SNR Walls for Signal Detection
Rahul Tandra and Anant Sahai

Abstract—This paper considers the detection of the presence/absence of signals in uncertain low SNR environments. Small modeling uncertainties are unavoidable in any practical system and so robustness to them is a fundamental performance metric. The impact of these modeling uncertainties can be quantified by the position of the “SNR wall” below which a detector will fail to be robust, no matter how long it can observe the channel. We propose simple mathematical models for the uncertainty in the noise and fading processes. These are used to show what aspects of the model lead to SNR walls for differing levels of knowledge of the signal to be detected. These results have implications for wireless spectrum regulators. The context is opportunistically sharing spectrum with primary users that must be detected in order to avoid causing harmful interference on a channel. Ideally, a secondary system would be able to detect primaries robustly without having to know much about their signaling strategies. We argue that the tension between primary and secondary users is captured by the technical question of computing the optimal tradeoff between the primary user’s capacity and the secondary user’s sensing robustness as quantified by the SNR wall. This is an open problem, but we compute this tradeoff for some simple detectors.

Index Terms—Capacity robustness tradeoff, cognitive radio, coherent detection, noise calibration, noise uncertainty, radiometer, robust sensing, SNR wall.

I. INTRODUCTION

IN ORDER TO recycle underutilized spectrum, the operation of unlicensed secondary devices within licensed primary bands such as “unused” television (TV) broadcast bands has been proposed [1]. The main constraint for opportunistic secondary devices/cognitive radios is guaranteeing noninterference to the primary system. One possibility is to dynamically align the secondary transmissions such that the resulting interference is mostly orthogonal to the primary’s received signal [2] or by having the secondary user counteract its own interference in the direction of the primary signals [3]. Such design strategies face several practical problems. The two systems must be jointly engineered and may require a huge amount of coordination between the systems involved. Furthermore, even simple phase uncertainties can significantly lower the performance of such collaborative designs [4].

Another possible strategy is for the secondary user to opportunistically use a primary channel only if sensing declares the channel to be vacant. This approach has the advantage of minimal coordination with primary systems and hence is more flexible, but this flexibility comes at the cost of lower performance when many primaries are 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 noninterference with potentially hidden primary receivers and deeply faded channels, the secondary system needs to be able to detect the presence/absence of very weak primary signals [5]–[7]. The IEEE 802.22 is the first international standards process for a cognitive-radio based PHY/MAC/air interface for use in the television bands. Under the currently understood requirements, secondary devices are required to sense TV transmissions as low as $-116$ dBm ($SNR = -22$ dB) [8], [9].

In spectrum sensing, the goal is to meet a given “receiver operating characteristic” (ROC) constraint at very low SNR. Classical detection theory suggests that degradation in the ROC due to reduced SNR can be countered by increasing the sensing time [10], [11]. Hence, the sensitivity is considered to be limited by higher layer design considerations. For instance, the QoS requirements of the application drive the protocol layer design, which in turn dictates the time available for sensing by the physical layer. This traditional perspective implies that a cognitive radio system can always be engineered at the cost of low enough QoS.

In a real-world physical system, parameters are never known to infinite precision. To list just a few, real-world background “noise” is neither perfectly Gaussian, perfectly white, nor perfectly stationary. The channel fading is neither flat nor is it constant over time. Real-world filters are not ideal, A/D converters have finite precision, $I$ and $Q$ signal pathways in a receiver are never perfectly matched and local oscillators are never perfect sine-waves. This paper argues that these model uncertainties impose fundamental limitations on detection performance. The 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 uncertainties.

To illustrate the impact of model uncertainties, consider a simple thought experiment. Imagine a detector that computes a test-statistic and compares it to a threshold to decide if the primary is present/absent. The threshold is set so the target false alarm probability ($P_{FA}$) is met using the nominal model. Now, this detector is deployed in physically distinct scenarios with their own slight variations in local noise and fading characteristics. The actual performance can deviate significantly from the prediction. An example using the radiometer is illustrated in Fig. 1. In fact, below a certain SNR threshold, at least one of the error probabilities can become worse than 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 [12]. To make progress for general classes of signals and detection algorithms, it is important to distill the relevant uncertainties into tractable mathematical models that enable us to focus on those uncertainties that are most critical. The basic sensing problem is posed in Section II with the explicit models for practical wireless system uncertainties in noise and fading being introduced gradually throughout this paper. Section III then explores the robustness of two extreme cases—noncoherent detectors and pilot-detection. The former corresponds to very limited knowledge about the primary signal whereas the latter corresponds to complete knowledge of it. Both of these detectors suffer from SNR walls. In particular, the best possible noncoherent detector is essentially as nonrobust as the radiometer and this nonrobustness arises from the distributional uncertainty in the background noise. In the coherent case, the SNR wall is pushed back, but only to an extent limited by the finite coherence time of the fading process. It is important to note that these SNR walls are not modeling artifacts. They have been experimentally verified in [13] using controlled experiments that were carefully calibrated to limit uncertainties.

Once it is clear that uncertainties limit detector performance, it is natural to attempt to learn the characteristics of the environment at run-time and thereby reduce the uncertainty. We call this approach noise calibration and it is explored in Section IV in the context of pilot detection. The key idea is that the pilot tone does not occupy all degrees of freedom at the receiver and so the other degrees of freedom can be used to learn the noise model. Such strategies can improve performance only so far, with the limit coming from the nonwhiteness of the background noise and the finite coherence-time of the fading process.

The robustness results in this paper have significant implications for policymakers. Suppose that a band of spectrum is opened for cognitive use. The rules governing the band 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, while the secondary should be able to robustly sense the primary. Under these rules, the primary could potentially use all its degrees of freedom by transmitting a “white” signal. As shown in Section III-A2, the secondary would be forced to use noncoherent detection algorithms that are highly nonrobust. The secondary systems face SNR walls and hence are forced to presumptively shut-up since they cannot robustly detect the absence of potential primary signals. This reduces overall spectrum utilization since the gaps between primary users will be unable to be filled by opportunistic secondary users. But, if the rules mandate the primary to transmit a known pilot tone at a certain power, then secondary systems can operate more easily at the cost of potentially lower performance for the primary user. The general tradeoff is posed technically as the capacity-robustness tradeoff in Section V and briefly discussed for the detectors considered here.

