SNR walls for signal detection

Rahul Tandra   Anant Sahai

tandra@eecs.berkeley.edu   sahai@eecs.berkeley.edu

Dept. of Electrical Engineering and Computer Sciences, U C Berkeley

Abstract

This paper considers the problem of detecting the presence/absence of signals in low SNR environments. The context is cognitive radios trying to opportunistically share spectrum along with potential primary users that must be detected in order to avoid causing harmful interference. Small modeling uncertainties are unavoidable in any practical system and so robustness to modeling uncertainties is a fundamentally important performance metric.

We propose simple mathematical models for these uncertainties and establish two main results. If we have no knowledge of the modulation scheme of the signal we are trying to detect, then there exists an absolute "SNR wall" below which every detector will fail to be robust, no matter how long the detector can observe the channel. Upper and lower bounds are computed that show any detector is essentially as non-robust as the radiometer. Given knowledge of the signal, we prove that both coherent and cyclostationary feature detectors are also non-robust to model uncertainties when channel coherence times are finite. The scaling of the SNR wall with coherence time is worse than for cyclostationary detectors as compared to coherent detectors. These results strongly suggest that all detectors are non-robust.

This has implications for wireless spectrum regulators. We argue that the tension between primary and secondary users is captured by the technical question of computing the optimal tradeoff between *capacity* and *robustness* as quantified by the SNR wall. This is an open problem, but we compute this tradeoff for some simple detectors.

I. Introduction

In order to reuse underutilized spectrum, the operation of unlicensed secondary devices within licensed primary bands such as "unused" television broadcast bands has been proposed [5]. The main constraint for opportunistic secondary devices/cognitive radios is guaranteeing non-interference to the primary system. One possibility is to align the secondary transmissions such that the resulting interference is mostly orthogonal to the primary's received signal [6] or by having the secondary counteract its own interference in the direction of the primary signals [7]. Such design strategies face several practical problems. The two systems must be jointly engineered and may require

a huge overhead in terms of the amount of coordination between the primary and secondary systems. Furthermore, simple phase uncertainties can significantly lower the performance of such collaborative designs [8].

Another possible strategy is for the secondary to sense the primary signal and opportunistically use the primary channel only if the sensing declares the channel to be vacant. This approach has the advantage of minimal coordination with the primary system and hence is more flexible. This flexibility comes at the cost of lower performance when the primary is present. The opportunistic approach is also fair in the sense that the onus is on the secondary system to sense for the primary if it wants to use the spectrum. However, in order to guarantee non-interference with potentially hidden primary receivers, the cognitive radio system needs to be able to detect the presence/absence of very weak primary signals [3], [9]. The IEEE 802.22 is the first international standards process for a cognitive-radio based PHY/MAC/air interface for use by license-exempt devices on a non-interfering basis in spectrum that is allocated to Television Broadcasts. Under the currently understood requirements, secondary devices (CPEs) must be able to sense TV transmissions as low as -116 dBm ($SNR = -22$ dB) [10], [11].

This paper focuses on the latter approach for designing a cognitive radio system: we assume that the secondary must sense for potential primary users before using the spectrum. The goal is to design detection algorithms that meet a given '*receiver operating characteristic*' (ROC) constraint at very low SNR. Classical detection theory [12], [13] tells us that degradation in the ROC due to reduced SNR can be countered by running the detection algorithm for longer sensing times. Hence, the sensing performance of a system is considered to be limited by sensing time constraints. These constraints usually arise from higher layer design considerations. For instance, the QoS requirements of an application drive the protocol layer design, which in turn dictates the time for sensing by the physical layer. This implies that a cognitive radio system can be engineered at the cost of slightly relaxed QoS requirements.

This paper argues that the degradation in QoS may be complete once the primary SNR that must be detected is low enough. We argue that *model uncertainties exist in any practical system and they impose fundamental limitations on detection performance. These limitations cannot be countered by increasing the sensing duration*. At very low SNRs, the ergodic view of the world is no longer valid, i.e., one cannot count on infinite averaging to combat the relevant uncertainty.

To illustrate the limitations caused by model uncertainties, consider a simple thought experiment. Imagine a cognitive radio running a detection algorithm to sense for the presence/absence of the primary signal. Assume that the detector computes a test-statistic and compares it to a threshold to decide if the primary is present/absent. The threshold is set such that the target false alarm probability (P_{FA}) is met using the nominal noise model. Now, we take this detector and deploy it in several physically distinct scenarios (with their own slight variations in local noise characteristics). An example of such a plot for the energy detector is given in Figure 1. The plot clearly illustrates

the limitations due to the uncertainty in the noise model. The performance achieved when the detector is deployed in the field deviate significantly. In fact, below a certain SNR threshold, at least one of the error probabilities can be worse than $\frac{1}{2}$. We call this sort of failure a lack of *robustness* in the detector. The nominal SNR threshold below which this phenomenon manifests is called the *SNR wall* for the detector. Such robustness limits were first shown in the context of radiometric (energy) detection of spread spectrum signals [14].
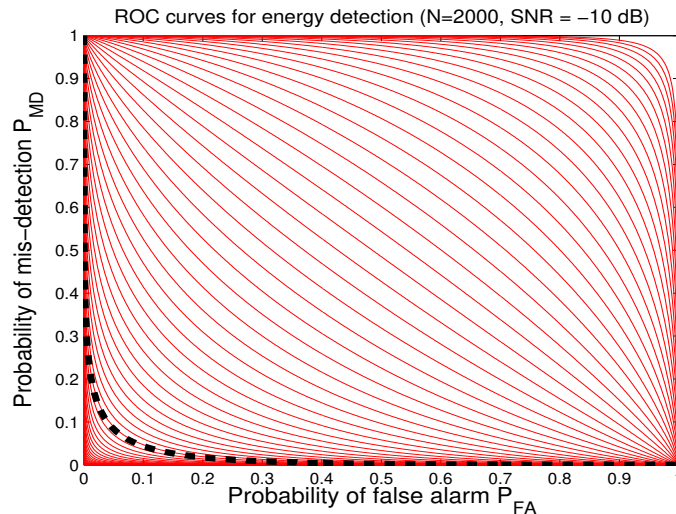

Fig. 1. The dashed (black) line is the theoretical ROC for the energy detector. The solid (red) curves are the ROC curves when this detector is deployed in the field with SNR's varying between -9 dB and -11 dB ($x = 1$ in Sec III-A).

The conjecture that this paper explores is that every detection algorithm suffers from similar robustness limitations in low SNR environments. The problem is posed in Section II and Section III proposes tractable mathematical models for practical wireless system uncertainties in noise and fading. These uncertainty models generalize the simple noise power uncertainty model of [14]. Section IV then explores the robustness of two extreme cases for signal detection — non-coherent detectors and matched filters. The former corresponds to very limited knowledge about the primary signaling scheme whereas the latter corresponds to complete knowledge of the primary's potential signal. Both these detectors suffer from SNR walls under model uncertainties. In particular the best possible non-coherent detector is as non-robust as the energy detector. In the coherent case, the wall is pushed back, but only to an extent limited by the finite coherence time T_c of the fading. These SNR wall effects have been experimentally verified in [15].

It is natural to wonder whether there is anything between the two extremes of nearly total ignorance and complete knowledge. Section V considers the robustness properties of a double-sideband signal to illustrate what is possible when the primary signal is constrained in some way that prevents it from filling all degrees of freedom. Noise calibration is identified as a way to overcome some of the *a priori* model uncertainty. This example also motivates the progression to general cyclostationary feature detection (first proposed by William Gardner [16], [17]) in Section VI. There, we derive one of the main results of this paper — the non-robustness of feature detection to uncertainties

at low SNR when the fading is frequency selective.

A summary of the results is given in Table I. The energy detector makes minimal assumptions on the signal but as a consequence is the least robust to noise uncertainty. This manifests as a stringent SNR wall that is insensitive to the channel coherence time. The matched filter makes maximal assumptions on the signal and hence its SNR wall is lower than that of the energy detector. The location of the SNR wall scales proportional to $\frac{1}{N_c}$, where N_c is the guaranteed coherence time of the channel. Feature detectors occupy an intermediate position in that they make structural assumptions on the signal that quantify how the primary signal does not fully occupy all its potential degrees of freedom. This gives rise to an SNR wall that is lower than that of the energy detector, but which scales only as $\frac{1}{\sqrt{N_c}}$ with the relevant coherence time.

Signal model	Channel model	Strategy	SNR wall location
Unknown signal	Arbitrary and unknown	Energy detection	Energy detection wall (See Section IV-A)
with a power constraint	frequency selective fading	Optimal non-coherent detection	Almost energy detection wall (See Theorem 1)
Unknown signal with	Arbitrary and unknown	Energy detection	Energy detection wall (See Section IV-A)
deterministic component	frequency selective fading	Coherent detection	$\frac{1}{N_c}$ improvement in the SNR wall (See Section IV-B)
Double side band carrier modulation in phase only	Phase changes randomly from one coherence time to another	Noise Calibration using coherence time	$\frac{1}{\sqrt{N_c}}$ improvement in the SNR wall (See Section V-A)
Cyclostationary signal	Flat fading	Feature detection ignoring coherence time	No SNR wall (See Theorem 3)
	Frequency-selective fading	Feature detector, but using coherence time	$\frac{1}{\sqrt{N_c}}$ improvement in the SNR wall (See Theorem 5)

TABLE I

THE ROBUSTNESS RESULTS FOR THE CASES CONSIDERED IN THIS PAPER ARE SUMMARIZED IN THIS TABLE. THE RESULTS (IN THE LAST COLUMN) ARE STATED IN TERMS OF THE LOCATION OF THE SNR WALL.

The robustness results in this paper have significant implications for policymakers. Suppose that the Federal Communications Cooperation (FCC) decides to open a band of spectrum for cognitive use. This must be accompanied by rules governing the band. The rules must be flexible enough to allow interoperability with future technologies while also being compatible with high performance. One can think of a simple and flexible rule: the primary can transmit any signal within a given spectral mask, similarly the secondary should be able to robustly sense the primary with high probability of detection. In this case the primary could potentially use all its degrees of freedom by transmitting a 'white' signal. As shown in Section IV-A, the secondary is then forced to non-coherently detect for the primary, and such detection algorithms are highly non-robust. Complete flexibility for the primary user comes

at a cost of inefficient spectral usage (the secondary systems face SNR limitations and hence will be forced to presumptively shut-up since they cannot robustly detect the absence of potential primary signals). But, if the rules are written such that the primary is mandated to transmit a known pilot tone at a certain power, then the secondary systems can operate more easily. The general tradeoff can be posed technically as the *Capacity robustness tradeoff*. We briefly discuss this tradeoff in Section VII-C for some of the detectors considered here.

II. PROBLEM FORMULATION

Let $X(t)$ denote the band-limited signal we are trying to sense, and let the additive noise be denoted by $W(t)$. Most of this paper focuses on real-valued signals. The analysis easily extends to complex signals. The discrete-time version is obtained by sampling the received signal at the appropriate rate. The sensing problem can be formulated as a binary hypothesis testing problem with the following hypotheses:

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

Here $X[n]$ are the samples of the signal of interest, $W[n]$ are samples of noise and $Y[n]$ are the received signal samples. Throughout this paper, the *signal is assumed to be independent of the noise*. The distribution of $X[n]$ depends on the particular signal model considered in the problem. The noise samples are modeled as having any distribution W from a set of possible distributions $\mathcal{W}_x$. This set is called the noise uncertainty set and a model for the distributions in this set is given in Section III-A.

Any detection strategy/algorithm can be written as a possibly random function $F : \mathbb{R}^N \to \{0, 1\}$, where F maps the N dimensional received vector $\mathbf{Y} = (Y[1], Y[2], \cdots, Y[N])$ onto the set $\{0, 1\}$. Here '0' stands for the decision that the received signal is noise and '1' stands for the decision that the received signal is signal plus noise. For each choice of a noise distribution $W \in \mathcal{W}_x$, the error probabilities are

$$
P_{FA}(W) = \mathbb{E}_W \left[\mathbf{1}_{\{F=1\}} | \mathcal{H}_0 \right]; \qquad P_{MD}(W) = \mathbb{E}_W \left[\mathbf{1}_{\{F=0\}} | \mathcal{H}_1 \right].
$$

- A decision strategy *robustly* achieves a given target probability of false alarm, P_{FA}, and probability of missed detection, P_{MD} if the algorithm satisfies

$$
\max_{W \in \mathcal{W}_x} P_{FA}(W) \leq P_{FA}; \qquad \max_{W \in \mathcal{W}_x} P_{MD}(W) \leq P_{MD}.
$$

- The average signal to noise ratio is defined as

$$
SNR = \frac{P}{\sigma_n^2}, \text{ where } P = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} X[n]^2.
$$

Although the actual noise variance might vary over distributions in the set $\mathcal{W}_x$, we assume that there is a single nominal noise variance σ_n^2 associated with the noise uncertainty set $\mathcal{W}_x$.

