \alpha$ for all k, which implies that $W_a \in \mathcal{W}_x$. Therefore, we have shown that there exists an snr^*_{wall} threshold below which detection is *absolutely impossible*. In fact we have shown that

$$snr^*_{wall} \geq \frac{\alpha - 1}{k_{min}}\left[1 - \frac{\alpha - 1}{2 - \alpha}\right] \tag{4.10}$$

Thus, we encounter absolute walls due to noise uncertainty under finite dynamic range operation.

Case II:

$$f_{W_n}(w) = f_{W_a+X}(w), \quad w \in (-M, M) \tag{4.11}$$

M	k_{min}	SNR_{wall}(x=0.5dB)	k_{min}	SNR_{wall}(x=1dB)
2 σ_n	4	–15.8	3	-12.5
3 σ_n	8	-18.1	6	-13.6
4 σ_n	12	-19.9	10	-15.8
5 σ_n	17	-21.4	15	-17.6
6 σ_n	25	-23.1	22	-19.2

Table 4.1. We list k_{min} as a function of the finite dynamic range M for two different values of the noise uncertainty x. For each value of M, k_{min} has been numerically computed such that $\frac{\mathbb{E}W_a^{2k}\mathrm{I}_{|W_a|<M} + M^{2k}\mathbb{P}(|W_a|=M)}{\mathbb{E}W_n^{2k}} < (\alpha - 1)$. Also, the SNR_{wall} has been computed from equation (4.9) for each value of k_{min}.

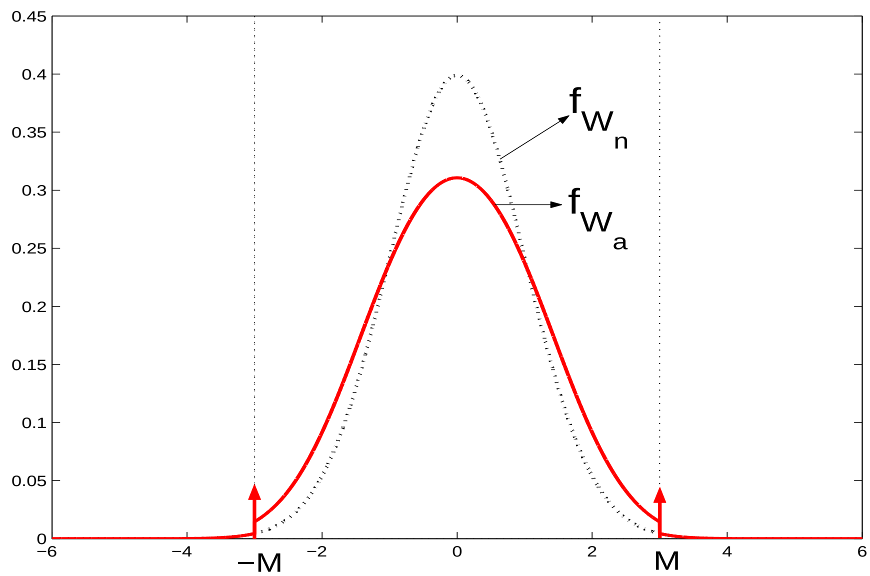
In this case the receiver is fooled to believe that the signal is absent even when the signal is actually present. In this case, we construct a noise random variable $W_a \in \mathcal{W}_x$ satisfying the condition (4.11). Begin by choosing W_a such that $f_{W_n} = f_{W_a+X}$ within $(-M, M)$. In this case the mass in W_a is smaller than the mass in W_n outside the finite dynamic range $(-M, M)$. Therefore, the delta function approach of case I will not work. But, this problem can be overcome by making $f_{W_a}(w) = f_{W_n}(w)$ for $w \in (-M_1, M_1)^c$, where we choose M_1 such that $\mathbb{P}(|W_a| > M_1) = \mathbb{P}(|W_n| > M_1)$. Now, we need to verify that $W_a \in \mathcal{W}_x$. This can be done exactly as in the proof for case I. Therefore as in case I, if $snr < \frac{\alpha-1}{k_{min}}\left[1 - \frac{\alpha-1}{2-\alpha}\right]$, detection is absolutely impossible. ■