II. PROBLEM FORMULATION

Let $X(t)$ denote the band-limited (to total bandwidth $B$) primary signal, let $H$ denote the fading process, and let the additive noise process be $W(t)$. Although this paper focuses on real-valued signals for ease of exposition, the analysis easily extends to complex signals. There are presumed to be different potential primary users in adjacent bands. Thus, our entire universe of consideration is the channel of width $B$ and this is sampled ideally at Nyquist to give a discrete-time model. Sensing is a binary hypothesis testing problem with the following hypotheses:

$$
\mathcal{H}_0 : Y[n] = W[n],
$$
$$
\mathcal{H}_1 : Y[n] = H(X)[n] + W[n].
$$

(1)

Here, $X[n]$ are the samples of the primary signal of interest, $H(\cdot)$ is a possibly random linear time-varying operator representing fading, $W[n]$ are samples of noise and $Y[n]$ are the received signal samples. Throughout this paper, the primary signal is assumed to be independent of both the noise and fading processes. Random processes are also assumed to be stationary and ergodic unless otherwise specified. All three aspects of the system ($W, H, X$) admit statistical models, but it is unrealistic to assume complete knowledge of their parameters to infinite precision. To understand the issue of robustness to uncertainty, we assume knowledge of their distributions within some bounds and are interested in the worst case performance of detection algorithms over the uncertain distributions.\(^1\)

Background noise is an aggregation of various sources like thermal noise, leakage of signals from other bands due to receiver nonlinearity, aliasing from imperfect front end filters, quantization noise, interference due to transmissions from licensed users far away, interference from other opportunistic systems in the vicinity, etc. A stationary white Gaussian assumption is often

\(^1\)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.
tion is only an approximation and so the noise process \(W_t\) is modeled as having any distribution \(W\) from a set of possible distributions \(\mathcal{W}\). This set is called the noise uncertainty set and specific models for it are considered in subsequent sections. The goal is to introduce only as much uncertainty as needed to show that detection becomes nonrobust.

Fading is modeled with the same philosophy. The fading process \(H \in \mathcal{H}_f\) is considered a possibly random linear time-varying filter and models for it will be developed as needed. The general idea is that filter coefficients vary on the time-scale given by the channel coherence time, but 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, the primary signal \(X\) can be modeled by imposing a cap on its power spectral density, instead of actually modeling its specific signal constellation, waveform, etc. The advantage of under-modeling is two-fold. Firstly, it leads to less complex systems. Secondly, intentional under-modeling keeps the model flexible. For example, a simple power spectral density cap on a chunk of spectrum gives the primary user flexibility to choose from a diverse set of signaling strategies and adapt to the circumstances. The set of distributions for \(X\) is denoted \(\mathcal{X}_P\).

Any detection strategy/algorithm can be written as a family \(F\) of functions \(F_N : \mathbb{R}^N \to \{0, 1\}\), where \(F_N\) maps the \(N\) dimensional received vector \(Y = (Y[1], Y[2], \ldots, Y[N])\) 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 noise distribution \(W \in \mathcal{W}\), fading model \(H \in \mathcal{H}_f\) and primary signal \(X \in \mathcal{X}_P\), the error probabilities are

\[
P_{FA}(W, N) = \mathbb{E}_W \left[ 1_{\{F_N = 1\}} | \mathcal{H}_0 \right],
\]

\[
P_{MD}(W, H, X, N) = \mathbb{E}_{W, H, X} \left[ 1_{\{F_N = 0\}} | \mathcal{H}_1 \right].
\]

- The average signal-to-noise ratio is defined as
  \[
  SNR = \frac{P}{\sigma_n^2}, \quad \text{where} \quad P = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} |X[n]|^2.
  \]

Although the actual noise variance might vary over distributions in the set \(\mathcal{W}\), we assume that there is a single nominal noise variance \(\sigma_n^2\) associated with the noise uncertainty set \(\mathcal{W}\). By convention, we consider all the \(X \in \mathcal{X}_P\) to share the same average power \(P\).

- A decision strategy \(F_N\) is said to robustly achieve a given target probability \(P_{FA}\) of false alarm and probability \(P_{MD}\) of missed detection if the algorithm satisfies

\[
\sup_{W \in \mathcal{W}} P_{FA}(W, N) \leq P_{FA},
\]

\[
\sup_{W \in \mathcal{W}, H \in \mathcal{H}_f, X \in \mathcal{X}_P} P_{MD}(W, H, X, N) \leq P_{MD}.
\]

\(^2\)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 turns out not to be true in our problem of spectrum sensing.

- A detection algorithm \(F\) is nonrobust at a fixed SNR if the algorithm cannot robustly achieve any pair \((P_{FA}, P_{MD})\), where \(0 < P_{FA} < 1/2\) and \(0 < P_{MD} < 1/2\), even when \(N\) is made arbitrarily large.

- Suppose there exists a signal-to-noise ratio threshold, \(SNR_t\), such that the detector \(F\) is nonrobust for all \(SNR < SNR_t\). The maximum value of such a threshold is defined as the \(SNR\) wall for the detector

\[
SNR_{wall} = \sup\{SNR_t\},
\]

s.t., the detector is nonrobust for all \(SNR < SNR_t\).

For detectors that are based on computing a test statistic \(T(Y)\), an equivalent condition for testing robustness can be given. A detector is nonrobust iff the sets of means of \(T(Y)\) under both hypotheses overlap \([14], [15]\). Mathematically, for a fixed \(N\), define \(A_N := \{E_W[T(Y)|\mathcal{H}_0] : W \in \mathcal{W}\}\) and \(B_N := \{E_{W,H,X}[T(Y)|\mathcal{H}_1] : W \in \mathcal{W}, H \in \mathcal{H}_f, X \in \mathcal{X}_P\}\). The detector is nonrobust iff \(A_N \cap B_N \neq \emptyset\) for all \(N > 0\).

The goal of this paper is to analyze detection algorithms and prove the existence of \(SNR\) walls.

III. DETECTOR ROBUSTNESS: TWO EXTREME CASES

A. Unknown Signal Structure

Consider a situation where the spectrum sensor 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.

1) Radiometer Robustness: It is useful to review the radiometer under noise level uncertainty \([12]\). The test statistic is given by \(T(Y) = (1/N) \sum_{n=1}^{N} Y[n]^2\). If there is no uncertainty and the noise variance is completely known, the central limit theorem (see \([16]\)) gives the following approximations:

\[
T(Y)|\mathcal{H}_0 \sim \mathcal{N}\left(\frac{\sigma^2}{N}, \frac{1}{N} \sigma^4\right),
\]

\[
T(Y)|\mathcal{H}_1 \sim \mathcal{N}\left(P + \sigma^2, \frac{1}{N} 2(P + \sigma^2)^2\right)
\]

where \(P\) is the average signal power and \(\sigma^2\) is the noise variance. Using these approximations gives

\[
P_{FA} = P_{rob} (T(Y) > \gamma|\mathcal{H}_0)
\]

\[
= Q\left(\frac{\gamma - \sigma^2}{\sqrt{\frac{1}{N} \sigma^4}}\right)
\]

(2)

where \(\gamma\) is the detector threshold and \(Q(\cdot)\) is the standard Gaussian complementary CDF. Similarly

\[
P_{MD} = 1 - Q\left(\frac{\gamma - (P + \sigma^2)}{\sqrt{\frac{1}{N} (P + \sigma^2)^2}}\right).
\]

(3)
Fig. 2. Understanding noise uncertainty for a radiometer. The shaded area in the figure represents the uncertainty in the noise power. It is clear that if the test statistic falls within the shaded region, there is no way to distinguish between the two hypotheses.