- A detection algorithm is *non-robust* at a fixed SNR if the algorithm cannot robustly achieve any pair (P_{FA}, P_{MD}), where $0 \leq P_{FA} < \frac{1}{2}$ and $0 \leq P_{MD} < \frac{1}{2}$, irrespective of the sampling duration N.

- Suppose there exists a signal to noise ratio threshold, SNR_t, such that the detector is non-robust for all $SNR \leq SNR_t$. The maximum value of such a threshold is defined as the SNR *wall* for the detector.

$$SNR_{wall} = \max\{SNR_t\}, \quad \text{s.t., the detector is non-robust for all} \quad SNR \leq SNR_t.$$

An equivalent condition for testing robustness of a detector can be given in terms of the test-statistic $T(\mathbf{Y})$ for the detector. The detector is non-robust iff the sets of means of $T(\mathbf{Y})$ under both hypotheses overlap [18], [19]. Mathematically, for a fixed N, define $A_N := \{\mathbb{E}_W[T(\mathbf{Y})|\mathcal{H}_0] \ : \ W \in \mathcal{W}_x\}$ and $B_N := \{\mathbb{E}_W[T(\mathbf{Y})|\mathcal{H}_1] \ : \ W \in \mathcal{W}_x\}$. The detector is non-robust iff $A_N \cap B_N \neq \emptyset$ for all $N > 0$. Furthermore, if the set of distributions for the observations $\mathbf{Y}$ under both hypotheses overlap then every detection algorithm will be non-robust. This criterion is used throughout this paper to verify if a given detection algorithm is non-robust. The goal of this paper is to analyze detection algorithms and prove the existence of SNR walls.

III. MODELING UNCERTAINTIES

There are aspects of the system that admit statistical models, but it is unrealistic to assume complete knowledge of their parameters to infinite precision. Instead, we assume knowledge of the model to within some bounds and are interested in worst case performance of detection algorithms over the distributions within these bounds.[1] Background noise falls under this category. Noise is an aggregation of various sources like thermal noise, leakage of signals from other bands due to receiver non-linearity, aliasing from imperfect front end filters, interference due to transmissions from licensed users far away, interference from other opportunistic systems in the vicinity, etc. A Gaussian assumption is only an approximation[2] and it is reasonable to assume that there is always some *residual uncertainty* in our noise estimate, even after run-time calibration.

Fading is modeled with the same philosophy. The fading process is considered a linear time varying filter, where the filter coefficients vary on the time-scale given by the channel coherence time. Lacking any further model, we adopt a block-fading perspective that the fading process takes independent realizations after each coherent time.

---

[1]If a Bayesian perspective is taken and a prior is assumed over these parameters, then nothing much changes. The prior turns into a set of possible parameter values once a desired probability of missed detection and false alarm are set. Having a prior just makes the notation and definitions more complicated and so we do not assume priors here for system parameters.

[2]In information theory, it turns out that the Gaussian distribution is a saddle point for many interesting optimization problems. Therefore, we are sometimes safe in ignoring the distributional uncertainty and just overbounding with a worst case Gaussian. However, this is not true in our case.

Furthermore, it is unreasonable to assume that the true channel coherence time is known exactly. All we can hope for are bounds on it.

Finally, uncertainty arises due to *intentional under-modeling* of system parameters. For instance, we can model the primary signal by imposing a cap on its power spectral density, instead of actually modeling its specific signal constellation, waveform, etc. The advantage of under-modeling can be two-fold. Firstly, it leads to less complex systems. Secondly, intentional under-modeling is done to keep the model flexible. For example, a simple power spectral density cap on a chunk of spectrum is advantageous because it gives the user flexibility to choose from a diverse set of signaling strategies. This property is useful for ease of system evolution.

To summarize, we have statistical models for things which we believe to be modeled accurately, and have uncertainty sets for parameters that are not modeled exactly.

A. Noise uncertainty model

To keep the analysis tractable we propose a bounded moment noise uncertainty model. A noise distribution $W_a \in \mathcal{W}_x$ if:

- The noise process is 'white' and symmetric, $\mathbb{E}W_a^{2k-1} = 0, \forall k = 1, 2, \ldots$.
- Even moments of the noise must be close to the nominal noise moments in that $\mathbb{E}W_a^{2k} \in [\frac{1}{\rho^k}\mathbb{E}W_n^{2k}, \rho^k \, \mathbb{E}W_n^{2k}]$, where $W_n \sim \mathcal{N}(0, \sigma_n^2)$ is a nominal Gaussian noise random variable and $\rho = 10^{x/10} > 1$.

The parameter x is used to quantify the amount of uncertainty in the noise power, i.e., we allow for x dB of uncertainty in the noise variance. We believe that our model is reasonable for the following reasons:

- Suppose the nominal noise distribution is believed to be $W_n \sim \mathcal{N}(0, \sigma_n^2)$. Then, any reasonable uncertainty model must include Gaussian distributions with variance slightly different than σ_n^2, i.e., any Gaussian with variance $\sigma^2 \in (\frac{1}{\rho}\sigma_n^2, \rho\sigma_n^2)$, for some $\rho > 1$.
- Background noise also includes weak signals from potential interferers and other random disturbances. This means that a perfectly Gaussian noise uncertainty model is unrealistic. It is reasonable to include other distributions that are close to Gaussian. For mathematical convenience, closeness to Gaussian distributions is expressed in terms of the moments. See Figure 2(a) for a pictorial description of the noise uncertainty set.

B. Fading model and coherence time

The signal to be detected is received over a wireless channel, so it will be faded at the cognitive user. As we are interested in the low SNR regime, we cannot assume that the fading coefficients can be tracked. We model the wireless channel as an L tap fading channel with the fading coefficients h_l, $l = 0, 1, \cdots, L-1$ that are locally constant, but vary on the time scale of the channel coherence time. The actual coherence time is unknown, but is assumed to be lower bounded by T_c (in seconds) or N_c (in number of samples).
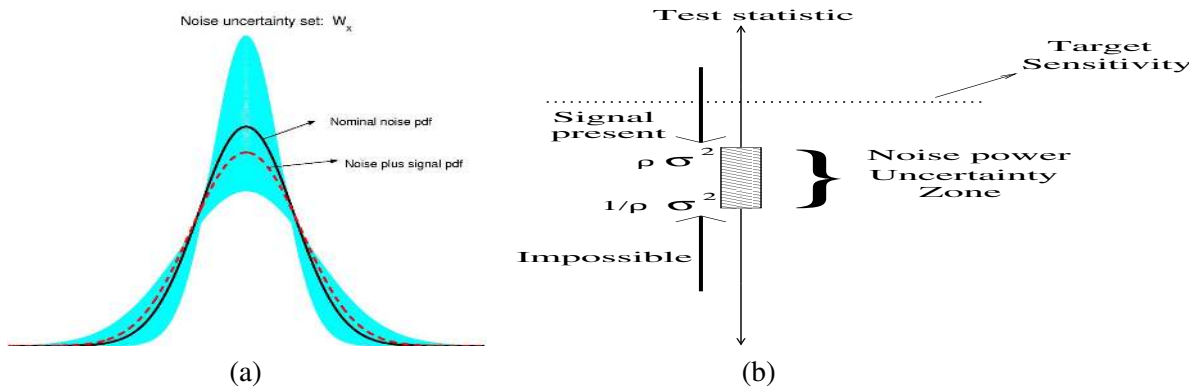
Fig. 2. Figure (a) is an example of a noise uncertainty set $\mathcal{W}_x$, for $x = 10 \log_{10} 2$. The shaded region is the set of all possible pdfs for the noise random variable. The bold curve inside the region is the pdf of a nominal Gaussian, W_n. The dashed curve is the pdf of $X + W_n$ for the case when X is BPSK modulated at $SNR = -6dB$. It is clear from this example that at low signal powers, it is impossible to robustly distinguish the signal from noise. Figure (b) helps understand noise uncertainty in the case of a radiometer. The shaded area in the figure represents the uncertainty in the noise power.

IV. DETECTOR ROBUSTNESS: TWO EXTREME CASES

A. Unknown signal structure

Consider a situation where the cognitive radio detector knows very little about the primary signal. Assume that the primary signaling scheme is unknown, except with a known power within the band of interest. This corresponds to a primary licensee that has absolute freedom to choose its signaling strategy with only the bandwidth and power specified in the license. To robustly detect such a primary user, a single detector must be able to detect the presence of any possible primary signal that satisfies the power and bandwidth constraint.

Under this limited information, the intuitively hardest signal to detect is a zero-mean white signal in the frequency band of interest. Hence, in this section we assume that $X[n]$'s in (1) are zero-mean *iid* random variables taking values from an unknown signal constellation $\mathcal{X}$. To keep the analysis simple, we assume that the signal constellation is finite and has a peak amplitude constraint.

1) Radiometer robustness: It is useful to review the radiometer (energy detector) under noise level uncertainty [14]. The test-statistic is given by $T(\mathbf{Y}) = \frac{1}{N} \sum_{i=1}^{N} Y[n]^2$. Since the radiometer only sees energy, the distributional uncertainty of noise can be summarized in a single interval $[\frac{1}{\rho}\sigma_n^2, \rho\sigma_n^2]$ where σ_n^2 is the nominal noise power. As illustrated in Figure 2(b), it is clear that the radiometer will not be able to robustly detect the signal iff the signal power is less than the uncertainty in the noise power, i.e., $P \leq (\rho - \frac{1}{\rho})\sigma_n^2$.

The presence of the signal is indistinguishable from a slightly larger value for the noise. Hence,

$$SNR_{wall}^{energy} = \frac{\rho^2 - 1}{\rho}. \tag{2}$$

Figure 3 plots the SNR wall for the radiometer as a function of the noise uncertainty and also shows what happens to the number of samples required to meet a performance target in the vicinity of the SNR wall. Extensive

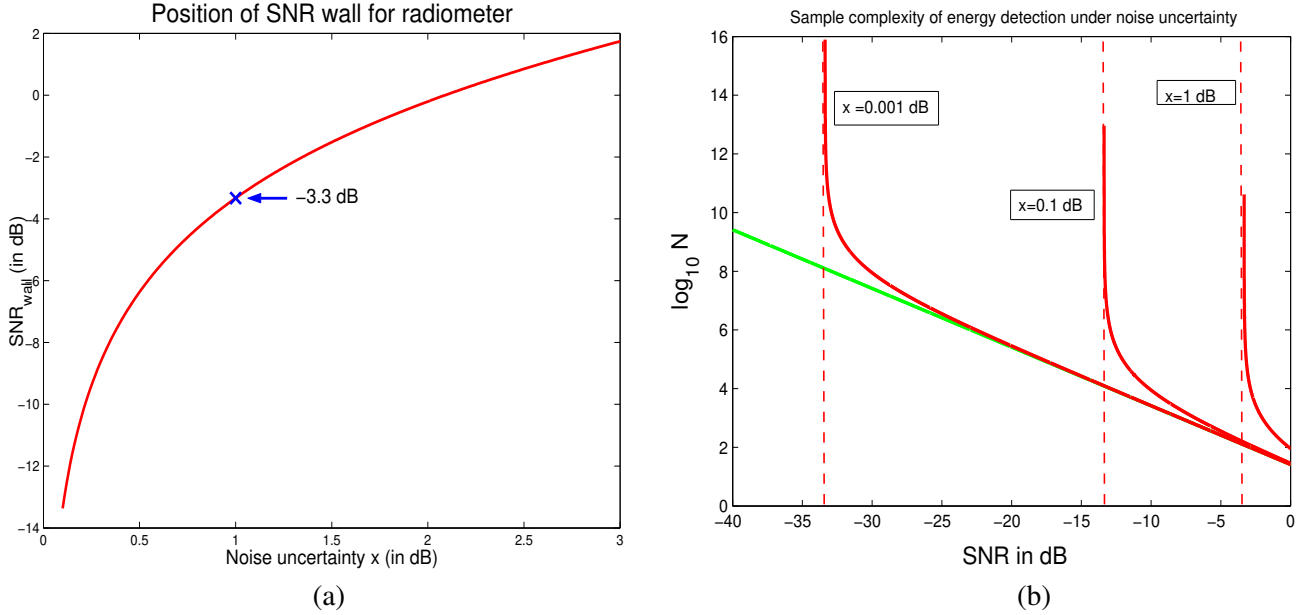simulations results showing the limits on radiometer performance under noise uncertainty have been done in [20].

Fig. 3. Figure (a) plots the SNR_{wall}^{energy} in (2) as a function of noise level uncertainty x where $\rho = 10^{x/10}$. The point marked by a "x" on the plot corresponds to x=1 dB of device level noise uncertainty. This shows that the radiometer is highly non-robust, even to simple device noise level uncertainties. Figure (b) shows how the sample complexity for energy detection varies as the SNR approaches the SNR wall. From the plot we can see that the sample complexity goes to infinity as we approach the SNR wall.