4.4 Discussion

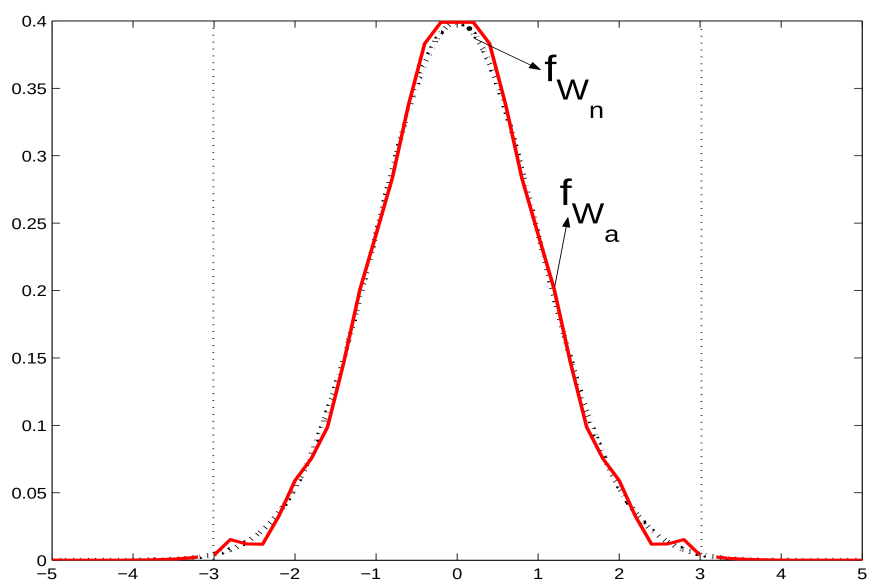
Theorem 3 shows that the seeming gain coming from using higher moment detectors is illusory. This is because higher moment detectors use rare large values of the test statistic for detection, but a finite dynamic range of operation prevents this.

Intuitively, this result comes from the fact that the classes of distributions under both hypotheses overlap due to the finite dynamic range assumption. Therefore, detection is impossible for any possible detector. Note that this kind of absolute impossibility result was not possible in chapter 3 because the distribution classes under our model don't overlap without the finite dynamic range limitation.

Also, we have shown that the actual value of the SNR_{wall} depends on the finite dynamic



(a)



(b)

Figure 4.3. The density functions of the constructed noise random variable in both cases to prove theorem 3 is shown in this figure. The figure on the left corresponds to Case I discussed in section 4.3. Note that there are two delta functions in the density of W_a at $\pm M$. The figure on the right corresponds to case II in section 4.3. In this figure note that there is a small hole in the density function of W_a at $\pm M$. In both these figures, the red curve corresponds to the density function of W_a and the black curve corresponds to the density function of W_n.

range M, as shown in table 3. The values in the table suggests that if we build a receiver with a larger finite dynamic range then the value of the SNR_{wall} decreases. So, it is important to know how large a value can M take. Lets look into this little more carefully.

Recall that the error term in the Central Limit Theorem, decreases as $\frac{1}{\sqrt{N}}$, where N is the number of independent terms constituting noise. We also discussed that the value of N is moderate. Further as M increases we are looking into higher and higher values of the noise, whose probability is equal to the tail probability of the noise. Therefore, it is important to ask ourselves, if we are confident about the values of noise which have such a low probability? And the answer to this question is: we are not confident about events whose probability is comparable to the error term in the CLT. Therefore, it makes sense to consider only those values of M for which the tail probability of noise $O(1/\sqrt{N})$.

Taking the above discussion into consideration, we can easily check that, a reasonable value for M is $2\sigma_n$. This implies that the realistic values of the SNR_{wall} are of the order of -12 dB (see table 3). Hence the limitations proved in this thesis are very realistic.