Eliminating $\gamma$ from (2) and (3) gives 

$$N = 2 \left[ Q^{-1}(P_{FA}) - Q^{-1}(1 - P_{MD}) \right] (1 + SNR)^2 SNR^{-2}.$$ 

This shows that if the noise power $\sigma^2$ were completely known, then signals could be detected at arbitrarily low SNRs by increasing the sensing time $N$.

Now, consider the case with uncertainty in the noise model. Since the radiometer only sees energy, the distributional uncertainty of noise can be summarized in a single interval $\sigma^2 \in [(1/\rho) \sigma^2_0, \rho \sigma^2_0]$ where $\sigma^2_0$ is the nominal noise power and $\rho > 1$ is a parameter that quantifies the size of the uncertainty.

To see the sample complexity required to achieve a target $P_{FA}$ and $P_D$ robustly, (2) and (3) are modified to get

$$P_{FA} = \max_{\sigma^2 \in [(1/\rho) \sigma^2_0, \rho \sigma^2_0]} Q \left( \frac{\gamma - \sigma^2}{\sqrt{2/\pi} \sigma^2} \right) = Q \left( \frac{\gamma - \rho \sigma^2_0}{\sqrt{2/\pi} \sigma^2_0} \right),$$

$$P_{MD} = 1 - \min_{\sigma^2 \in [(1/\rho) \sigma^2_0, \rho \sigma^2_0]} Q \left( \frac{\gamma - (P + \sigma^2)}{\sqrt{2/\pi} (P + \sigma^2)} \right) = 1 - Q \left( \frac{\gamma - (P + \rho \sigma^2_0)}{\sqrt{2/\pi} (P + \rho \sigma^2_0)} \right).$$

Approximating $1 + SNR \approx 1$ and eliminating $\gamma$ gives

$$N \approx 2 \cdot \frac{[Q^{-1}(P_{FA}) - Q^{-1}(1 - P_{MD})]^2}{[SNR - (\rho / (1 - \rho))]^2}.$$ 

From the above expression it is clear that $N \to \infty$ as $SNR \downarrow (\rho - (1/\rho))$ and this is illustrated in Fig. 3. The figure shows why we call it an “SNR Wall.”

Fig. 2 illustrates that the SNR Wall reflects the fact that the radiometer cannot robustly detect the signal if the signal power is less than the uncertainty in the noise power, i.e., $P \leq (\rho - (1/\rho)) \sigma^2_0$. The presence of the signal is indistinguishable from a slightly larger value for the noise. This prevents us from setting the threshold within the noise uncertainty zone and hence

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

Fig. 4 plots (6) as a function of the noise uncertainty $x = 10 \log_{10} \rho$, as expressed in dB terms. Extensive simulations showing the limits on radiometer performance under noise uncertainty have been reported in [17].

2) Optimal Noncoherent Detection: When both the primary signal and the noise are modeled as white Gaussian processes, then clearly nothing can perform better than the radiometer. However, real-world primary signals are decidedly not Gaussian. We can attempt to exploit knowledge of the
underlying signal constellation, even if the dependence through time is unknown and hence allowed to be white.

Even in this case, it turns out that all detectors suffer from an SNR wall if the uncertainty set is sufficiently rich. To keep the analysis tractable, we propose a bounded moment noise uncertainty model that captures the idea of approximate Gaussianity. See Fig. 5 for a pictorial description of the noise uncertainty set. A white noise distribution \( W \) is in \( W_\mu \) if its i.i.d. marginals \( W_n \) satisfy the following.

- The noise process is symmetric \( EW_a^{2k-1} = 0, \forall k = 1, 2, \ldots \).
- Even moments of the noise must be close to the nominal noise moments in that \( EW_n^{2k} = [(1/\rho^2)EW^2_n, \rho^2 EW^2_n] \), where \( W_n \sim \mathcal{N}(0, \sigma_n^2) \) is a nominal Gaussian noise random variable and \( \rho = 10^{\beta/10} > 1 \).

The parameter \( \beta \) is used to quantify the amount of uncertainty in the noise power, i.e., we allow for \( \pm \beta \) dB of uncertainty in the noise variance. The above 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 \( \sigma_n^2 \); i.e., any Gaussian with variance \( \sigma^2 \in [(1/\rho)\sigma_n^2, \rho\sigma_n^2] \), for some \( \rho > 1 \).
- Background noise also includes weak signals from interferers, quantization noise, and other random disturbances. This means that assuming perfect Gaussianity is unrealistic and the noise uncertainty set should include other distributions that are close to Gaussian. For mathematical tractability, closeness to Gaussian distributions is expressed in terms of the moments. It is clear that similar results hold if the uncertainty set is modeled as a small ball (in the variational distance) around the nominal Gaussian CDF for \( W_n \).

**Theorem 1:** Consider the robust hypothesis testing problem defined in (1) and the above noise uncertainty model \( W_\mu \) with \( x > 0 \). Let \( \rho = 10^{\beta/10} \). Assume that there is no fading and the primary signal model \( \mathcal{F}_p \) includes signals \( X[n] \) satisfying 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 \( X \).
3. \( X \) is zero-mean and all its odd moments are zero, i.e.,
   \[ E[X^{2k-1}] = 0 \text{ for all } k = 1, 2, \ldots \]
4. \( X \) has bounded support so \( |X| \leq \bar{X} \).

Define \( SNR_{\text{peak}} = \bar{X}^2/\sigma_n^2 \). Under these assumptions, there exists a threshold \( SNR_{\text{wall}}^* \) such that robust detection is impossible if \( SNR_{\text{peak}} \leq SNR_{\text{wall}}^* \). Furthermore, there are easy to compute bounds on this SNR wall

\[
\rho - 1 \leq SNR_{\text{wall}}^* \leq \frac{\rho^2 - 1}{\rho}.
\]

**Proof:** See the Appendix.

The upper bound in (7) is the SNR wall for the radiometer from (6). Furthermore, the gap between the upper and lower bounds is at most a ratio of 2 when the uncertainty parameter \( \rho \) is very close to 1, and the ratio approaches 1 if the uncertainty factor \( \rho \) gets large. Thus, Theorem 1 asserts the existence of an absolute SNR wall that is essentially like that of the radiometer. This wall holds for white primary signals with bounded support in cases without fading.