2) Optimal non-coherent detection: It turns out that the radiometer is not unique — all detectors will suffer from an SNR wall under these assumptions.

Theorem 1: Consider the robust hypothesis testing problem defined in Eqn. (1). Assume that the signal samples $X[n]$ satisfy the following properties.

1) $X[n]$ are independent of the noise samples $W[n]$.

2) $X[n]$ are independent and identically distributed as a random variable X taking values in a finite set $\mathcal{X}$.

3) X is zero-mean and all its odd moments are zero, i.e., $\mathbb{E}[X^{2k-1}] = 0$ for all $k = 1, 2, \cdots$

Define $SNR_{peak} = \frac{P_{peak}}{\sigma_n^2}$, where $P_{peak} = \max_{x \in \mathcal{X}} x^2$. Under these assumptions, there exists an SNR threshold SNR_{wall}^* such that robust detection is impossible if $SNR_{peak} \leq SNR_{wall}^*$.

Proof: For $k \geq 1$, define $f(k, SNR)$ to be the following uni-variate degree k polynomial

$$f(k, SNR) := \left[\sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}} \right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)} \right) SNR^i \right]. \qquad (3)$$

Note that all the coefficients of $f(k, SNR)$ are non-negative and hence $f(k, SNR)$ is monotonically increasing in SNR. This means that the equation $f(k, SNR) = \rho^k$ has a unique real valued root $SNR_{wall}^{(2k)}$. Because $f(k, 0) = \frac{1}{\rho^k} < \rho^k$ since $\rho > 1$, the root $SNR_{wall}^{(2k)} > 0$ for all k. Define SNR_{wall}^* to be

$$SNR_{wall}^* := \inf_{k=1,2,3,\cdots\infty} SNR_{wall}^{(2k)}. \qquad (4)$$

Let $W_n \sim \mathcal{N}(0, \sigma_n^2)$, $W_1 \sim \frac{1}{\sqrt{\rho}} W_n$. Let $W_2 := X + W_1$. If $W_1 \in \mathcal{W}_x$, and $W_2 \in \mathcal{W}_x$, then it is clear that robust detection is impossible. We now prove that $W_2 \in \mathcal{W}_x$ by showing that $\mathbb{E}W_2^{2k-1} = 0$ and $\frac{1}{\rho^k} \mathbb{E}W_n^{2k} \leq \mathbb{E}W_2^{2k} \leq \rho^k \mathbb{E}W_n^{2k} \quad \forall\, k$.

Since $\mathbb{E}[X^{2k-1}] = 0$ and $W_2 =: X + W_1$, we have $\mathbb{E}W_2^{2k-1} = 0$ and $\mathbb{E}W_2^{2k} \geq \frac{1}{\rho^k} \mathbb{E}W_n^{2k}$. Therefore, $W_2 \in \mathcal{W}_x$

$$\Leftrightarrow \quad \mathbb{E}W_2^{2k} \leq \rho^k \; \mathbb{E}W_n^{2k}$$

$$\Leftrightarrow \quad \frac{\mathbb{E}(X + W_1)^{2k}}{\mathbb{E}W_n^{2k}} \leq \rho^k$$

$$\Leftrightarrow \quad \sum_{i=0}^{k} \binom{2k}{2k-2i} \frac{\mathbb{E}W_1^{2k-2i}}{\mathbb{E}W_n^{2k}} \mathbb{E}X^{2i} \leq \rho^k$$

$$\Leftrightarrow \quad \sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}}\right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)\; \sigma_n^{2k-2i}}{1 \cdot 3 \cdot 5 \cdots (2k-1)\; \sigma_n^{2k}}\right) \mathbb{E}X^{2i} \leq \rho^k$$

$$\Leftrightarrow \quad \sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}}\right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}\right) \frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}} \leq \rho^k \quad \forall\, k = 1, 2, \cdots \qquad (5)$$

So, we just need to show that the last inequality in (5) is true.

By definition of 'peak SNR', we have $\frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}} \leq SNR_{peak}^i$. Now, if $SNR_{peak} \leq SNR_{wall}^*$, by the monotonicity of $f(k, SNR)$, we have

$$\sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}}\right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}\right) \frac{\mathbb{E}X^{2i}}{\sigma_n^{2i}}$$

$$\leq \quad \sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}}\right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}\right) SNR_{peak}^i$$

$$= \quad f(k, SNR_{peak})$$

$$\leq \quad f(k, SNR_{wall}^*) = \rho^k.$$

which is the last inequality in (5).

∎

3) Bounds on SNR_{wall}^*: Theorem 1 asserts that there exists an SNR threshold SNR_{wall}^* below which robust detection is absolutely impossible. However, the result is vacuous if the infimum in (4) is zero. Recall that $SNR_{wall}^{(2k)}$ is the root of the polynomial $f(k, SNR) - \rho^k$. The coefficient of SNR^i in $f(k, SNR) - \rho^k$ is $\binom{2k}{2k-2i} \left(\frac{1}{\rho^{k-i}}\right) \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}\right)$. Replace $\left(\frac{1}{\rho^{k-i}}\right)$ by 1 to get a new polynomial, $\widetilde{f}(k, SNR) - \rho^k$. Since $\rho > 1$, we have $\left(\frac{1}{\rho^{k-i}}\right) \leq 1$ for all $i = 1, 2, \cdots, k$, and hence the unique positive root of $\widetilde{f}(k, SNR) - \rho^k = 0$ must be smaller than $SNR_{wall}^{(2k)}$. Call this root $\widetilde{SNR}_{wall}^{(2k)}$.

This gives a lower bound on SNR_{wall}^*,

$$SNR_{wall}^* \geq \inf_{k=1,2,\cdots,\infty} \widetilde{SNR}_{wall}^{(2k)}. \qquad (6)$$

where $\widetilde{SNR}_{wall}^{(2k)}$ is the unique positive root of

$$\widetilde{f}(k, SNR) := \left[\sum_{i=0}^{k} \binom{2k}{2k-2i} \left(\frac{1 \cdot 3 \cdot 5 \cdots (2k-2i-1)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}\right) SNR^i\right] = \rho^k. \tag{7}$$

Numerical calculations for this family of polynomials reveal that the $\inf_{k=1,2,\cdots,\infty} \widetilde{SNR}_{wall}^{(2k)} = \widetilde{SNR}_{wall}^{(2)} = \rho - 1$. See [21] for the details. Combining the lower bound with the energy detection wall,

$$\rho - 1 \leq SNR_{wall}^* \leq \frac{\rho^2 - 1}{\rho}. \tag{8}$$

B. Unknown signal structure + deterministic pilot

The other extreme signal detection scenario is when the entire signal $X(t)$ is known. Equivalently, assume that the primary signal allocates a fraction of its total power to transmit a known and deterministic pilot tone and just focus on that part for detection. This model covers many practical communication schemes which use pilot tones/training sequences for data frame synchronization. For example, a Digital Television (ATSC) signal has a pilot tone that is -11 dB weaker than the average signal power [22].

For simplicity, assume the primary's signal is given by $\sqrt{\theta}X_p[n] + \sqrt{(1-\theta)}X_d[n]$. Here $X_p[n]$ is a known pilot tone with θ being the fraction of the total power allocated to the pilot tone and $X_d[n]$, $W[n]$ are zero-mean *iid* processes as before. The test-statistic for the matched filter is given by $T(\mathbf{Y}) = \frac{1}{\sqrt{N}}\sum_{n=1}^{N} Y[n]\widehat{X}_p[n]$, where $\widehat{\mathbf{X}}_p$ is a unit vector in the direction of the pilot tone. In the case of completely known noise statistics, the matched filter is asymptotically optimal. It achieves a dwell time $N \approx [\mathcal{Q}^{-1}(P_D) - \mathcal{Q}^{-1}(P_{FA})]^2 \ \theta^{-1}SNR^{-1}$ [4].

We now analyze the robustness properties of the matched filter. It is easy to see that the matched filter is robust to noise uncertainties alone. Notice that $\mathbb{E}[T(\mathbf{Y})|\mathcal{H}_0] = 0$ and $\mathbb{E}[T(\mathbf{Y})|\mathcal{H}_1] = \frac{1}{\sqrt{N}}\sum_{n=1}^{N} X_p[n]\widehat{X}_p[n] \neq 0$. This shows that the distributional uncertainty classes of $T(\mathbf{Y})$ under both hypotheses do not overlap and hence the detector can robustly achieve any (P_{FA}, P_{MD}) pair for sufficiently large N.

We now incorporate the multipath channel uncertainty model discussed in Section III-B. The primary signal in (1) is now given by $\sum_{l=0}^{L-1} h_l[n]\widetilde{X}[n-l]$, where $\widetilde{X}[n] = \sqrt{\theta}X_p[n] + \sqrt{(1-\theta)}X_d[n]$, and $h_l[n]$ are the multipath fading coefficients. In this case, it is clear that we cannot reap the gains of coherent signal processing forever. As soon as the channel taps assume independent realizations, we can no longer gain from coherent signal processing. By considering time in units of coherence times, it is clear that a good test statistic in the fading channel case is $T(\mathbf{Y}) = \frac{1}{M}\sum_{n=0}^{M-1} \left[\frac{1}{\sqrt{N_c}}\sum_{k=1}^{N_c} Y[n]\widehat{X}_p[nN_c + k]\right]^2$, where N_c is the number of samples in each coherent block and M is the number of coherent blocks over which we can listen for the primary. This detector can be visualized as a combination of two detectors. First, the signal is coherently combined within each coherence time T_c. Coherent processing gain boosts the signal power by N_c while the noise uncertainty is unchanged. Second, the new boosted

signal is detected at the receiver by passing it through a radiometer. The radiometer aspect remains non-robust to noise uncertainties, in spite of the boost in the signal strength. The effective SNR of the coherently combined signal is given by $SNR_{eff} = SNR \cdot \theta \cdot N_c$.

Hence, the modified matched filter will be non-robust if

$$
\begin{aligned}
SNR_{eff} &\leq \left(\frac{\rho^2 - 1}{\rho}\right) \\
\Rightarrow SNR \cdot \theta \cdot N_c &\leq \left(\frac{\rho^2 - 1}{\rho}\right) \\
\Rightarrow SNR_{wall}^{mf} &= \frac{1}{N_c \cdot \theta}\left(\frac{\rho^2 - 1}{\rho}\right).
\end{aligned}
\tag{9}
$$

Coherent processing gains could also be limited due to implementational complexity. The clock-instability of both the sensing radio and the primary transmitter imposes a limit on the coherent processing time. For instance, suppose that there is 1000 Hz of frequency uncertainty and we are doing coherent processing over 1 ms. Since the pilot frequency is uncertain, suppose that we need to search over 4 bins to achieve a target probability of missed detection. If we want to do coherent processing over 10 ms, we now need to search over 40 frequency bins. Since searching over frequency bins involves computing FFTs, it is clear that the coherent processing time is limited by the complexity of the cognitive radio receiver. Finally, coherent processing gains might be limited due to the lengths of the pilots themselves. In packet-based primary systems, the pilots are embedded within each packet and are only coherent for a packet duration since the inter-packet arrival times are usually random on the time-scale of the carrier frequency. No further coherent processing gain is available. Simulations for evaluating the performance of coherent detectors under noise uncertainty have been done in the IEEE 802.22 community. Extensive simulation results evaluating performance of coherent detectors for captured DTV signal data can be found in [23], [24].

V. DETECTOR ROBUSTNESS: NON-WHITE SIGNALS

The extreme of complete ignorance of primary signaling seems excessively conservative. Only counting on the fraction of signal power that is in the known pilots for detection seems like a bad idea, despite the improved robustness. It is natural to want to rule out truly white primary signals, but the model for the primary signal must recognize that the data payload of the signal is likely unknowable to a secondary user facing a low SNR. This is because the primary data rate will likely be far higher than the Shannon capacity of the weak channel between primary and secondary. After all, the secondary user must detect the primary even when it is slightly outside the service area of the primary system [9].

Rather than jump into a general framework, it is useful to consider the simple example of a double sideband modulated primary signal. Our aim is twofold. First, we want to verify if such cases of non-white signals still suffer

from SNR wall limitations. Second, we want to give a qualitative understanding of the nature of SNR walls. We work in continuous time in this section for mathematical convenience.

Consider a real valued random data carrying signal $X(t)$ modulated using a double sideband carrier modulation. Let the carrier frequency be f_c. Since the data is random and unknown, we can assume $X(t)$ to be a wide sense stationary process (WSS) [25]. We also assume that $X(t)$ has zero-mean, i.e., $\mathbb{E}X(t) = 0$.