Chapter 5

Comments and Conclusions

In this thesis we have considered fundamental bounds to the detection of the presence or absence of signals in very low SNR. We discussed the detection problem with and without considering the uncertainty in noise. In the case when the noise statistics are assumed to be completely known to the receiver, we showed that detection of zero-mean random signals without any pilot tones is hard. However, we showed that in the presence of a even a weak pilot tone the sample complexity of the detector decreases considerably.

We also discussed the problem of detection of sinusoidal signals. We show that, while imperfect synchronization does cause us to need more samples, the increase in the number of samples is small compared to the difference in sample complexities of the radiometer and the matched filter. It is a general feature of coherent detection that processing gain gives us robustness to a wide range of uncertainties in the system.

While we have assumed that the signal and the noise are independent of each other and that the noise is white, we assume that we know the distribution of the noise only to within some uncertain set. Even with a minimalist model of the noise uncertainty, we found that this implies a fundamental bound on the detectable SNR if we assume that the radio can only observe the wireless signals and noise to some finite dynamic range. Furthermore, we showed that detection is absolutely impossible even without this finite dynamic range limitation if we assume that the noise has a bigger distributional uncertainty class.

While our arguments have been focused on the BPSK transmission case, theorem 1, 2 and 3 are not limited to this case. It is not hard to extend our proofs to general zero-mean signal constellations with low signal amplitude. In addition, we believe that these results reflect the generic difficulty of the low SNR signal detection problem and not peculiarities of our model. For example, let us consider the case when the signal is no longer white.[1] In that case, it is possible to use feature detectors (see [8], [9]) to exploit the colored nature of the signal. However, in such cases, it will also be natural to consider our noise as being only approximately white. After all, it reflects the sum of many different physical sources of undesired signals, not all of which are white! For example in real life, mixers in receivers are non-linear and hence actual signals from other bands are also mixed into the original signal. For low enough SNR, we suspect the structure brought by signals from other bands will be indistinguishable from the uncertain low level structure of the noise.

These sorts of fundamental bounds make us further appreciate the usefulness and robustness of the coherent signal processing possible whenever the primary signal has known frequency specific training data or pilot tones. Coherent processing enables us to take long averages that bring the SNR up to a reasonable value for detection.

Finally, we must point out that throughout this thesis we have ignored the presence of fading, which is inherent in most of the wireless channels [27]. One expects that the effect of channel fading should not make the problem of detection any easier. Since most of our results are impossibility results, the presence of fading does not make a qualitative difference to the nature of our results.

However, we must be cautious in ignoring the effect of fading in the case of coherent detection, especially in the problem of detecting a weak pilot signal. It is possible that the presence of channel fading can make it very hard for detecting pilot signals. Factors like, Doppler shift, pilot tones from multiple primary users, multi-path fading can have an adverse effect on these weak pilot tones and can make it impossible to detect them. We suspect that, a careful analysis of this problem might lead to similar impossibility results

[1]This just means that we are not going to sample it faster than the Nyquist rate.

for the problem of detection of weak pilot signals. However, the position of these new SNR walls might be lower than the ones for non-coherent detection obtained in this thesis.

References

[1] T. Bayes, "An essay toward solving a problem in the doctrine of chances," *Phil. Trans. Roy. Soc.*, vol. 53, pp. 370–418, 1763.

[2] R. W. Broderson, A. Wolisz, D. Cabric, S. M. Mishra, and D. Willkomm, "(2004) white paper: Corvus: A cognitive radio approach for usage of virtual unlicensed spectrum." [Online]. Available: http://bwrc.eecs.berkeley.edu/Research/MCMA/CR_White_paper_final1.pdf

[3] R. Durrett, *Probability: Theory and Examples*, 3rd ed. Belmont: Duxbury Press, 2004.

[4] A. H. El-Sawy and V. D. VandeLinde, "Robust detection of known signals," *IEEE Trans. Inform. Theory*, vol. 23, pp. 222–229, March 1977.

[5] FCC, "Et docket no. 03-237," November 2003. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-289A1.pdf

[6] ——, "Et docket no. 03-322," December 2003. [Online]. Available: http://hraunfoss.fcc.gov/edocspublic/attachmatch/FCC-03-322A1.pdf

[7] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. Roy. Soc.*, vol. A 222, pp. 309–368, 1922.