At first glance, the assumption of no fading seems quite severe since fading is natural in wireless environments and furthermore, most models of multipath fading give rise to potentially unbounded signals. However, consider the following possible regulations.

- The primary and secondary users are allowed, but not required, to move. Their peak speed is \( v \).
- The primary user is free to transmit any desired signal, constrained only to stay within its allocated band and to transmit at its assigned total power.

The above seem quite natural and in the spirit of “minimalist regulation” principles. However, this allows both the primary and secondary to be stationary. A lack of motion in turn can lead to infinite coherence times—the channel has some particular fade and this is not going to change during our observation period. There is also no reason to rule out the case of flat fading over the band of interest, i.e., the fading can be represented by a single fading coefficient \( h \). Under the above rules, the hypothetical primary can also transmit a white signal using a BPSK constellation. At this point, the problem is almost covered by the theorem, except for the issue of having a bounded fade.

Recall that the interest in sensing very weak primary signals comes from the need to overcome deeply negative fades [5]–[7]. Suppose that the goal of the secondary system designer is to have probabilities of missed detection and false alarm less than 10%. Set \( \bar{X} \) in the statement of Theorem 1 such that \( P_{\text{false}}(\hat{h}|X) \geq \bar{X} \) = 0.3. If the fading gives rise to constructive or even less destructive interference and increases the received signal strength beyond this level, then the probability of detection can be considered 1. Therefore, to have any hope of meeting the desired specifications, the spectrum sensor must
be able to robustly achieve a probability of missed detection of less than 1/3 for such deeply faded signals. If the resulting $SNR_{\text{peak}}$ is below the $SNR_{\text{wall}}$ of Theorem 1, then it is impossible to do so regardless of how many samples it takes.

The only way around the limitations induced by Theorem 1 is to impose additional structure on the primary signal, to force some diversity in the fading (enforce minimal motion\textsuperscript{3} or switch to a multiuser cooperative setting\textsuperscript{18}), or to somehow reduce the noise uncertainty. Of these three, it is signal structure that seems the most promising from a practical signal processing perspective.

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 the sensor just focuses on that part for detection. This model covers many practical communication schemes that use pilot tones/training sequences for data frame synchronization and timing acquisition. For example, a digital television (ATSC) signal is mandated to have a pilot tone that is \(-11\) dB weaker than the average signal power\textsuperscript{19}.

For simplicity, assume the all potential primary signals $X[n] \in X_P$ are of the form $\sqrt{\theta}X_P[n] + \sqrt{(1-\theta)}X_d[n]$. Here $X_P[n]$ is a known pilot tone with $\theta$ being the fraction of the total power guaranteed to be allocated to the pilot tone. Continue to model $X_P[n]$, $W[n]$ as zero-mean i.i.d. processes as in the previous section.

If there was no fading, then the test statistic for the matched filter is given by $T(Y) = (1/N) \sum_{n=1}^{N} Y[n]X_P[n]$, where $X_P$ is a unit vector in the direction of the known pilot tone. In the case of completely known noise statistics, the matched filter is asymptotically optimal. It achieves a dwell time of $N \approx \left[Q^{-1}(P_D) - Q^{-1}(P_{F,\alpha})\right]^2\theta^{-1}SNR^{-1}$\textsuperscript{20}.

It is easy to see that when there is no fading, the matched filter is robust to uncertainties in the noise distribution alone. Notice that $E[T(Y) | H_0] = 0$ and $E[T(Y) | H_1] = (1/N) \sum_{n=1}^{N} X_P[n]X_P[n] \neq 0$. This shows that the set of means of $T(Y)$ under both hypotheses do not overlap and hence the detector can robustly achieve any $(P_{F,\alpha}, P_{MD})$ pair by choosing a sufficiently large $N$.

The SNR wall in this case is a consequence of time-selectivity of the fading process. Let us consider the simplest possible time-selective fading process: $H$ is modeled as acting on the signal by simple scalar multiplication $h[n]X[n]$ where $h[n]$ is the unknown time-varying flat fade. Even with this simple fading model, it is clear that we cannot reap the gains of coherent signal processing forever. As soon as the channel assumes an independent realization, we can no longer gain from coherent signal processing.

To simplify analysis we assume a block fading model for the fades $h[n]$. That is, $h[n]$ is assumed to be piecewise constant taking an independent realization every coherence time. Furthermore, assume that the length of the coherence time is $N_c$.

It is clear that a good test statistic in this case is

$$T(Y) = \frac{1}{M} \sum_{n=0}^{M-1} \sum_{k=1}^{N_c} Y[n+N_c+k]X_P[n+N_c+k]$$

where $M$ 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 block. Coherent processing gain boosts the signal power by $N_c$ while the noise uncertainty is unchanged. Second, the new boosted signal is detected by passing it through a radiometer. The radiometer aspect remains nonrobust to noise uncertainties, in spite of the boost in the signal strength. The effective $SNR$ of the coherently combined signal is given by

$$SNR_{\text{eff}} = SNR \cdot \theta \cdot N_c.$$  

Hence, the modified matched filter will be nonrobust if

$$SNR \cdot \theta \cdot N_c \leq \left(\frac{\rho^2 - 1}{\rho}\right)$$

$$\Rightarrow SNR_{\text{wall}}^{\text{eff}} = \frac{1}{N_c} \cdot \theta \cdot \left(\frac{\rho^2 - 1}{\rho}\right).$$  

(8)

Coherent processing gains could also be limited due to implementation issues. For example, the clock-instability of both the sensing radio and the primary transmitter imposes a limit on the coherent processing time. The primary signal itself might also contain pilots that are only coherent for a certain length of time. For example, the primary “user” might actually be a collection of wireless nodes doing TDMA among each other and sending packets within their assigned time-slots. Each packet has its own pilot, but the different primary nodes are not all phase-synchronized with each other so no coherent processing gain is available across packets.

Extensive simulation results evaluating the performance of coherent detectors for captured DTV signals can be found in\textsuperscript{21}, \textsuperscript{22}.

IV. Noise Calibration

So far we have shown the existence of $SNR$ walls for detection algorithms that do not attempt to explicitly do anything about the noise uncertainty. The key question is whether it is possible to obtain accurate calibration of uncertain statistical quantities like noise or fading at run-time. The focus on run-time is important because both the fading process and the noise/interference are likely to be at least mildly nonstationary in practice. This means that calibration will get stale (leaving substantial residual uncertainty) if it is done far in advance.