Our goal is to detect the signal $X(t)\cos(2\pi f_c t + \phi)$ in the presence of a white noise process $W(t)$ whose distribution lies in the uncertainty set $\mathbb{W}_x$, which is the continuous time equivalent of the noise uncertainty set $\mathcal{W}_x$ described in Section III-A. We could use energy detection to distinguish the signal $X(t)\cos(2\pi f_c t + \phi)$ from noise. However, we would have to live with the SNR wall for energy detection from Section IV-A. Another possible detection strategy is *noise calibration*.

A. Noise calibration

The idea is to exploit the fact that there is no signal component in the quadrature phase. First consider the case when there is no unknown phase offset, i.e., $\phi = 0$. In this case we can get "perfect" noise calibration by estimating the noise variance in the 'quadrature phase' and using it as a proxy for the noise variance in the 'in phase' component. Clearly, as the sensing time $T \to \infty$, the noise variance estimation error goes to zero and hence robust sensing is possible at arbitrarily low signal to noise ratios.

Now, consider the case when there is a constant but unknown phase offset ϕ. The detection algorithm must search over ϕ:

1) Mix the received signal onto $\cos(2\pi f_c t + \theta)$ to get the candidate in-phase component and onto $\sin(2\pi f_c t + \theta)$ to get the candidate quadrature phase component.

2) Low-pass filter the two signals obtained in the previous step to the appropriate bandwidth.

3) Square the output of the low pass filters.

4) Time average the result.

5) Maximize the difference of the time average of the 'in phase' and quadrature components over θ and compare the maximum with the detector threshold.

First consider hypothesis $\mathcal{H}_0$, i.e., $Y(t) = W(t)$. In this case, the 'in phase' and 'quadrature phase' signal after step 2 of the algorithm can be written as $N_1(t)\cos(\theta) + N_2(t)\sin(\theta)$ and $N_2(t)\cos(\theta) - N_1(t)\sin(\theta)$ respectively. Here, $N_1(t)$ and $N_2(t)$ are independent White Gaussian noise processes with, $\mathbb{E}[N_1(t)^2] = \frac{\sigma_n^2}{2}$. Now, the test statistic under hypothesis $\mathcal{H}_0$ is given by

$$T(\mathbf{Y})|\mathcal{H}_0 = \max_{0 \le \theta \le \pi} \left[\frac{1}{T} \left(\int_0^T [c(\theta)N_1(t) + s(\theta)N_2(t)]^2 \ dt - \int_0^T [c(\theta)N_2(t) - s(\theta)N_1(t)]^2 \ dt \right) \right], \quad (10)$$

where $c(\theta) := \cos(\theta)$ and $s(\theta) := \sin(\theta)$.

Lemma 1:

$$\mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_0\right] = O\left(\frac{\sigma_n^4}{T}\right). \tag{11}$$

Proof: The proof is given in Appendix I. ∎

From Lemma 1, we have $\lim_{T\to\infty} \mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_0\right] = 0$. We maximize over θ to get the best phase explanation for ϕ. In particular if we set $\theta = \phi$, we get a lower bound for the test-statistic under hypothesis $\mathcal{H}_1$. So,

$$\begin{aligned}
\mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_1\right] \geq \;& \frac{1}{4}\mathbb{E}\left[\int_0^T X^2(t)\; dt\right] \\
+ \;& \mathbb{E}\left[\frac{1}{T}\left(\int_0^T [c(\phi)N_1(t) + s(\phi)N_2(t)]^2\; dt - \int_0^T [c(\phi)N_2(t) - s(\phi)N_1(t)]^2\; dt\right)\right].
\end{aligned}$$

Taking limits, $\lim_{T\to\infty} \mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_1\right] \geq \frac{1}{4}\mathbb{E}\left[\int_0^T X^2(t)\; dt\right] = \frac{P}{2} \neq 0$. Therefore, when the phase ϕ is an unknown constant, our detector is robust to noise power uncertainties and there is no SNR wall.

Now, assume that the coherence time is finite and the phase offset ϕ changes randomly across coherence times. Clearly, if one ignores the coherence time and averages over $T \gg T_c$ then $\mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_1\right] \to 0$ as $T \to \infty$. A simple way to avoid this problem is to compute the maximum over the phase offset θ within each coherence time and then average this test statistic over multiple coherence times. Denote this modified test-statistic by $\widetilde{T}(\mathbf{Y})$.

Since we are taking a maximum within each coherence time, $\lim_{T\to\infty} \mathbb{E}\left[\widetilde{T}(\mathbf{Y})|\mathcal{H}_0\right] \neq 0$, i.e., the mean of the test statistic is no longer zero under hypothesis $\mathcal{H}_0$. In fact, using the results from Lemma 1 it is straightforward to show that

$$\lim_{T\to\infty} \mathbb{E}\left[\widetilde{T}(\mathbf{Y})|\mathcal{H}_0\right] = O\left(\frac{\sigma_n^4}{T_c}\right) > 0.$$

Define

$$\Delta(T_c) := \max_{W_1, W_2 \in \mathbb{W}_x} \left(\lim_{T\to\infty} \mathbb{E}_{W_1}\left[\widetilde{T}(\mathbf{Y})|\mathcal{H}_0\right] - \lim_{T\to\infty} \mathbb{E}_{W_2}\left[\widetilde{T}(\mathbf{Y})|\mathcal{H}_0\right]\right),$$

where $\mathbb{E}_W[.]$ is the expectation when the noise process has a distribution $W \in \mathbb{W}_x$. As in the case of the radiometer, detection is impossible if $\lim_{T\to\infty} \mathbb{E}\left[\widetilde{T}(\mathbf{X})\right] \leq \Delta(T_c)$. This leads to an SNR wall below which detection is impossible. Furthermore, note that the uncertainty in the mean under hypothesis $\mathcal{H}_0$ varies as $\frac{\sigma_n^4}{T_c}$, hence for a fixed signal power P, σ_n^2 must increase like $\sqrt{T_c}$ to keep the SNR wall location constant. This implies for a fixed P and σ_n^2, the SNR wall decays to zero as $O(\frac{1}{\sqrt{T_c}})$.

VI. Feature detection

The example considered in section V is a special case of a class of signals that exhibit features which can be exploited for signal detection. In this section, we consider the general class of cyclostationary feature detectors and analyze their robustness properties. These detectors were first proposed by William Gardner and can cover the double sideband modulation signal as a special case [16], [17].

Until now we have considered stationary random processes as models for the signals to be detected. However, it can be argued that for the purposes of signal detection, the signal of interest is better modeled as a cyclostationary random process. Cyclostationary processes are random processes whose statistical properties vary periodically with time [25]. The data content of the modulated signal is unknown and is usually modeled as a discrete-time stationary random process modulated onto pulse trains like raised cosine pulses [26]. Suppose

$$X(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT_0), \tag{12}$$

where a_n is *iid* random data modulated onto the time-limited pulse $p(t)$ and T_0 is the inter-symbol time. This stationarity combined with the periodicity of pulse trains gives a cyclostationary model.

A. Cyclostationary signal model

We first review the basic properties of cyclostationary random processes. The notation used here is same as in [16]. Assume that the signal of interest, $X(t)$ is a complex cyclostationary process. Let

$$m_X(t) \quad := \quad \mathbb{E}[X(t)], \qquad R_X(t_1, t_2) := \mathbb{E}\left[X(t_1)X^*(t_2)\right].$$

Then $m_X(t)$ and $R_X(t + \frac{\tau}{2}, t - \frac{\tau}{2})$ are periodic in t with the same period T_0. This immediately implies that $R_X(t + \frac{\tau}{2}, t - \frac{\tau}{2})$ admits a Fourier series representation given by

$$R_X(t + \frac{\tau}{2}, t - \frac{\tau}{2}) = \sum_{\alpha} R_X^{\alpha}(\tau) e^{i2\pi\alpha t},$$

where $\alpha = \frac{k}{T_0}$, $k = 0, 1, 2, \ldots$.. Assume that the signal $X(t)$ is zero-mean, i.e., $m_X(t) = 0$. The Fourier coefficient $R_X^{\alpha}(\tau)$, which depends on the lag parameter τ, is called the *cyclic autocorrelation function*, and α is called the cycle frequency. For $\alpha = 0$, it reduces to the conventional autocorrelation function $R_X^0(\tau)$. One can visualize $R_X^0(\tau)$ as the DC component of the lag-product waveform $X(t + \frac{\tau}{2})X^*(t - \frac{\tau}{2})$ for each value of τ, and $R_X^{\alpha}(\tau)$ can be thought of as the AC component corresponding to the frequency α. The Fourier transform of $R_X^{\alpha}(\tau)$ given by

$$S_X^{\alpha}(f) := \int_{-\infty}^{\infty} R_X^{\alpha}(\tau) e^{-i2\pi f\tau} \, d\tau$$

is called the *spectral correlation function*. For $\alpha = 0$ this reduces to the conventional power spectral density. For $\alpha \neq 0$, $S_X^\alpha(f)$ can be thought of as the density of the correlation between spectral components at frequencies $f + \frac{\alpha}{2}$ and $f - \frac{\alpha}{2}$. Explicit formulas of $S_X^\alpha(f)$ for many classes of modulated signals can be found in [27]–[29].

Note that "noise" in a practical setting is likely to contain many interfering signals that are themselves cyclostationary. It is likely that any uncertainty set quantifying the degree of non-stationarity of the noise will introduce its own robustness limits. However, the goal here is to study the extent to which simple noise-distribution and fading uncertainties limit robustness and so noise will be modeled as wide-sense stationary. It is easy to show that if $W(t)$ is wide-sense stationary, then its spectral correlation function $S_W^\alpha(f) = 0$ for $\alpha \neq 0$ [25].

B. Single cycle feature detectors

Till now, we have assumed that sampling the received signal at the symbol-rate is sufficient for detection purposes. However, we need to over-sample the received signal in order to reconstruct the features from the signal samples since the features are capturing the *shape* of the pulses onto which the symbols have been modulated. We assume that the sampling rate $f_s > 2 \max\{\alpha, B\}$, where α is the cycle frequency of the feature we are trying to detect, and B is the bandwidth of $X(t)$.

The test statistic for a single cycle feature detector is defined as

$$Z_{sc}(N) := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_Y^\alpha(f, N) \widehat{S}_X^\alpha(f)^* \, df, \tag{13}$$

where $\widehat{S}_X^\alpha(f) := \frac{S_X^\alpha(f)}{\sqrt{\mathcal{E}}}$, with $\mathcal{E} = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |S_X^\alpha(f)|^2 \, df$. Note that $\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |\widehat{S}_X^\alpha(f)|^2 \, df = 1$. Intuitively, the detector (13) is doing matched filtering of the estimated spectral correlation function $\widetilde{S}_Y^\alpha(f)$ with the normalized spectral correlation function $\widehat{S}_X^\alpha(f)$ of the signal we are trying to detect. The empirical estimate of the spectral correlation function, $\widetilde{S}_Y^\alpha(f, N)$, can be computed from the received signal samples $Y[n]$ by [27], [30]

$$\widetilde{S}_Y^\alpha(f, N) = \frac{1}{N} \frac{1}{T} \sum_{n=0}^{N} Y_T\left((2n+1)\frac{T}{2}, f + \frac{\alpha}{2}\right) Y_T^*\left((2n+1)\frac{T}{2}, f - \frac{\alpha}{2}\right), \tag{14}$$

where $Y_T(n, f)$ is a T point DTFT of the signal $Y[n]$ centered at time n, i.e.,

$$Y_T(n, f) = \sum_{u=n-\frac{T}{2}}^{n+\frac{T}{2}} Y[u] e^{-i2\pi\left(\frac{f}{f_s}\right)u}. \tag{15}$$

C. Robustness of single cycle feature detection

Since $X[n]$ is independent of $W[n]$, it can easily be shown that $\mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = \mathbb{E}\left[\widetilde{S}_X^\alpha(f, N)\right] + \mathbb{E}\left[\widetilde{S}_W^\alpha(f, N)\right]$. Furthermore, as the sample size $N \to \infty$, it is easy to show that $\mathbb{E}\left[\widetilde{S}_X^\alpha(f, N)\right] \to S_X^\alpha(f)$ and $\mathbb{E}\left[\widetilde{S}_W^\alpha(f, N)\right] \to$

$S_W^\alpha(f) = 0$ since the noise is stationary. So, we have

$$\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_1\right] = \frac{1}{\sqrt{\mathcal{E}}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |S_X^\alpha(f)|^2 \, df = \sqrt{\mathcal{E}} \neq 0, \text{ and}$$

$$\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_0\right] = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} S_W^\alpha(f) \widehat{S}_X^\alpha(f)^* \, df = 0. \tag{16}$$

By increasing the number of samples N, the set of means of the test-statistic under both hypotheses can always be made disjoint. Hence a single cycle feature detector is robust to noise power level uncertainty alone.