[8] W. A. Gardner, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Trans. Commun.*, vol. 40, pp. 149–159, January 1992.

[9] ——, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, pp. 897–906, August 1998.

[10] C. F. Gauss, *Theoria Combinationis Observationum Erroribus Minimis Obnoxia.* Germany: Dieterich: Gottingen, 1823.

[11] N. Hoven, "On the feasibility of cognitive radio," Master's thesis, University of California, Berkeley, 2005.

[12] N. Hoven and A. Sahai, "Power scaling for cognitive radio," in *WirelessCom Symposium on Emerging Networks, Technologies and Standards*, June 2005.

[13] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, pp. 73–101, March 1964.

[14] ——, "A robust version of the probability ratio test," *Ann. Math. Statist.*, vol. 36, pp. 1753–1758, December 1965.

[15] T. Kailath and V. Poor, "Detection of stochastic processes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2230–2259, October 1998.

[16] S. A. Kassam and J. B. Thomas, "Asymptotically robust detection of a known signal in contaminated nongaussian noise," *IEEE Trans. Inform. Theory*, vol. 22, pp. 22–26, January 1976.

[17] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory.* Printice Hall PTR, 1998, vol. 2.

[18] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly gaussian noise," *IEEE Trans. Inform. Theory*, vol. 17, pp. 50–56, January 1971.

[19] D. Middleton, "On the detection of stochastic signals in additive normal noise - part i," *IEEE Trans. Inform. Theory*, vol. 3, pp. 86–121, June 1957.

[20] I. J. Mitola, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerosp. Electron. Syst. Mag*, vol. 8, pp. 25–36, April 1993.

[21] J. Neyman and E. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. Roy. Soc.*, vol. A 231, no. 9, pp. 492–510, 1933.

[22] R. Price and N. Abramson, "Detection theory," *IEEE Trans. Inform. Theory*, vol. 7, pp. 135–139, July 1961.

[23] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Forty-Second Allerton Conference on Communication, Contol and Computing*, September 2004.

[24] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[25] D. Slepian, "Some comments on the detection of gaussian signals in gaussian noise," *IEEE Trans. Inform. Theory*, vol. 4, pp. 65–68, June 1958.

[26] A. Sonnenschein and P. M. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 654–660, July 1992.

[27] D. Tse and P. Vishwanath, *Fundamentals of wireless communications.* Cambridge University Press, 2005.

[28] J. W. Tukey, *A survey of sampling from contaminated distributions*, ser. Essays in Honor of Harold Hotelling, I. Olkin et al, Stanford and Calif., Eds. Stanford Univ. Press, 1960.

[29] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, 1967.

Appendix A

Appendix: Sample complexity of BPSK detector

A.0.1 Detection Performance using the Central Limit Theorem

From (2.14), we have

$$\begin{aligned} T(\mathbf{Y}) &= \sum_{i=0}^{N-1} \ln \left[\exp \frac{\sqrt{P}Y[i]}{\sigma^2} + \exp \frac{-\sqrt{P}Y[i]}{\sigma^2} \right] \\ &=: \sum_{i=0}^{N-1} Y_i \end{aligned} \tag{A.1}$$

where

$$Y_i := \ln \left[\exp \frac{\sqrt{P}Y[i]}{\sigma^2} + \exp \frac{-\sqrt{P}Y[i]}{\sigma^2} \right]$$

and

$$Z_i := \left[\exp \frac{\sqrt{P}Y[i]}{\sigma^2} + \exp \frac{-\sqrt{P}Y[i]}{\sigma^2} \right]$$

Thus,

$$Y_i = f(Z_i) = \ln Z_i$$

Observe that Y_i, $i = 0, 1, \ldots, N-1$ are independent, identically distributed random variables and the test statistic $T(\mathbf{Y})$ is a sum of independent, identically distributed random variables (see (A.1)). Hence by the central limit theorem for large N,

$$\frac{T(\mathbf{Y}) - Nm_i}{\sqrt{N}} \sim \mathcal{N}\left(0, \sigma_i^2\right)$$

where $m_i := \mathbb{E}(Y_i)$ and $\sigma_i^2 = \mathrm{Var}(Y_i)$. In order to use the CLT we will have to find the mean and variance of the random variable Y_i. Unfortunately, we cannot express m_i and σ_i^2 in closed form. Hence we use the Taylor series expansion of the function $Y_i = f(Z_i)$ at the mean of Z_i.