This leads to a tension. Ideally, we would like access to $H_0$ in a parallel universe that is identical to our own except that the primary user is guaranteed to be absent. However, our detector is necessarily confined to our own universe where the presence or absence of the primary user is unknown. Any run-time calibration that aims to reduce the noise uncertainty must use data that might be corrupted by the presence of the primary signal. When the primary user is allowed to be white, then Theorem 1 can be
interpreted as telling us that noise-calibration is impossible. Intuitively, there is no degree of freedom that is uncorrupted by the primary.

Therefore, a prerequisite for getting any noise-calibration gains is that the primary signal must be constrained to discriminate among the available degrees of freedom. In a run-time setting, noise-calibration is only possible if the signal features allow it. Pilot-detection provides a natural case to explore since the pilot signal is confined to a specific frequency, leaving measurements at other frequencies available to do noise calibration.

One can think of noise calibration as a method for adjusting the detector threshold using the additional run-time measurements. For instance, in (4) the radiometer sets its detector threshold corresponding to the worst case noise variance in $[(1/\rho)^2 \sigma^2]$. If the radiometer had access to additional measurements of noise at run-time, then it could adjust its threshold accordingly and hence improve performance.

A. Pilots With White Noise

Consider the basic model of Section III-B, except with the additional assumption that the pilot $X_P[n] = \sqrt{2P\sin(2\pi f_p/B)} n$, where $P$ is the average signal power and $f_p$ is the pilot tone frequency. Rather than modeling the fading process using a block-fading model with a coherence time of $T_C$, it is convenient to think about the fading model in $H_f$ as being arbitrary but bandlimited to a Doppler bandwidth of $B_d$. This means that the coherence time $T_C \approx 1/B_d$. Since $T_C = T_B/B$ where $B$ is the original width of the channel, the Doppler bandwidth can be considered $B_d = B/T_C$. Thus, the coherent processing using the matched filter of the previous section can be reinterpreted as an ideal bandpass filter that selects a bandwidth that is a factor $1/T_C$ as wide as the total primary band and centered on $f_p$. The bandpass filter gives a reduction in noise power, and thus boosts the effective SNR by a factor of $T_C$.

The approach of the previous section throws away all the information in the received signal that lies outside that bandpass filter. Noise calibration is about using that information to improve detection robustness.

Fig. 6 pictorially describes the noise calibration technique in the case of signals with narrowband pilots. The figure plots the power spectral density (PSD) of the primary signal plus noise. The shaded region is the noise PSD and the solid region corresponds to the signal PSD. The sinusoidal pilot tone is shown as a Dirac delta in the frequency domain. Let $f_p$ be the frequency of the pilot tone and $f_m \neq f_p$ be any other frequency within the channel. The noise-calibration detection algorithm can be formally described as follows.

- Pass the received signal through two identical bandpass filters with disjoint passbands centered at $f_p$ and $f_m$, respectively. For mathematical convenience assume that these bandpass filters are ideal “brick-wall” filters and denote their bandwidth by $B_{PF}$. Let the output of these filters be denoted by $Y_{P}[n]$ and $Y_{m}[n]$, respectively.
- Measure the empirical power of both outputs. Let them be denoted by $\hat{P}(f_p, N)$ and $\hat{P}(f_m, N)$, respectively. That is $\hat{P}(f_p, N) = (1/N) \sum_{n=1}^{N} |Y_{P}[n]|^2$ and $\hat{P}(f_m, N) = (1/N) \sum_{n=1}^{N} |Y_{m}[n]|^2$.

- Compare $\hat{P}(f_p, N)$ to a detector threshold $\gamma$ to distinguish between the two hypotheses. The threshold $\gamma$ is set so as to robustly achieve the target probability of false alarm, $P_{FA}$. If we ignore the additional noise measurements $\hat{P}(f_m, N)$, then this detector would get coherent processing gains due to reduction in noise power, but would not get any gains from noise calibration. However, $\hat{P}(f_m, N)$ gives a reliable estimate of the actual noise level and hence can be used to adjust the detector threshold, i.e., set $\gamma = \gamma + \hat{P}(f_m, N)$.
- Adjusting the detector threshold is the same as considering the following new test statistic: $T(Y, \Lambda) := [\hat{P}(f_p, N) - \hat{P}(f_m, N)]$. Compare this test statistic to a static threshold $\gamma$ and declare the signal is present if it exceeds the threshold.

We call this detector the pilot power detector [21]. $\hat{P}(f_p, N)$ estimates the total power in a narrow band around the pilot tone and this includes the pilot power itself, signal power (the solid region in Fig. 6) and the noise power. Since we are concerned with the low SNR regime, the signal power in the band is much lower than the noise power and can be safely neglected. Similarly, $\hat{P}(f_m, N)$ estimates the signal power plus noise power in a narrow band around $f_m$. If we assume that the noise is guaranteed to be white, the noise power in the output of both the bandpass filters is the same. Hence, $E[T(Y, \Lambda) | H_0] = 0$ and $E[T(Y, \Lambda) | H_1] = \theta P \neq 0$. This clearly implies that this detector is robust to distributional uncertainties in the margins of a white noise process.

B. Colored Noise Uncertainty Model

The robustness of the pilot power detector described in the previous section critically depends on the absolute whiteness of noise. It is unreasonable to assume that noise is completely white. Since “noise” includes thermal noise and interference from other transmitters sharing the same band of spectrum, leakage from adjacent bands, etc., it has some a priori unknown
color to its spectrum. Real-world noise also contains seemingly arbitrarily placed spurs and other terms that can change quite rapidly across the channel [23].

Our noise model is motivated by the idea that the noise uncertainty is dominated by the unknown interference that is coming from an aggregation of many sources. The uncertainty in the level of the overall noise corresponds to the uncertainty in the activity level of the interference. This can vary wildly over possible sites because it is a function of how far away the sources of interference are. To understand the color, suppose that each interferer is roughly white to begin with, but hits our spectrum sensor after passing through its own frequency-selective channel representing multipath fading. The channels between our spectrum sensor and the different interferers are independent. Thus the autocorrelation function of the sum has a large term at a lag of zero, but the nonzero lag terms coming from different interferers tend to cancel out by the independence of the fades. However, there are a finite number of dominant interferers and so the autocorrelation function of the aggregate is not a perfect delta-function. This type of scenario naturally leads to a model of unknown, but slight, color for the noise.

Fig. 7(a) illustrates the set of possible noise power spectral densities under the earlier model \( \mathcal{W}_c \) of Section III-A2. The PSD is flat, but its variance can vary within the bounds as shown in the figure. Fig. 7(b) plots the set of possible power spectral densities under the new colored noise uncertainty model \( \mathcal{W}_{p,\lambda} \). The two solid (red) curves correspond to two different nominal levels around which the noise has some ripples.