More importantly, we want to answer the following question: Is feature detection robust to uncertainties in the fading process? Incorporating fading, the received signal under hypothesis $\mathcal{H}_1$ (See (1)) is given by $Y[n] = \sum_{l=0}^{\infty} h_l[n]X[n-l] + W[n]$, where $h_l[n]$ are unknown fading filter coefficients that can vary at the scale of the channel coherence time N_c. We assume that these coefficients assume independent realizations after every N_c samples. For simplicity, we first consider the time-invariant case, i.e., $N_c = \infty$.

Theorem 2: Assume $N_c = \infty$, denote $h_l := h_l[n]$, and define $H(f) := \sum_{l=0}^{\infty} h_l e^{-j2\pi \left(\frac{f}{f_s}\right)l}$, where f_s is the sampling frequency. Consider the single cycle feature detector defined in Section VI-B, and assume $\alpha \neq 0$ is fixed. In this case, there exists a choice of a fading response, $H(f)$ satisfying $\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |H(f)|^2 \, df = 1$, such that

$$\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_1\right] = 0,$$

and hence the single cycle feature detector is non-robust if the channel is modeled as completely unknown.

Proof: Let $\widetilde{X}[n] := h_l * X[n]$. Then the spectral correlation function of $\widetilde{X}[n]$ is given by (see Chapter 4 in [27])

$$S_{\widetilde{X}}^\alpha(f) = H(f + \frac{\alpha}{2})H^*(f - \frac{\alpha}{2})S_X^\alpha(f). \tag{17}$$

Since $Y[n] = h_l * X[n] + W[n] = \widetilde{X}[n] + W[n]$, we have $\mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = \mathbb{E}\left[\widetilde{S}_{\widetilde{X}}^\alpha(f, N)\right] + \mathbb{E}\left[\widetilde{S}_W^\alpha(f, N)\right]$. Taking the limits as $N \to \infty$, we get

$$\lim_{N \to \infty} \mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = S_{\widetilde{X}}^\alpha(f, N) + S_W^\alpha(f, N)$$

$$= H(f + \frac{\alpha}{2})H^*(f - \frac{\alpha}{2})S_X^\alpha(f) + S_W^\alpha(f)$$

$$= H(f + \frac{\alpha}{2})H^*(f - \frac{\alpha}{2})S_X^\alpha(f).$$

The second equality in the chain follows from (17) and the last equality is true because $W[n]$ is stationary and $\alpha \neq 0$. Substituting the expression for $\lim_{N \to \infty} \mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right]$ into (13), we get

$$\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_1\right] = \frac{1}{\sqrt{\mathcal{E}}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} H(f + \frac{\alpha}{2})H^*(f - \frac{\alpha}{2})|S_X^\alpha(f)|^2 \, df. \tag{18}$$

Choose $f_0 \in [\frac{-f_s}{2}, \frac{f_s}{2}]$ such that $\int_{-\frac{f_s}{2}}^{f_0} |S_X^\alpha(f)|^2 \, df = \int_{f_0}^{\frac{f_s}{2}} |S_X^\alpha(f)|^2 \, df$. For a fixed $\alpha \neq 0$, define

$$H(f) = \begin{cases} 1 & \text{if } \frac{-f_s}{2} \leq f < f_0 + \frac{\alpha}{2} \\ -1 & f_0 + \frac{\alpha}{2} \leq f < f_0 + \frac{3\alpha}{2} \\ 1 & f_0 + \frac{3\alpha}{2} \leq f < f_0 + \frac{5\alpha}{2} \\ -1 & f_0 + \frac{5\alpha}{2} \leq f < f_0 + \frac{7\alpha}{2} \\ \cdots & \text{alternate till we hit } \frac{f_s}{2} \end{cases}$$

Note that $H(f)$ has unit norm. For this choice of $H(f)$, (18) simplifies to

$$\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_1\right] = \frac{1}{\sqrt{\mathcal{E}}} \left[\int_{-\frac{f_s}{2}}^{f_0} |S_X^\alpha(f)|^2 \, df - \int_{f_0}^{-\frac{f_s}{2}} |S_X^\alpha(f)|^2 \, df \right] = 0.$$

Hence, $\lim_{N \to \infty} \mathbb{E}\left[Z_{sc}(N)|\mathcal{H}_1\right] = 0$ which makes the single cycle detector non-robust. ∎

Discussion: The above theorem shows that if the fading model is completely unknown, then in the worst case there exists a filter that completely kills the desired signal feature. This might seem too pessimistic since there is no reason to believe that the fading is *trying* to destroy the feature. In such a context, it is natural to assume a probabilistic model for the fading. It is easy to see that when the fading is frequency selective so that $H(f + \frac{\alpha}{2})$ and $H(f - \frac{\alpha}{2})$ have independent values, (18) implies that the expected value for the test statistic could have any complex phase and there is no reason to believe that the resulting phase distribution will be anything other than uniform on the unit circle. This means that if we blindly average the feature over several coherent blocks, the test-statistic will tend to zero, destroying robustness even with a probabilistic model for the fading.

The natural question is to find what fading models do not exhibit this phenomenon and what is the minimal fading model that does. In pursuit of this, it is natural to first consider fading that is *not* frequency selective over the band and features of interest.

***Theorem* 3:** Let $Y[n] = h[n]X[n] + W[n]$ under hypothesis $\mathcal{H}_1$ in (1), with $\mathbb{E}[|h[n]|^2] > 0$. Consider a block fading model for $h[n]$. Assume that the channel coherence time is finite, but it satisfies $T \ll N_c \ll N$ (See (15)). Under these assumptions the single cycle feature detector, with test-statistic given in Equation (13) is robust for all values of SNR.

Proof: For simplicity, we assume that $N = K_1 \cdot N_c$, $N_c = K_2 \cdot T$ for some integers $K_1, K_2 > 0$. This is done to avoid edge effects. From Equation (13) we have $\mathbb{E}[Z_{sc}(N)|\mathcal{H}_1] := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] \widehat{S}_X^\alpha(f)^* \, df$.

Under hypothesis $\mathcal{H}_1$, we have $Y_T(n, f) = \sum_{u=n-\frac{T}{2}}^{n+\frac{T}{2}} (h[u]X[u] + W[u]) e^{-i2}$. Using this in (14), we get

$$\widetilde{S}_Y^\alpha(f, N) = \frac{1}{N} \frac{1}{T} \sum_{n=0}^{N} \left(\sum_{u=nT}^{(n+1)T} (h[u]X[u] + W[u]) e^{-i2\pi\left(\frac{f+\frac{\alpha}{2}}{f_s}\right)u} \right) \left(\sum_{u=nT}^{(n+1)T} (h[u]X[u] + W[u]) e^{-i2\pi\left(\frac{f-\frac{\alpha}{2}}{f_s}\right)u} \right)^*$$

$$= \frac{1}{K_1} \sum_{k=0}^{K_1-1} \left[\frac{1}{N_C \cdot T} \sum_{n=kN_c}^{(k+1)N_c-1} \left(\sum_{u=nT}^{(n+1)T} (h[u]X[u] + W[u]) e^{-i2\pi\left(\frac{f+\frac{\alpha}{2}}{f_s}\right)u} \right) \left(\sum_{u=nT}^{(n+1)T} (h[u]X[u] + W[u]) e^{-i2\pi\left(\frac{f-\frac{\alpha}{2}}{f_s}\right)u} \right)^* \right].$$

The last equation is derived by splitting the summation over n into K_1 different summations over each coherence time. By taking an expectation over the last equation, the cross terms involving noise and signal vanish giving:

$$\mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = \frac{1}{K_1} \sum_{k=0}^{K_1-1} \left(\mathbb{E}\left[|h_k|^2 \widetilde{S}_{X^{(k)}}^\alpha(f, N_c)\right] + \mathbb{E}\left[\widetilde{S}_{W^{(k)}}^\alpha(f, N_c)\right]\right).$$

Here h_k is used to denote the fading coefficient in the kth coherence time, and $\widetilde{S}_{X^{(k)}}^\alpha(f, N_c)$ is the estimate of the spectral correlation function of $X[n]$ (defined in (14)) using the samples within the kth coherence times, i.e, using $X[n]$, $n = k \cdot N_c, \ldots, (k+1)N_c - 1$. $\widetilde{S}_{W^{(k)}}^\alpha(f, N_c)$ is similarly defined for noise. Note that in rearranging the previous equation, we have used the fact that each DTFT $Y_T(n, f)$ uses signal terms from the same coherence block, and hence their fading coefficients are the same. Now, it is straightforward to verify that

$$\widetilde{S}_X^\alpha(f, N) = \frac{1}{K_1} \sum_{k=0}^{K_1-1} \widetilde{S}_{X^{(k)}}^\alpha(f, N_c), \quad \widetilde{S}_W^\alpha(f, N) = \frac{1}{K_1} \sum_{k=0}^{K_1-1} \widetilde{S}_{W^{(k)}}^\alpha(f, N_c).$$

Using this and that the fading process is independent of noise and *iid* across coherence times, we get

$$\mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = \mathbb{E}[|h|^2]\mathbb{E}\left[\widetilde{S}_X^\alpha(f, N_c)\right] + \mathbb{E}\left[\widetilde{S}_W^\alpha(f, N_c)\right],$$

which implies that

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] &= \mathbb{E}[|h|^2] \cdot \mathbb{E}\left[S_X^\alpha(f)\right] + \mathbb{E}\left[S_W^\alpha(f)\right] \\
&= \mathbb{E}[|h|^2] \cdot \mathbb{E}\left[S_X^\alpha(f)\right].
\end{aligned}$$

Hence, $\lim_{N\to\infty} \mathbb{E}[Z_{sc}(N)|\mathcal{H}_1] = \frac{1}{\sqrt{\mathcal{E}}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbb{E}[|h|^2] \cdot |S_X^\alpha(f)|^2 \, df > 0$. This shows that the set of means under both hypothesis can be separated for sufficiently large N, and hence the detector is robust irrespective of the SNR. ∎

At one level, this is somewhat surprising since the matched-filter was not robust to this kind of flat-fading uncertainty. So clearly, it is frequency selectivity that limits robustness for feature detection. The simplest frequency-selective filter is a delay. So let the received signal under hypothesis $\mathcal{H}_1$ be $Y[n] = X[n - l[m]] + W[n]$. Define D_c to be the *delay coherence time* over which the delay $l[m]$ of the wireless channel assumes an independent realization.

Theorem 4: Let the received signal under hypothesis $\mathcal{H}_1$ in (1) be given by $Y[n] = Y[n] = X[n - l[m]] + W[n]$, with $X[n]$ a cyclostationary signal, and $W[n]$ stationary noise independent of the signal. Assume that the distribution of noise arises from the noise uncertainty set $\mathcal{W}_x$. Suppose the delay fading satisfies:

- $l[m]$ is constant within a delay coherence block and *iid* across different delay coherence blocks.
- $l[m]$ is uniformly distributed in $[0, L]$ for some $L = \frac{jf_s}{\alpha}$, f_s is the sampling frequency and some integer $j > 0$.
- $T \ll D_c \ll N$.

Under these assumptions, the single cycle feature detector defined in (13) is non-robust.

Proof: Substituting $Y[n] = X[n - l[m]] + W[n]$ in the definition for $\widetilde{S}_Y^\alpha(f, N)$ (see (14)) and following the same argument as in the proof of Theorem 3, we get

$$\mathbb{E}\left[\widetilde{S}_Y^\alpha(f, N)\right] = \left(\mathbb{E}\left[e^{i2\pi\left(\frac{\alpha}{f_s}\right)l}\right] \cdot \mathbb{E}\left[\widetilde{S}_X^\alpha(f, N)\right] + \mathbb{E}\left[\widetilde{S}_W^\alpha(f, N)\right]\right),$$

where l is the random variable used to denote the random delays $l[m]$. Substituting this in (13) and taking the limit as $N \to \infty$ we have

$$\lim_{N \to \infty} \mathbb{E}[Z_{sc}(N)|\mathcal{H}_1] = \frac{1}{\sqrt{\mathcal{E}}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbb{E}\left[e^{i2\pi\left(\frac{\alpha}{f_s}\right)l}\right] \cdot \mathbb{E}[|h|^2] \cdot |S_X^\alpha(f)|^2 \, df = 0,$$

where the last equality follows because $\mathbb{E}\left[e^{i2\pi\left(\frac{\alpha}{f_s}\right)l}\right] = 0$ when $\left(\frac{\alpha l}{f_s}\right) \sim Unif[0, j]$, for some integer $j > 0$. Hence the means of the test statistic under both hypotheses cannot be separated and so the single cycle detector is non-robust under this model for fading. ∎

The previous result tells us that blind averaging over multiple delay coherence times is not a wise decision and since the particular model of fading used here is so simple (just a variable delay), it will be included in any more natural model for block-fading as long as the delay spread of the channel is larger than $\frac{f_s}{\alpha}$. Notice that α is usually a small multiple of the inter-symbol time and so as long as there is any significant ISI in the channel, this condition will be met in practice. A simple fix to avoid the problem of feature fading across multiple coherence times is to compute the square of the test statistic within each coherence time and average that over multiple coherence times. The modified test statistic is given by

$$\widetilde{Z}_{sc}(N) := \frac{1}{K} \sum_{k=0}^{K-1} |Z_{sc}^{(k)}(D_c)|^2, \tag{19}$$

where $Z_{sc}^{(k)}(D_c)$ denotes the old test statistic (13) computed over the received signal samples in the kth coherence time, i.e., the test statistic computed using the expression given in Equation (13) over the received signal vector $(Y[kD_c + 1], Y[kD_c + 2], \cdots, Y[kD_c + D_c - 1])$.