$$Y_i = f[\mu_{Z_i}] + [Z_i - \mu_{Z_i}]f'(Z_i)|_{\mu_{Z_i}} + \frac{1}{2}[Z_i - \mu_{Z_i}]^2 f''(Z_i)|_{\mu_{Z_i}} + \ldots$$

where $\mathbb{E}(Z_i) =: \mu_{Z_i}$. We use only the first order terms to get

$$\begin{aligned} m_i &\approx f(\mu_{Z_i}) \\ \sigma_i^2 &\approx \left[f'(Z_i)|_{\mu_{Z_i}}\right]^2 Var(Z_i) \end{aligned} \tag{A.2}$$

Using the above approximations we evaluate the mean and variance of Y_i under both the hypothesis.

1. Under $\mathcal{H}_0$:

$$Y_i = \ln\left[\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right]$$

where $Y[i] \sim \mathcal{N}(0, \sigma^2)$. Thus, using (A.2) we have

$$\begin{aligned} \mathbb{E}(Y_i) = m_i &\approx f(\mu_{Z_i}) \\ &= \ln\left(\mathbb{E}\left[\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right]\right) \\ &= \ln\left(2\mathbb{E}\left[\exp\frac{\sqrt{P}Y[i]}{\sigma^2}\right]\right) \\ &= \ln\left(2\exp(\frac{1}{2}\frac{d^2}{\sigma^2})\right) \\ &= \ln\left(2\exp(\frac{snr}{2})\right) \end{aligned} \tag{A.3}$$

Again from (A.2), we have

$$\begin{aligned} Var(Y_i) = \sigma_i^2 &\approx \left[f'(Z_i)|_{\mu_{Z_i}}\right]^2 Var(Z_i) \\ &= \frac{Var(Z_i)}{{\mu_{Z_i}}^2} \end{aligned}$$

We know that

$$Var(Z_i) = \mathbb{E}(Z_i^2) - \mathbb{E}^2(Z_i)$$

Hence, we evaluate each term of the above equation separately.

$$\begin{aligned} \mathbb{E}(Z_i^2) &= \mathbb{E}\left[\left(\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right)^2\right] \\ &= \mathbb{E}\left[\exp\frac{2\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-2\sqrt{P}Y[i]}{\sigma^2} + 2\right] \\ &= 2 + 2\mathbb{E}\left[\exp\frac{2\sqrt{P}Y[i]}{\sigma^2}\right] \\ &= 2 + 2\exp(2snr) \end{aligned}$$

Therefore,

$$\begin{aligned} Var(Z_i) &= 2 + 2\exp(2snr) - 4\exp(snr) && \text{(From } (A.3)) \\ &= 2\left[\exp(snr) - 1\right]^2 \end{aligned}$$

Hence,

$$Var(Y_i) = \sigma_i^2 \approx \frac{2\left[\exp(snr)-1\right]^2}{{\mu_{Z_i}}^2}$$

and,

$$\begin{aligned}
\sigma_i &\approx \frac{\sqrt{2}\left[\exp(snr)-1\right]}{\mu_{Z_i}} \\
&= \frac{\left[\exp(snr)-1\right]}{\sqrt{2}\exp(\frac{snr}{2})} \\
&\approx \frac{1}{\sqrt{2}}\frac{\left[1+snr-snr\right]}{1+\frac{snr}{2}} \\
&\approx \frac{snr}{\sqrt{2}} \qquad \text{(A.4)}
\end{aligned}$$

Here we have used the fact that, at low snr, i.e., when $snr \ll 1$, $\exp(snr) \approx 1 + snr$