Formally, a noise distribution \( W \in \mathcal{W}_{p,\lambda} \) if the marginals satisfy the conditions from Section III-A2, but the process is not necessarily white. We do assume that the noise process is wide sense stationary, and its power spectral density \( S_W(f) \) satisfies

\[
(1/\lambda)N_a \leq S_W(f) \leq \lambda N_a \text{ for some } \rho > \lambda > 1. \text{ Here } N_a \text{ satisfies } (\lambda/\rho)N_0 \leq N_a \leq (\rho/\lambda)N_0 \text{ and represents the unknown overall PSD level actually experienced by the cognitive radio in this particular site at the time of sensing. } N_0 \text{ is the nominal PSD level at design time and this is related to the discrete-time } \sigma_n^2 \text{ using } N_0B = \sigma_n^2.
\]

C. Pilots With Colored Noise

Consider the detector in Section IV-A under our new colored noise uncertainty model. Recall that the test statistic for the detector is \( T(Y, N) = \hat{P}(f_p, N) - \hat{P}(f_m, N) \), i.e., the test statistic measures the difference in powers at the pilot frequency \( f_p \) and the noise calibration frequency \( f_m \). Since the noise can be colored, it is clear that the difference between the noise powers in the two passbands (centered at \( f_p \) and \( f_m \)) need not be zero. The gains from noise calibration are limited and the detector faces SNR wall limitations.

**Theorem 2:** Consider the detection problem in (1). Assume the nominal \( X[n] = \sqrt{\theta}X_p[n] + \sqrt{1-\theta}X_d[n] \), the fading uncertainty set \( \hat{\mathcal{H}}_f \) contains arbitrary bandlimited processes \( H \) with Doppler bandwidth \( B/N_c \) and unit gain, and \( W[n] \) are noise samples whose distribution lies in the colored noise uncertainty set \( \mathcal{W}_{p,\lambda} \). Assume \( X_p[n] = \sqrt{2P}\sin 2\pi(f_p/B)n \).

In this case the pilot power detector defined in Section IV-A is nonrobust if

\[
SNR \leq \frac{\rho}{\lambda\theta} \cdot \left( \frac{\lambda^2-1}{\lambda} \right) \cdot \left( \frac{2B_{BPF}}{B} \right)
\]  

(9)

where \( B \) is the bandwidth of the primary signal and \( B_{BPF} \geq B/N_c \) is the bandwidth of the bandpass filter in the detection algorithm. If the filter’s width \( B_{BPF} = B/N_c \) is set to be as narrow as possible, then (9) simplifies to

\[
SNR \leq \frac{\rho}{\lambda\theta} \cdot \left( \frac{\lambda^2-1}{\lambda} \right) \cdot \left( \frac{2}{N_c} \right).
\]  

(10)

**Proof:** The test statistic in Section IV-A is \( T(Y, N) = \hat{P}(f_p, N) - \hat{P}(f_m, N) \). This detector is nonrobust if

\[
\inf_{W \in \mathcal{W}_{p,\lambda}, H \in \hat{\mathcal{H}}_f} E_W[H|\mathcal{H}_i] 
\leq \sup_{W \in \mathcal{W}_{p,\lambda}} E_W[H|\mathcal{H}_0]
\leq \theta P + \sup_{W \in \mathcal{W}_{p,\lambda}} E_W[T(W, N)]
\leq \theta P + \left( \frac{1}{\lambda}N_aB_{BPF} - \lambda N_aB_{BPF} \right)
\leq \left( \frac{\lambda N_aB_{BPF}}{(\lambda-1)} \right) \cdot \left( \frac{2B_{BPF}}{B} \right) \cdot \left( \frac{N_a}{N_0} \right)
\leq \frac{\rho}{\lambda\theta} \cdot \left( \frac{\lambda^2-1}{\lambda} \right) \cdot \left( \frac{2B_{BPF}}{B} \right).
\]
• In step (a) above, \( W \) is an \( N \) dimensional vector of noise samples drawn from a distribution \( W \sim \mathcal{W}_{\rho \lambda} \). The result follows from the fact that the test statistic under hypothesis \( \mathcal{H}_1 \) is the sum of the faded pilot power and test statistic on the noise alone. This is true because the bandpass filter is wide enough to capture all the pilot power even after it is spread in frequency by the channel’s Doppler spread. So \( \mathbb{E}[|H(X_p)|^2] = \mathbb{E}[|H|^2]\mathbb{E}[|X_p|^2] = \theta P \) since the unknown bandlimited fading process \( H \) is restricted to have unit gain on average.

In addition, we are neglecting the power of the data part of the signal \( (X_d[n]) \) in the passband. This assumption is reasonable because we are interested in the regime where \( P \ll \sigma_n^2 \).

• Step (b) is obtained by using the fact that the minimum difference in noise powers in the two bands is \((1/\lambda)N_0B_{BPF} - \lambda N_0B_{BPF}\). Similarly, the maximum possible difference is \((N_0B_{BPF} - (1/\lambda)N_0B_{BPF}\).

• Step (c) follows from the fact that \( N_0B = \sigma_n^2 \). The final factor of \( \rho/\lambda \) is just present to account for the fact that \( N_0B \approx \sigma_n^2 \). This is an artifact of the fact that we are expressing the wall in terms of the nominal SNR and the uncertainty is modeled multiplicatively.

From Theorem 2, the only way to have the pilot power detector’s SNR wall go to zero is to let \((B_{BPF}/B) \to 0\). This in turn requires the coherence-time \( N_c \to \infty \). This is because the pilot tone is a Dirac delta in frequency domain and is smeared out due to the fading process. The pilot power is spread over the bandwidth of the fading process in some unknown way. Under this uncertainty model \( \mathcal{H}_f \), tightening the bandpass filter any further could result in the loss of all the pilot power and hence missed detection.

It is easy to see that because the time allowed for sensing is unbounded, the above noise calibration strategy does not really depend on the width of the bandpass filter used for calibration. Even if a detector used all the information outside of the tight window of width \( B_d \) containing the pilot signal, the detector cannot completely calibrate the noise process within that tight window due to the model \( \mathcal{W}_{\rho \lambda} \) of unknown color. Within the tight window of width \( B_d \), all detectors essentially face the same problem as Theorem 1— an arbitrary (and hence possibly white) signal in the presence of noise with a residual unknown level coming from the color ripple.

D. Primary Signals With Perfect Guard-Bands

The case of pilot tones is just one way that the primary signal can leave room for noise calibration. Suppose instead that the primary signal leaves empty a contiguous \((1 - \beta)B\) of its total licensed bandwidth \( B \) for a guard-band, but is free to use the remaining \( \beta B \) bandwidth as it sees fit. It is immediately obvious that the pilot power detector can be generalized to this case.