Theorem 5: Let $Y[n] = \sum_{l=0}^{L-1} h_l X[n - l] + W[n]$ under hypothesis $\mathcal{H}_1$ in (1). The fading coefficients $h[n]$ satisfy a block fading model. Define $H(f) := \sum_{l=0}^{L-1} h_l e^{-j2\pi\left(\frac{f}{f_s}\right)l}$, where f_s is the sampling frequency. Let $\widetilde{X}[n] := h_l * X[n]$. Also, assume that $T \gg \frac{f_s}{\alpha}$, and $D_c \gg T$. Then, the modified single cycle feature detector defined in (19) is non-robust if

$$SNR := \frac{P}{\sigma_n^2} \leq \frac{-\frac{1}{\rho}\frac{\widehat{K}}{D_c} + \sqrt{\frac{1}{\rho^2}\frac{\widehat{K}^2}{D_c^2} + 4\frac{(\rho^2 - \frac{1}{\rho^2})K_1}{D_c}}}{2K_1}, \tag{20}$$

where $P =: \mathbb{E}\left[\frac{1}{D_c}\sum_{n=0}^{D_c-1}|\widetilde{X}[n]|^2\right]$ is the effective power in the signal after fading, and ρ is the noise uncertainty parameter defined in Section III-A. K_1 and $\widehat{K}$ are constants that depend on the shape of the waveform as well as the detailed model of the frequency-selective fading process. They do not depend on the SNR or the coherence time D_C.

Proof: Our goal is to compute $\mathbb{E}\left[\widetilde{Z}_{sc}(N)|\mathcal{H}_i\right]$ for $i = 0, 1$. From (19), $\mathbb{E}\left[\widetilde{Z}_{sc}(N)|\mathcal{H}_i\right] = \mathbb{E}\left[|Z_{sc}^{(k)}(D_c)|^2|\mathcal{H}_i\right]$, $i = 1, 2$, for any $k = 0, \cdots, K - 1$. The last equality follows from the fact that $\{Z_{sc}^{(k)}(D_c)\}$, $k = 0, 1, \cdots, K - 1$ are *iid*. Hence, we only need to consider the first coherence block

$$\mathbb{E}\left[\widetilde{Z}_{sc}(N)|\mathcal{H}_i\right] = \mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_i\right], \quad i = 0, 1$$

$\mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_0\right]$ is computed in the following lemma.

Lemma 2:

$$\mathbb{E}[Z_{sc}(D_c)|\mathcal{H}_0] = 0, \qquad \mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_0\right] = \frac{\sigma^4}{D_c}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}|\widehat{S}_X^\alpha(f)|^2 \, df = \frac{\sigma^4}{D_c}.$$

Proof: The proof is given in Appendix II ∎

Since we are working in a single coherence block, the fading coefficients are time-invariant. So, denote $h_l = h_l[n]$. Under hypothesis $\mathcal{H}_1$, $Y[n] = \widetilde{X}[n] + W[n]$. From (14), we have

$$\widetilde{S}_Y^\alpha(f, D_c) = \widetilde{S}_{\widetilde{X}}^\alpha(f, D_c) + \widetilde{S}_W^\alpha(f, D_c) + \widetilde{S}_{\widetilde{X},W}^\alpha(f, D_c) + \widetilde{S}_{W,\widetilde{X}}^\alpha(f, D_c), \tag{21}$$

where

$$\widetilde{S}_{\widetilde{X},W}^\alpha(f, D_c) := \frac{1}{D_c}\frac{1}{T}\sum_{n=0}^{D_c-1}\left[\widetilde{X}_T\left((2n+1)\frac{T}{2}, f + \frac{\alpha}{2}\right)\cdot W_T\left((2n+1)\frac{T}{2}, f - \frac{\alpha}{2}\right)^*\right]$$

$\widetilde{S}_{W,\widetilde{X}}^\alpha(f, D_c)$ is defined analogously. Substituting (21) in the definition of $Z_{sc}(D_c)$, we get

$$Z_{sc}(D_c)|\mathcal{H}_1 = Z_{\widetilde{X}}(D_c) + Z_W(D_c) + Z_{\widetilde{X},W}(D_c) + Z_{W,\widetilde{X}}(D_c), \tag{22}$$

where $Z_{\widetilde{X}}(D_c), Z_W(D_c), Z_{\widetilde{X},W}(D_c)$, and $Z_{W,\widetilde{X}}(D_c)$ are random variables defined as

$$Z_{\widetilde{X}}(D_c) := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\widetilde{S}_{\widetilde{X}}^\alpha(f, D_c)\cdot\widehat{S}_X^\alpha(f)^* \, df, \quad Z_W(D_c) := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\widetilde{S}_W^\alpha(f, D_c)\cdot\widehat{S}_X^\alpha(f)^* \, df$$

$$Z_{\widetilde{X},W}(D_c) := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\widetilde{S}_{\widetilde{X},W}^\alpha(f, D_c)\cdot\widehat{S}_X^\alpha(f)^* \, df, \quad Z_{W,\widetilde{X}}(D_c) := \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\widetilde{S}_{W,\widetilde{X}}^\alpha(f, D_c)\cdot\widehat{S}_X^\alpha(f)^* \, df. \tag{23}$$

Lemma 3:

$$\mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_1\right] = \mathbb{E}\left[|Z_{\widetilde{X}}(D_c)|^2\right] + \mathbb{E}\left[|Z_W(D_c)|^2\right] + \mathbb{E}\left[|Z_{\widetilde{X},W}(D_c)|^2\right] + \mathbb{E}\left[|Z_{W,\widetilde{X}}(D_c)|^2\right].$$

Proof: The proof is given in Appendix III ∎

We now compute expressions for each of the four terms.

Lemma 4:

$$\mathbb{E}\left[|Z_{\widetilde{X}}(D_c)|^2\right] = P^2 K_1, \qquad \mathbb{E}\left[|Z_W(D_c)|^2\right] = \frac{\sigma^4}{D_c},$$

$$\mathbb{E}\left[|Z_{\widetilde{X},W}(D_c)|^2\right] = \frac{\sigma^2}{D_c} P K_2, \qquad \mathbb{E}\left[|Z_{W,\widetilde{X}}(D_c)|^2\right] = \frac{\sigma^2}{D_c} P K_3,$$

where $K_i, \quad i = 1, 3, 4$ are constants that can be computed numerically given explicit models for both the fading and the signal random processes.

Proof: The proof is given in Appendix IV. ∎

From Lemmas 2-4, we have

$$\mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_0\right] = \frac{\sigma^4}{D_c}, \qquad \mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_1\right] = P^2 K_1 + \frac{\sigma^4}{D_c} + \frac{\sigma^2}{D_c} P K_2 + \frac{\sigma^2}{D_c} P K_3.$$

The detector in (19) will be non-robust if the set of means under both hypotheses are not disjoint, i.e.,

$$\min_{W \in \mathcal{W}_x} \mathbb{E}_W\left[|Z_{sc}(D_c)|^2|\mathcal{H}_1\right] \leq \max_{W \in \mathcal{W}_x} \mathbb{E}_W\left[|Z_{sc}(D_c)|^2|\mathcal{H}_0\right]$$

$$\Rightarrow P^2 K_1 + \frac{\sigma_l^4}{D_c} + \frac{\sigma_l^2}{D_c} P K_2 + \frac{\sigma_l^2}{D_c} P K_3 \leq \frac{\sigma_h^4}{D_c},$$

where $\sigma_l^2 = \frac{1}{\rho}\sigma_n^2$ and $\sigma_h^2 = \rho\sigma_n^2$. The last inequality is quadratic in P, which can be solved to give

$$P \leq \frac{-\frac{\sigma_l^2 \cdot \widehat{K}}{D_c} + \sqrt{\frac{\sigma_l^4 \widehat{K}^2}{D_c^2} + 4\frac{(\sigma_h^4 - \sigma_l^4)K_1}{D_c}}}{2K_1}$$

$$SNR := \frac{P}{\sigma_n^2} \leq \frac{-\frac{\frac{1}{\rho}\widehat{K}}{D_c} + \sqrt{\frac{\frac{1}{\rho^2}\widehat{K}^2}{D_c^2} + 4\frac{(\rho^2 - \frac{1}{\rho^2})K_1}{D_c}}}{2K_1} =: SNR_{wall}^{feature},$$

where $\widehat{K} = K_2 + K_3$. Observe that $SNR_{wall}^{feature} = O\left(\frac{1}{\sqrt{D_c}}\right)$. ∎

VII. DISCUSSION AND CONCLUSIONS

A. Delay vs Amplitude coherence time

Flat fading with random delay is the simplest model under which signal features can be completely killed (see Theorem 4). To understand the nature of the delay coherence time D_c and when it can be different from the regular coherence time, it is useful to consider two cases:

- Assume that the delay spread of the channel is significantly larger than $\frac{\alpha}{f_s}$ in terms of samples and all of the channel taps arise from the sum of many paths. (There is no dominant line-of-sight path) As these path-lengths
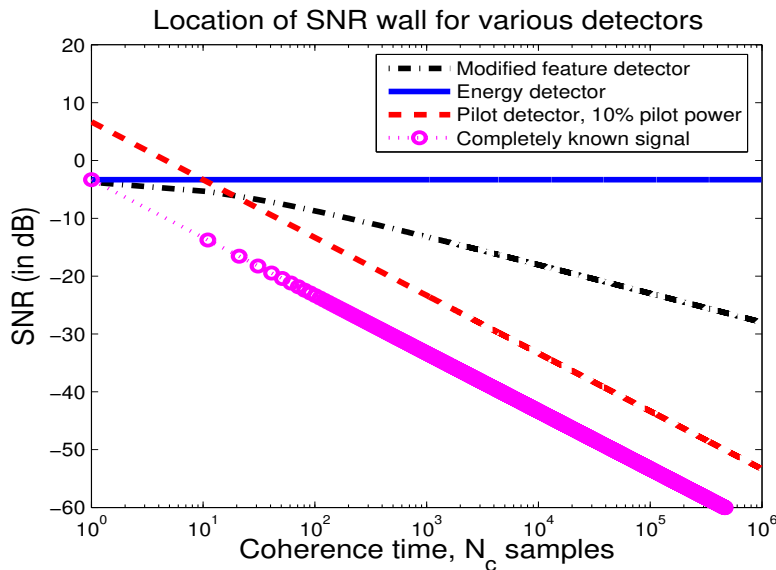
Fig. 4. The location of SNR walls for various detection algorithms is plotted as a function of the channel coherence time N_c. A detailed description of this figure is given in VII-B .

change over time, all of the taps will get different values. In such a case, the delay coherence time D_c and the amplitude coherence time N_c are the same since the relative phase response can change significantly in this time. This is the model that is likely to hold for wideband primary users being detected at significant distances from their transmitters.

- Now consider if there is only one dominant line-of-sight path that is arriving at the secondary user, but is subject to local reflections in the near neighborhood of the secondary. (e.g. A user in an airplane) The delay spread of the channel can be smaller than $\frac{1}{\alpha}$ when viewed in seconds and $\frac{f_s}{\alpha}$ when viewed in samples. In this case, there will just be a single tap in the filter. The amplitude might change rapidly since the local paths can go in and out of phase with each other, but the overall delay is not going to be changing as fast. In such cases, the delay coherence time $D_c > N_c$ and the dominant effect may be the clock skew of the secondary's local oscillator relative to the primary oscillator. In such cases, the factor difference between D_c and N_c is comparable to the factor difference between the signal bandwidth and the carrier frequency.

B. Comparing different SNR walls

In this paper we have shown the existence and computed the location of SNR walls for several detection algorithms. Figure 4 plots the SNR wall locations for energy detection, coherent detection and feature detection as a function of the channel coherence time N_c. In this example, the delay coherence time is assumed to be the same as the amplitude coherence time, i.e., $N_c = D_c$. The data symbols a_n in this example are *iid* Bernoulli-(1/2) random variables modulated onto a rectangular pulse with a 50% duty cycle (See (12)). The strongest feature at $\alpha = \frac{1}{T_0}$ acts like it contains a 0.4 fraction of the total signal energy. There is also a pilot tone with 10% of the

total signal energy, i.e., $\theta = 0.1$ (See (9)). Assume $x = 1$dB uncertainty in the noise level, which corresponds to $\rho = 10^{(1/10)} = 1.259$.