2. Under $\mathcal{H}_1$:

$$Y_i = \ln\left[\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right]$$

where $Y[i] \sim \mathcal{N}(a_i, \sigma^2)$ and $a_i \sim$ Bernoulli(1/2). Thus, using the Taylor series approximation in (A.2) we have

$$\begin{aligned}
\mathbb{E}(Y_i) = m_i &\approx f(\mu_{Z_i}) \\
&= \ln\left(\mathbb{E}\left[\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right]\right) \\
&= \ln\left(\mathbb{E}\left[\exp\frac{\sqrt{P}(a_i+\tilde{x}[i])}{\sigma^2} + \exp\frac{-\sqrt{P}(a_i+\tilde{x}[i])}{\sigma^2}\right]\right) \\
&= \ln\left(\mathbb{E}\left[\exp(\frac{\sqrt{P}a_i}{\sigma^2})\right]\mathbb{E}\left[\exp(\frac{\sqrt{P}\tilde{x}[i]}{\sigma^2})\right] + \mathbb{E}\left[\exp(\frac{-\sqrt{P}a_i}{\sigma^2})\right]\mathbb{E}\left[\exp(\frac{-\sqrt{P}\tilde{x}[i]}{\sigma^2})\right]\right) \\
&= \ln\left(2\left[\frac{(\exp(snr)+\exp(-snr))}{2}\right]\exp(\frac{snr}{2})\right) \\
&= \ln\left[\frac{(\exp(snr)+\exp(-snr))}{2}\right] + \ln\left[2\exp(\frac{snr}{2})\right] \qquad \text{(A.5)}
\end{aligned}$$

Where $\tilde{x}[i] \sim \mathcal{N}(0, \sigma^2)$ and independent of a_i and hence we justify writing the expectation as a product in the 4th step of the above series of equations.

We now need to calculate the Variance of Y_i. As before we have

$$\begin{aligned}
Var(Y_i) &\approx \left[f'(Z_i)|_{\mu_{Z_i}}\right]^2 Var(Z_i) \\
&= \frac{Var(Z_i)}{{\mu_{Z_i}}^2}
\end{aligned}$$

and

$$Var(Z_i) = \mathbb{E}(Z_i^2) - \mathbb{E}^2(Z_i) \qquad \text{(A.6)}$$

So, we evaluate each term in the above expression.

$$
\begin{aligned}
\mathbb{E}(Z_i^2) &= \mathbb{E}\left[\left(\exp\frac{\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-\sqrt{P}Y[i]}{\sigma^2}\right)^2\right] \\
&= \mathbb{E}\left[\exp\frac{2\sqrt{P}Y[i]}{\sigma^2} + \exp\frac{-2\sqrt{P}Y[i]}{\sigma^2} + 2\right] \\
&= 2 + 2\mathbb{E}\left[\exp\frac{2\sqrt{P}Y[i]}{\sigma^2}\right] \\
&= 2 + 2\mathbb{E}\left[\exp\frac{2d(a_i + \tilde{x}[i])}{\sigma^2}\right] \\
&= 2 + 2\mathbb{E}\left[\exp\left(\frac{2da_i}{\sigma^2}\right)\exp\left(\frac{2d\tilde{x}[i]}{\sigma^2}\right)\right] \\
&= 2 + 2\mathbb{E}\left[\exp\left(\frac{2da_i}{\sigma^2}\right)\right]\mathbb{E}\left[\exp\left(\frac{2d\tilde{x}[i]}{\sigma^2}\right)\right] \\
&= 2 + 2\left[\frac{\exp(2snr) + \exp(-2snr)}{2}\right]\exp(2snr) \\
&= 2 + [\exp(2snr) + \exp(-2snr)]\exp(2snr) \qquad \text{(A.7)}
\end{aligned}
$$

Substituting (A.7) in (A.6) we get

$$
\begin{aligned}
Var(Z_i) &= 2 + [\exp(2snr) + \exp(-2snr)]\exp(2snr) - \mathbb{E}^2(Z_i) \\
&= 2 + [\exp(2snr) + \exp(-2snr)]\exp(2snr) - \\
&\quad [\exp(snr) + \exp(-snr)]^2\exp(snr) \\
&\qquad \text{(From } (A.5)) \\
&= 3 + \exp(4snr) - \exp(3snr) - \exp(-snr) - 2\exp(snr)
\end{aligned}
$$