• Pass the received signal through two different ideal bandpass filters. Let the first encompass the tightest band guaranteed to contain the entire primary signal. By the assumptions above, this is of width \( B_{BPF} = (\beta + (1/N_c))B \) since the \( \beta B \) must be expanded by \( B_d \) to account for the fading Doppler. Let the second bandpass filter have width \( 0 < B_c \leq (1 - \beta - (1/N_c))B \) and be placed entirely in the guard-band. Let the output of these filters be denoted by \( Y_p[n] \) and \( Y_m[n] \), respectively.

• Measure the empirical power \( \hat{P}(f_p, N) \) and \( \hat{P}(f_m, N) \) of both filter outputs. Since there is no limit on the sample complexity, assume that we listen for long enough for both to have converged to their limits.

• The test statistic is given by \( T(Y, N) := [\hat{P}(f_p, N) - (B_{BPF}/B_c)\hat{P}(f_m, N)]. \) Compare this test statistic to a threshold and declare the signal is present if it exceeds the threshold. The normalization term \( B_{BPF}/B_c \) would make the test statistic perfectly zero if there were no signal and the noise were perfectly white.

It is easy to see that Theorem 2 generalizes to this scenario and detector. The SNR wall from (9) holds with \( \theta = 1 \) and \( B_{BPF} \) as above. This shows that the noise calibration gains add to the bandwidth reduction gains in a straightforward way by reducing the noise uncertainty to a factor that depends on the coloring uncertainty \( \lambda \).

V. CAPACITY-ROBUSTNESS TRADEOFF

It is natural to wonder "Who must pay the price for robust detection?" The goal of the primary transmitter is to maximize its data rate to its own receiver, while the secondary user wants to robustly 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 this paper’s results indicate that robustness can be quantified in terms of the SNR wall.

For a target SNR wall, the ultimate goal is to find a primary signaling scheme that maximizes the data rate achievable for the primary system while still being robustly detectable above the wall. Conversely, given a target primary data rate, the 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 should look white [24] and Theorem 1 indicates that this gives very poor robustness for the secondary. Thus, it appears that this tradeoff is fundamental from a system design point of view and can be called the "capacity-robustness tradeoff."

It is a nontrivial open problem to determine the optimal tradeoff. Instead, here the tradeoff is explored in the context of a given sensing algorithm and class of primary signals.

A. Capacity-Robustness Tradeoff: Radiometer

Let the channel between the primary transmitter and primary receiver be a classic bandlimited AWGN channel. The capacity of this channel is \( C = B \log(1 + SNR_P) \) assuming complex signaling. Here, \( B \) is the channel bandwidth and the signal-to-noise ratio \( SNR_P = P_P/N_0B \) where \( P_P \) is the received signal power at the primary receiver. Assume that the spectrum sensor uses a radiometer to detect the primary signal. As seen in Section III-A1, the relevant \( SNR_{P,wall} = (\rho^2 - 1)/\rho \). The nominal \( SNR_P \) at the sensor can be written as \( SNR_P = P_S/N_0B, \) where \( P_S \) is the received primary power at the spectrum sensor.
It is clear that the common parameter that affects both the primary data rate and the sensing \( SNR \) is the bandwidth \( B \). The primary user can tradeoff its rate in exchange for robustness at the spectrum sensor by reducing the fraction of degrees of freedom used. Let \( B_{\text{actual}} = \beta B \), where \( \beta \in [0,1] \). We also assume that the fading at the spectrum sensor is bandlimited and has a Doppler of \( B_d = B/N_c \). In this case, the total bandwidth occupied by the primary signal at the spectrum sensor is \( \beta B + (B/N_c) \). The effective SNR wall is given by \( (\beta + (1/N_c))(\rho^2 - 1)/\rho) \). Hence the capacity-robustness tradeoff for the simple radiometer is given by the parametric curve
\[
\left( \beta + \frac{1}{N_c} \right) \left( \frac{\rho^2 - 1}{\rho} \right), \beta B\log \left( 1 + \frac{P_P}{\beta N_0 B} \right) \tag{11}
\]
where \( \beta \in [0,1] \). The first coordinate in (11) is the robustness of secondary detection and the second coordinate is an upper bound on the data rate achievable for the primary system. The factor \( \beta + (1/N_c) \) in the robustness is the processing gain obtained because the primary does not use all the degrees of freedom.

One can also obtain noise calibration gains by taking measurements in the frequencies over which the primary does not spread its signal. Using the model in Section IV-D, the uncertainty in the noise is reduced from \( \rho \) to some smaller number \( \lambda \). Therefore, the tradeoff with noise calibration is given by
\[
\left( \beta + \frac{1}{N_c} \right) \left( \frac{2\rho}{\lambda} \right) \left( \frac{\lambda^2 - 1}{\lambda} \right), \beta B\log \left( 1 + \frac{P_P}{\beta N_0 B} \right) \tag{12}
\]

\[B. \text{ Capacity-Robustness Tradeoff: Coherent Detector}\]

Assume that the spectrum sensor uses a coherent detector and the primary signal has a deterministic pilot tone as in Section III-B. Any power spent on the pilot improves robustness but reduces the power available for communicating messages. Let \( \theta \) be the fraction of total power spent in the pilot tone. The primary data rate is bounded by \( R_P = B \log(1 + ((1 - \theta)P_P/N_0 B)) \). From (8), the SNR wall is given by \( SNR_{\text{wall}} = (1/(N_c \cdot \theta))(\rho^2 - 1)/\rho) \), where \( \theta \in [0,1] \) and \( N_c \) is the channel coherence time. Thus, the capacity-robustness tradeoff for the matched filter is given by the parametric curve
\[
\left[ \frac{1}{N_c \cdot \theta} \left( \frac{\rho^2 - 1}{\rho} \right), \beta B\log \left( 1 + \frac{(1 - \theta)P_P}{N_0 B} \right) \right] \tag{13}
\]
where \( \theta \in [0,1] \). The gains in robustness given in (13) are coherent processing gains. As discussed in Section IV-C, noise calibration can also be used to reduce the uncertainty down to the \( \lambda \) level. Hence, by (10) the capacity-robustness tradeoff with noise calibration is given by
\[
\left[ \frac{2\rho}{\lambda \cdot N_c \cdot \theta} \left( \frac{\lambda^2 - 1}{\lambda} \right), \beta B\log \left( 1 + \frac{(1 - \theta)P_P}{N_0 B} \right) \right]. \tag{14}
\]

\[C. \text{ Numerical Examples}\]

Fig. 8 compares the capacity-robustness tradeoff curves for the radiometer and coherent detection. These curves were plotted using (11), (12), (13) and (14). These plots correspond to a bandwidth \( B \) of 1 MHz, the signal-to-noise ratio of the primary channel \( SNR_P = 20 \) dB, the channel coherence time \( N_c = 1000 \), the noise uncertainty parameters \( \rho = 10 \) (to account for potentially significant interference) and \( \lambda = 10^{11} \). Qualitatively, one can see that coherent detection is much better than the radiometer, i.e., for a given loss in primary data rate, it gives better robustness gains for the secondary system. Essentially, the primary system is not power limited. So, sacrificing some power to the pilot tone is not a significant cost, but it gives significant robustness gains. However, shrinking the bandwidth and using fewer degrees of freedom to improve radiometer robustness is painful since it is degrees of freedom that are the critical resource for high SNR primary users.