In Figure 4, the solid line corresponds to the energy detection SNR wall, and this is independent of the channel coherence time. The dashed curve corresponds to the pilot detector. The dashed-dotted curved corresponds to the modified single cycle feature detector trying to detect the strongest feature above. The dotted curve represents the best possible case corresponding to the matched-filter when the primary signal is completely known.

The feature detector SNR wall curve from (20) has the constants K_1 and $\widehat{K}$ numerically computed. The expression for the feature detector SNR wall is accurate for large values of N_c. When N_c is small the feature detector SNR wall curve is an approximation. The approximation corresponds to the best case scenario and in reality the actual feature detector wall is worse.

Notice that in Fig 4, the pilot detector SNR wall is strictly better than the feature detector SNR wall for $N_c \geq 60$. For $N_c \leq 60$, the feature detector SNR wall is essentially as bad as the energy detector SNR wall. This is despite the fact that the 50% duty cycle BPSK signal considered in this example has a very strong feature. Such a signaling scheme has a significant loss in primary data rate as compared to an optimal capacity achieving signaling scheme since it is effectively throwing away more than half the degrees of freedom. The robustness of feature detection degrades as the feature we are trying to detect gets weaker and the signaling scheme makes better use of the degrees of freedom. This means that an energy and a coherent detector are jointly more robust than a feature detector in most practical situations.

C. The capacity robustness tradeoff

It is natural to wonder 'Is there any price paid for robustness?' Recall that the goal of the primary transmitter is to maximize its data rate to the primary receiver, while the secondary receiver wants to design a robust algorithm to detect the presence of the primary. The relevant tradeoff is between the performance loss to the primary system and the robustness gains to the secondary system. The primary performance can be quantified using capacity and robustness can be quantified in terms of the SNR wall.

For a target SNR wall our goal is to find the primary signaling scheme that maximizes the data rate achievable to the primary system while still being robustly detectable above the wall. Conversely, given a target primary data rate, our goal is to find the primary signaling scheme that maximizes the robustness of secondary detection. If the primary were only concerned with a given bandwidth and power constraint, then information theory reveals that the optimal signaling is likely to look white and this would give very poor robustness. Thus, this tradeoff is fundamental from a system design point of view and can be called the 'capacity robustness tradeoff'.

Our eventual goal is to derive this optimal tradeoff. However, this is a non-trivial open problem. Instead of aiming for the optimal tradeoff directly, we can find tradeoffs for a given secondary detection algorithm. For

instance, assume that the secondary receiver uses a coherent detector and the primary signal has a deterministic pilot tone. Any power spent on the pilot improves robustness but reduces the power available for data. Similarly, we can derive these tradeoff curves for other commonly used detection algorithms together with primary signaling strategies. This gives us a comprehensive way of comparing individual detection algorithms.
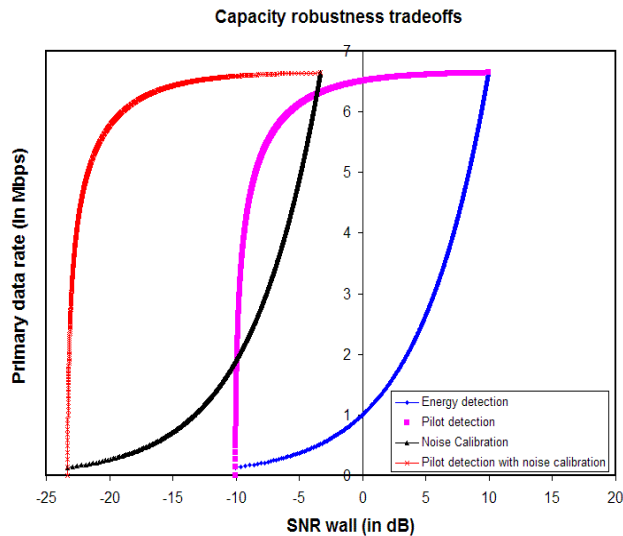

Fig. 5.   The capacity robustness tradeoff curves are plotted for energy and coherent detectors. The primary channel is an AWGN channel with bandwidth 1 MHz, and $SNR = 20$ dB and uncertainty at the secondary is $x = 10$ dB. The rightmost (blue) curve is the tradeoff for energy detection. The tradeoff is obtained by varying the fraction of the total bandwidth used by the primary signal. The second curve from the right (magenta) is the capacity robustness tradeoff for a pilot detector with $N_c = 100$. The tradeoff is obtained by varying the fraction of the primary power used for the pilot tone. The 3rd (black) and the 4th (red) curves from the right are the tradeoffs for energy and coherent detection with noise calibration, which effectively reduces the uncertainty to $x = 1$ dB.

Figure 5 compares the capacity robustness tradeoff curves for energy detection, coherent detection and noise calibration. Qualitatively, one can see that coherent detection is much better than energy detection, i.e., for a given loss in primary data rate, it gives better robustness gains for the secondary system. Also, these curves show that noise calibration can provide significant robustness gains, especially when there is lot of uncertainty due to interference from other secondary users operating in the same band.

D. Concluding remarks

When a system attempts to detect the presence/absence of weak signals in noise, the model uncertainties cannot be ignored since they can be comparable to the desired signals. So while traditionally the performance metric to compare various detection algorithms is their sample complexity, at low SNR the robustness becomes a more significant issue. The robustness can be quantified in terms of an SNR wall giving the threshold below which weak signals cannot be detected reliably no matter how many samples are taken. The SNR wall can be quantified for different detectors and this shows that there is a tradeoff between the amount of knowledge assumed about the primary and the level of robustness. Knowledge of deterministic components of the primary signal such as pilot

tones give rise to SNR walls that scale as $\frac{1}{N_c}$ with amplitude coherence time N_c whereas knowledge of purely structural features of the data signaling strategy give rise to SNR walls that scale down only as $\frac{1}{\sqrt{D_c}}$ with delay coherence time D_c. A simple example where $N_c = D_c$ shows that feature detection by itself can be dominated by simple energy detection and coherent pilot detection at all SNRs.

The example of double-sideband modulation suggests that the robustness gains for feature detection are coming implicitly from run time calibration of the noise model based on the observations themselves. This relies on the primary signal not using all the degrees of freedom itself. Rather than trying to combine traditional cyclostationary feature detection, this suggests that there might be a better way to incorporate the noise calibration gains into purely coherent strategies to get the best possible robustness in practical settings.

APPENDIX I

PROOF OF LEMMA 1

Define

$$
\begin{aligned}
f(\theta) &:= \left[\frac{1}{T} \left(\int_0^T [c(\theta)N_1(t) + s(\theta)N_2(t)]^2 \ dt - \int_0^T [c(\theta)N_2(t) - s(\theta)N_1(t)]^2 \ dt \right) \right] \\
\bar{N}_1 &:= \frac{1}{T} \int_0^T N_1^2(t) \ dt, \quad \bar{N}_2 := \frac{1}{T} \int_0^T N_2^2(t) \ dt, \quad \text{and} \quad \bar{N}_{12} := \frac{1}{T} \int_0^T N_1(t)N_2(t) \ dt.
\end{aligned}
$$

Then, we can write

$$
\begin{aligned}
f(\theta) &= [c(\theta)^2 - s(\theta)^2](\bar{N}_1 - \bar{N}_2) + 4c(\theta)s(\theta)\bar{N}_{12} && (24) \\
&= c(2\theta)(\bar{N}_1 - \bar{N}_2) + 2s(2\theta)\bar{N}_{12}. && (25)
\end{aligned}
$$

Now we can rewrite Equation (10) as

$$
T(\mathbf{Y})|\mathcal{H}_0 = \max_{0 \leq \theta \leq \pi} f(\theta). \tag{26}
$$

Differentiating $f(\theta)$ with respect to θ, the maximizing $\widehat{\theta}$ satisfies, $c^2(2\widehat{\theta}) = \frac{\lambda}{1+\lambda}$ where $\lambda = \left[\frac{\bar{N}_1 - \bar{N}_2}{2\bar{N}_{12}} \right]^2$. This implies

$$
c(2\widehat{\theta}) = \pm\sqrt{\frac{\lambda}{1+\lambda}}, \quad s(2\widehat{\theta}) = \pm\sqrt{1 - \left(\frac{\lambda}{1+\lambda} \right)^2}.
$$

The signs of $c(2\widehat{\theta})$, and $s(2\widehat{\theta})$ are chosen such that $[c(2\widehat{\theta})(\bar{N}_1 - \bar{N}_2)]$ and $[s(2\widehat{\theta})\bar{N}_{12}]$ are non-negative respectively.

Substituting the optimal values in (25) and (26), we get

$$
T(\mathbf{Y})|\mathcal{H}_0 = \frac{(\bar{N}_1 - \bar{N}_2)^2 + 4|\bar{N}_{12}|^3 \cdot \sqrt{2\bar{N}_{12}^2 + (\bar{N}_1 - \bar{N}_2)^2}}{4\bar{N}_{12}^2 + (\bar{N}_1 - \bar{N}_2)^2}. \tag{27}
$$

Note that

$$
\begin{aligned}
\mathbb{E}[\bar{N}_{12}] &= 0, \quad \mathbb{E}(\bar{N}_1 - \bar{N}_2) = 0 \\
\mathbb{E}[\bar{N}_{12}^2] &= O\left(\frac{\sigma_n^4}{T}\right), \quad \mathbb{E}(\bar{N}_1 - \bar{N}_2)^2 = O\left(\frac{\sigma_n^4}{T}\right).
\end{aligned}
\tag{28}
$$

For sufficiently large T, the Central Limit Theorem can approximate $\bar{N}_{12}$ and $(\bar{N}_1 - \bar{N}_2)$ as Gaussian random variables. Hence, taking expectations in (27) and using (28) we get

$$
\mathbb{E}\left[T(\mathbf{Y})|\mathcal{H}_0\right] = O\left(\frac{\sigma_n^4}{T}\right).
$$

This proves the Lemma 1.

APPENDIX II

PROOF OF LEMMA 2

Under hypothesis $\mathcal{H}_0$, $Y[n] = W[n]$, where $W[n] \sim \mathcal{CN}(0, \sigma^2)$. Substituting $Y[n] = W[n]$ in (13) we have

$$
Z_{sc}(D_c)|\mathcal{H}_0 = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_W^\alpha(f, D_c)\widehat{S}_X^\alpha(f)^* \, df.
\tag{29}
$$

Recall from (15)

$$
W_T(n, f) = \sum_{u=n-\frac{T}{2}}^{n+\frac{T}{2}} W[u]e^{-j2\pi\frac{f}{f_s}u}.
$$

Hence, $W_T(n, f) \sim \mathcal{CN}(0, T\sigma^2)$. Also recall that

$$
\widetilde{S}_W^\alpha(f, D_c) = \frac{1}{D_c}\frac{1}{T}\sum_{n=0}^{D_c-1} W_T\left((2n+1)\frac{T}{2}, f+\frac{\alpha}{2}\right) W_T^*\left((2n+1)\frac{T}{2}, f-\frac{\alpha}{2}\right).
$$

Since $W_T\left((2n+1)\frac{T}{2}, f\right)$ are *iid*, the Central Limit Theorem can approximate $\widetilde{S}_W^\alpha(f, D_c)$ as a Gaussian random variable with $\mathbb{E}\left\{\widetilde{S}_W^\alpha(f, D_c)\right\} = 0$, and $\mathbb{E}\left\{|\widetilde{S}_W^\alpha(f, D_c)|^2\right\} = \frac{\sigma^4}{D_c}$. Using this in (29), we have

$$
\mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_0\right] = \frac{\sigma^4}{D_c}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |\widehat{S}_X^\alpha(f)|^2 \, df = \frac{\sigma^4}{D_c},
$$

which proves the lemma.