Therefore,

$$
Var(Y_i) \approx \frac{3 + \exp(4snr) - \exp(3snr) - \exp(-snr) - 2\exp(snr)}{[\exp(snr) + \exp(-snr)]^2\exp(snr)}
$$

Since we are concerned with the low snr scenario, we again use the fact that, when $snr \ll 1$, $\exp(snr) \approx 1 + snr$ and simplify the above expression to obtain:

$$
\begin{aligned}
Var(Y_i) &\approx snr^2 \\
\Rightarrow \sqrt{Var(Y_i)} &\approx snr \qquad \text{(A.8)}
\end{aligned}
$$

Finally, to summarize we define, $m_k := \mathbb{E}(Y_i|\mathcal{H}_k)$ for $k = 0, 1$ and all $i = 0, 1, \ldots, N-1$. Similarly define, $\sigma_k := \sqrt{Var(Y_i|\mathcal{H}_k)}$ for $k = 0, 1$. Thus, summarizing the results of the previous two section:

$$
\begin{aligned}
m_0 &\approx \ln\left[2\exp(\frac{snr}{2})\right] \\
m_1 &\approx \ln\left[\frac{\exp(snr) + \exp(-snr)}{2}\right] + \ln\left[2\exp(\frac{snr}{2})\right] \\
\sigma_0 &= O(snr) \\
\sigma_1 &= O(snr) \qquad \text{(A.9)}
\end{aligned}
$$

From the above equations we also have

$$
\begin{aligned}
m_1 - m_0 &= \ln\left[\frac{\exp(snr) + \exp(-snr)}{2}\right] \\
&\approx snr^2
\end{aligned}
\tag{A.10}
$$

For $snr \ll 1$. Now, we apply the use these expressions to derive the detection performance. We have

$$
\begin{aligned}
P_{FA} &= Pr\left(T(\mathbf{Y}) > \gamma' | \mathcal{H}_0\right) \\
&= Pr\left(\frac{T(\mathbf{Y} - Nm_0)}{\sqrt{N}\sigma_0} > \frac{\gamma' - Nm_0}{\sqrt{N}\sigma_0} | \mathcal{H}_0\right) \\
&= Q\left(\frac{\gamma' - Nm_0}{\sqrt{N}\sigma_0}\right) \qquad \text{(By CLT)}
\end{aligned}
$$

Simplifying this,

$$
\gamma' = \sqrt{N}\sigma_0 Q^{-1}(P_{FA}) + Nm_0 \tag{A.11}
$$

Similarly,

$$
\begin{aligned}
P_D &= Pr\left(T(\mathbf{Y}) > \gamma' | \mathcal{H}_1\right) \\
&= Pr\left(\frac{T(\mathbf{Y} - Nm_1)}{\sqrt{N}\sigma_1} > \frac{\gamma' - Nm_1}{\sqrt{N}\sigma_1} | \mathcal{H}_1\right) \\
&= Q\left(\frac{\gamma' - Nm_1}{\sqrt{N}\sigma_1}\right) \qquad \text{(By CLT)}
\end{aligned}
$$

and

$$
\gamma' = \sqrt{N}\sigma_1 Q^{-1}(P_D) + Nm_1 \tag{A.12}
$$

Eliminating γ' from (A.11) and (A.12) we have,

$$
N = \left[\frac{\sigma_0 Q^{-1}(P_{FA}) - \sigma_1 Q^{-1}(P_D)}{m_1 - m_0}\right]^2
$$

Using (A.9) and (A.10) in the above equation we have,

$$
\begin{aligned}
N &\approx \left[\frac{snrQ^{-1}(P_{FA}) - snrQ^{-1}(P_D)}{snr^2}\right]^2 \\
&= \left[Q^{-1}(P_{FA}) - Q^{-1}(P_D)\right]^2 snr^{-2} \\
&\approx snr^{-2}
\end{aligned}
\tag{A.13}
$$