APPENDIX III

PROOF OF LEMMA 3

From (22), we have

$$
\begin{aligned}
\mathbb{E}\left[|Z_{sc}(D_c)|^2|\mathcal{H}_1\right] &= \mathbb{E}\left([Z_{\widetilde{X}}(D_c) + Z_W(D_c) + Z_{\widetilde{X},W}(D_c) + Z_{W,\widetilde{X}}(D_c)] \cdot [Z_{\widetilde{X}}(D_c) + Z_W(D_c) + Z_{\widetilde{X},W}(D_c) + Z_{W,\widetilde{X}}(D_c)]^*\right) \\
&= \mathbb{E}\left[|Z_{\widetilde{X}}(D_c)|^2\right] + \mathbb{E}\left[|Z_W(D_c)|^2\right] + \mathbb{E}\left[|Z_{\widetilde{X},W}(D_c)|^2\right] + \mathbb{E}\left[|Z_{W,\widetilde{X}}(D_c)|^2\right] \\
&\quad + \mathbb{E}[Z_{\widetilde{X}}(D_c)Z_W^*(D_c)] + \mathbb{E}[Z_{\widetilde{X}}(D_c)Z_{\widetilde{X},W}^*(D_c)] + \mathbb{E}[Z_{\widetilde{X}}(D_c)Z_{W,\widetilde{X}}^*(D_c)] \\
&\quad + \mathbb{E}[Z_W(D_c)Z_{\widetilde{X},W}^*(D_c)] + \mathbb{E}[Z_W(D_c)Z_{W,\widetilde{X}}^*(D_c)] + \mathbb{E}[Z_{\widetilde{X},W}(D_c)Z_{W,\widetilde{X}}^*(D_c)] \\
&\quad + \text{six terms conjugate of the previous six terms.} \quad (30)
\end{aligned}
$$

We show that the first six cross terms are zero, which implies that their conjugates are also zero. First, notice that $\mathbb{E}[Z_{\widetilde{X}}(D_c)Z_W^*(D_c)] = \mathbb{E}[Z_{\widetilde{X}}(D_c)]\mathbb{E}[Z_W^*(D_c)] = 0$, since the signal is independent of noise and $\mathbb{E}[Z_W(D_c)] = \mathbb{E}[Z_{sc}(D_c)|\mathcal{H}_0] = 0$ from Lemma 2. Now,

$$
\begin{aligned}
\mathbb{E}[Z_{\widetilde{X}}(D_c)Z_{\widetilde{X},W}^*(D_c)] &= \mathbb{E}\left[\left(\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_{\widetilde{X}}^\alpha(f,D_c) \cdot \widehat{S}_X^\alpha(f)^* \, df\right)\left(\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_{\widetilde{X},W}^\alpha(f,D_c) \cdot \widehat{S}_X^\alpha(f)^* \, df\right)^*\right] \\
&= \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbb{E}\left[\widetilde{S}_{\widetilde{X}}^\alpha(f_1,D_c) \cdot \left(\frac{1}{D_c}\frac{1}{T}\sum_{n=0}^{D_c-1}\left[\widetilde{X}_T\left((2n+1)\frac{T}{2}, f_2+\frac{\alpha}{2}\right) \cdot W_T\left((2n+1)\frac{T}{2}, f_2-\frac{\alpha}{2}\right)^*\right]\right)^*\right] \\
&\quad \widehat{S}_X^\alpha(f_1)^* \cdot \widehat{S}_X^\alpha(f_2)^* \, df_1 df_2.
\end{aligned}
$$

The random variables $W_T\left((2n+1)\frac{T}{2}, f_2-\frac{\alpha}{2}\right)$ within the expectation are independent of the rest of the terms in the expectation. Also, from Lemma 2, $W_T\left((2n+1)\frac{T}{2}, f_2-\frac{\alpha}{2}\right)$ are zero-mean Gaussian random variables, and are independent in both n and f_2. From these observations, $\mathbb{E}[Z_{\widetilde{X}}(D_c)Z_{\widetilde{X},W}^*(D_c)] = 0$. By similar analysis we can show that each of the cross terms in (30) vanish, which proves the lemma.

APPENDIX IV

PROOF OF LEMMA 4

Observe that $Z_W(D_c) = Z_{sc}(D_c)|\mathcal{H}_0$. Therefore, from Lemma 2, we get $\mathbb{E}[|Z_W(D_c)|^2] = \frac{\sigma^4}{D_c}$. From (23), we have

$$
\begin{aligned}
\mathbb{E}\left[|Z_{\widetilde{X},W}(D_c)|^2\right] &= \mathbb{E}\left[\left(\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_{\widetilde{X},W}^\alpha(f,D_c) \cdot \widehat{S}_X^\alpha(f)^* \, df\right) \cdot \left(\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S}_{\widetilde{X},W}^\alpha(f,D_c) \cdot \widehat{S}_X^\alpha(f)^* \, df\right)^*\right] \\
&= \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbb{E}\left[\widetilde{S}_{\widetilde{X},W}^\alpha(f_1,D_c) \cdot \widetilde{S}_{\widetilde{X},W}^\alpha(f_2,D_c)^*\right] \widehat{S}_X^\alpha(f_1)^* \widehat{S}_X^\alpha(f_2) \, df_1 \, df_2. \quad (31)
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{S}_{\widetilde{X},W}^\alpha(f_1,D_c) \cdot \widetilde{S}_{\widetilde{X},W}^\alpha(f_2,D_c)^*\right] &= \frac{1}{D_c}\mathbb{E}\left[\sum_{n_1=0}^{D_c-1}\sum_{n_2=0}^{D_c-1} \widetilde{X}_T\left((2n_1+1)\frac{T}{2}, f_1+\frac{\alpha}{2}\right) \cdot \widetilde{X}_T^*\left((2n_2+1)\frac{T}{2}, f_2+\frac{\alpha}{2}\right)\right] \\
&\quad \cdot \frac{1}{D_cT^2}\mathbb{E}\left[\sum_{n_1=0}^{D_c-1}\sum_{n_2=0}^{D_c-1} W_T^*\left((2n_1+1)\frac{T}{2}, f_1-\frac{\alpha}{2}\right) \cdot W_T\left((2n_2+1)\frac{T}{2}, f_2-\frac{\alpha}{2}\right)\right]. \quad (32)
\end{aligned}
$$

From Lemma 2, $W_T\left((2n+1)\frac{T}{2}, f_2 - \frac{\alpha}{2}\right)$ are zero-mean Gaussian random variables, and are independent in both n and f_2. Hence, $\frac{1}{D_c T^2}\mathbb{E}\left[\sum_{n_1=0}^{D_c-1}\sum_{n_2=0}^{D_c-1} W_T^*\left((2n_1+1)\frac{T}{2}, f_1 - \frac{\alpha}{2}\right) \cdot W_T\left((2n_2+1)\frac{T}{2}, f_2 - \frac{\alpha}{2}\right)\right] = \frac{\sigma^2}{D_c}$, if $n_1 = n_2$ and $f_1 = f_2$, and is zero otherwise. Substituting this in (32) with $f_1 = f_2 = f$, we get

$$\mathbb{E}\left[\widetilde{S}_{\widetilde{X},W}^{\alpha}(f, D_c) \cdot \widetilde{S}_{\widetilde{X},W}^{\alpha}(f, D_c)^*\right] = \left(\frac{\sigma^2}{D_c}\right) \cdot \frac{1}{D_c}\sum_{n=0}^{D_c-1}\mathbb{E}\left[\left|\widetilde{X}_T\left((2n+1)\frac{T}{2}, f+\frac{\alpha}{2}\right)\right|^2\right]. \tag{33}$$

$\widetilde{X}[n]$ is cyclostationary and hence its statistics are periodic with period $\frac{f_s}{\alpha}$. Since $T \gg \frac{f_s}{\alpha}$, the statistics of a block of T samples of $\widetilde{X}[n]$ can be safely assumed to be independent of each other. Hence $\mathbb{E}\left[\left|\widetilde{X}_T\left((2n+1)\frac{T}{2}, f+\frac{\alpha}{2}\right)\right|^2\right]$ is approximately same for all n.

Observe that $\frac{1}{D_c}\sum_{n=0}^{D_c-1}\mathbb{E}\left[\left|\widetilde{X}_T\left((2n+1)\frac{T}{2}, f+\frac{\alpha}{2}\right)\right|^2\right]$ is similar to the standard periodogram estimate. Since $\widetilde{X}_T(n, f)$ is a linear combination of the time samples $\widetilde{X}[n]$, and D_c is assumed to be large, $\mathbb{E}\left[\frac{|\widetilde{X}_T(n,f)|^2}{P}\right]$ is approximately constant. Therefore,

$$\frac{1}{D_c}\sum_{n=0}^{D_c-1}\mathbb{E}\left[\left|\widetilde{X}_T\left((2n+1)\frac{T}{2}, f+\frac{\alpha}{2}\right)\right|^2\right] = P \cdot K_2,$$

for some constant K_2 which depends on the fading distribution, the signal feature shape, and does not change with D_c and the signal power P. Using this in (33), we get $\mathbb{E}\left[\widetilde{S}_{\widetilde{X},W}^{\alpha}(f, D_c) \cdot \widetilde{S}_{\widetilde{X},W}^{\alpha}(f, D_c)^*\right] = \frac{\sigma^2}{D_c}PK_2$. Substituting this in (31), we have

$$\mathbb{E}\left[|Z_{\widetilde{X},W}(D_c)|^2\right] = \frac{\sigma^2}{D_c}PK_2 \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |\widehat{S}_X^{\alpha}(f)|^2 \, df = \frac{\sigma^2}{D_c}PK_2.$$

By identical reasoning we can show $\mathbb{E}\left[|Z_{\widetilde{X}}(D_c)|^2\right] = P^2 K_1$, and $\mathbb{E}\left[|Z_{W,\widetilde{X}}(D_c)|^2\right] = \frac{\sigma^2}{D_c}PK_3$.

ACKNOWLEDGMENTS

REFERENCES

[1] R. Tandra and A. Sahai, "Snr walls for feature detectors," in *Proc. of 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, April 2007, pp. 559–570.

[2] ——, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. of the WirelessCom 05 Symposium on Signal Processing*, Maui, HI, June 13–16 2005.

[3] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Forty-second Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2004.

[4] R. Tandra, "Fundamental limits on detection in low SNR," Master's thesis, University of California, Berkeley, 2005. [Online]. Available: http://www.eecs.berkeley.edu/~tandra/pub1.htm

[5] FCC, "FCC 04-113," May 2004. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-04-113A1.pdf

[6] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the k user interference channel," *Submitted to the transaction of Information theory*, 2007.

[7] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio," *IEEE Trans. Inform. Theory*, May 2006.

[8] P. Grover and A. Sahai, "On the need for knowledge of the phase in exploring known primary transmissions," in *Proc. of 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, April 2007.

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

[10] S. Shellhammer, "Numerical spectrum sensing requirements," in *IEEE 802.22-06/0088r0*, June 2006.

[11] E. S. Carl R. Stevenson, Carlos Cordeiro and G. Chouinard, "Functional requirements for the 802.22 wran standard," in *IEEE 802.22-05/0007r46*, September 2005.

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

[13] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed.   Springer, March 1998.

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

[15] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," submitted to IEEE MILCOM 2006.

[16] W. Gardner, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, pp. 897–906, Aug. 1988.

[17] ——, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Trans. Commun.*, vol. 40, pp. 149–159, Jan. 1992.

[18] P. J. Huber, "A robust version of the probability ratio test," *The Annals of Mathematical Statistics*, vol. 36, pp. 1753–1758, 1965.

[19] ——, *Robust Statistics*.   Wiley-Interscience, February 1981.

[20] S. Shellhammer and R. Tandra, "Performance of the power detector with noise uncertainty," in *IEEE 802.22-06/0134r0*, July 2006. [Online]. Available: "http://ieee802.org/22/Meeting_documents/2006_July/22-06-0134-00-0000-Performance-of-the-power-detector-with-Noise-Uncertainty.ppt"

[21] R. Tandra and A. Sahai, "Snr walls for signal detectors," University of California, Berkeley, Tech. Rep., September 2007. [Online]. Available: "www.eecs.berkeley.edu/~tandra/pub1.htm"

[22] "Atsc digital television standard." [Online]. Available: "http://www.atsc.org/standards.html"

[23] S. Shellhammer and R. Tandra, "An evaluation of dtv pilot power detection," in *IEEE 802.22-06/0188r0*, September 2006. [Online]. Available: "http://ieee802.org/22/Meeting_documents/2006_Sept/22-06-0188-00-0000-An-Evaluation-of-DTV-Pilot-Power-Detection.ppt"

[24] S. Mathur and S. Shellhammer, "An evaluate of the pn sequence based detection of dtv signals in the draft," in *IEEE 802.22-06/0189r0*, September 2006. [Online]. Available: "http://ieee802.org/22/Meeting_documents/2006_Sept/22-06-0189-00-0000-An-Evaluation-of-the-PN-Sequence-based-detection-of-DTV-signals.ppt"

[25] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*.   New York: Macmillan, 1985.

[26] J. G. Proakis, *Digital Communications*, 4th ed.   McGraw Hill Higher Education, December 2000.

[27] W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*.   Englewood Cliffs, NJ: Prentice-Hall, 1987.

[28] ——, "Spectral correlation of modulated signals, part i-analog modulation," *IEEE Trans. Commun.*, vol. COM-35, pp. 584–594, June 1987.

[29] W. A. Gardner, W. A. Brown, and C. K. Chen, "Spectral correlation of modulated signals, part ii-digital modulation," *IEEE Trans. Commun.*, vol. COM-35, pp. 595–601, June 1987.

[30] W. A. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Processing*, vol. 11, pp. 13–36, July 1986.