Fig. 9 compares the capacity-robustness tradeoffs for the radiometer and coherent detection when the primary \( SNR_P \) is low. These tradeoffs were plotted for \( SNR_P = -10 \) dB. The
Finally, in order to make a fair comparison, the nominal SNRs in all these plots correspond to the worst possible noise power within the noise uncertainty set as opposed to the middle noise power. This way, it is clear that after noise calibration, it makes no difference what the original \( \rho \) was. But noise-calibration makes less of a difference when \( \rho \) and \( \lambda \) are comparable to start with.

VI. CONCLUSIONS

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 used to compare various detection algorithms is the sample complexity, at low SNR, robustness becomes a more significant issue. The robustness can be quantified in terms of the SNR wall giving the threshold below which weak signals cannot be detected reliably no matter how many samples are taken.

The SNR wall is quantified for different detectors and the results show that there is a tradeoff between the amount of knowledge assumed about the primary and the level of robustness. If the primary signal does not use all the degrees of freedom equally, detection robustness can be improved by explicitly trying to reduce model uncertainty at run-time through noise calibration. While this mitigates the SNR wall, it does not eliminate it.

The kinds of uncertainties that cause the SNR wall vary with the degree of knowledge of the primary signal and the resulting detection strategies. If nothing is known, then the uncertainty in the marginal noise distribution alone is enough to cause an SNR wall. If any degree of freedom is available for attempts at noise-calibration, then the relevant uncertainty shifts to the uncertain color of the noise rather than the uncertain marginals alone. If a pilot tone is known, then it is the uncertain time-selectivity of the fading process that combines with the noise uncertainty to cause the SNR wall.

Frequency is not the only way to partition degrees of freedom. For instance, we could consider primary signals whose symbol pulses have only a 50% duty cycle. This is a simple example of signals using only half the time degrees of freedom. Such pulse amplitude modulated signals come under the general category of cyclostationary signals [25]–[27]. Feature detectors are a general class of detectors that have been proposed to distinguish cyclostationary signals from noise [28], [29]. Fundamental results on robustness of feature detectors are shown in [30], [31] by extending the results here. It turns out that in such cases, the SNR walls are due to the uncertain frequency selectivity of the fading.

The problem of low SNR signal detection arises naturally when considering the opportunistic reuse of wireless spectrum that might be used by primary users. In many practical settings, there is a significant amount of geographic area (“white-space”) that could potentially be occupied by secondary systems without causing harmful interference, if they could somehow verify that they were indeed within this white-space. Attempting to detect very weak signals is one approach to doing this when the secondary system faces fading [5]–[7]. To balance the goals of the primary and secondary users, the capacity-robustness tradeoffs

![Capacity robustness tradeoffs](image)

Fig. 10. Capacity-robustness tradeoff curves for the radiometer and coherent detectors where the primary is operating at high spectral efficiency and moderate noise uncertainty (\( \rho = 10^{0.2}, \lambda = 10^{0.4} \)).

![Capacity robustness tradeoffs](image)

Fig. 11. Capacity-robustness tradeoff curves for the radiometer and coherent detectors where the primary is operating at low spectral efficiency and moderate noise uncertainty (\( \rho = 10^{0.2}, \lambda = 10^{0.4} \)).

remaining parameters are the same as in Fig. 8. Since the primary is power limited, the loss of power to the pilot has a non-trivial rate penalty. Meanwhile, giving up degrees of freedom incurs less of a penalty to help the radiometer. The robustness curves for the two strategies come closer to each other, but deterministic pilot tones and coherent detection still win out. Also, both Figs. 8 and 9 are plotted for the case when the ratio of the noise uncertainty parameters \( \rho = 10 \) and \( \lambda = 10^{0.4} \) is large. In this case it is clear that run-time noise calibration gives significant gains to both the radiometer and coherent detector.

Figs. 10 and 11 plot the capacity robustness tradeoff curves for the radiometer and coherent detection when the ratio of the noise uncertainty parameters \( \rho = 10^{0.2} \) and \( \lambda = 10^{0.4} \) is small. As before the two figures correspond to high and low spectral efficiency for the primary system.
given here suggest that primary users should be mandated to transmit known pilot tones. Otherwise, the SNR walls indicate that it is impossible for spectrum sensors to robustly detect the absence of even moderately weak unstructured signals.

Requiring secondary users to detect very weak primary signals has problems even without SNR wall limitations. If infinitely robust spectrum sensors listen for very weak primary signals, they will start picking up signals from distant primary transmitters. Even though the secondary users are far away from primary systems and can safely use the spectrum, they will be forced to shut-up because they cannot distinguish between this situation and being close but deeply faded [6], [7]. In effect, the fading margin for sensing at the secondary user enlarges the guard-space around each primary transmitter. If the fading margin is too large, then all of the white-space will be lost to these guard-spaces. Cooperative approaches to sensing will likely be required to avoid giving up too much area, and quantifying the impact of modeling uncertainty in those contexts is an open research question [18], [32].

APPENDIX

PROOF OF THEOREM 1

The goal is to show the existence of an absolute SNR wall, below which every detector is nonrobust. This requires all possible test statistics to have overlapping means under the two hypotheses. The easiest way for this to happen is if the set of joint distributions of the received signal vector \( Y \) under both hypotheses overlap. As the noise and signal are both modeled as i.i.d., it is sufficient to show that the set of marginal distributions of the scalar received signal samples \( Y[n] \) overlap under both hypotheses. Mathematically, we just need to show the existence of two noise distributions \( W_1, W_2 \in W_x \), such that \( W_1 + X \) has the same distribution as \( W_2 \).

Let \( W_1 \sim \mathcal{N}(0, \sigma_n^2) \), \( W_1 \sim (1/\sqrt{\rho})W_1 \). Let \( W_2 := X + W_1 \). If \( W_1 \in W_x \), and \( W_2 \in W_x \), then it is clear that robust detection is impossible. We now prove \( W_2 \notin W_x \) by showing that \( E[X^{2k-1}] = 0 \) and \( \rho^{k}E[W_2^{k}] \leq E[W_2^{k}] \leq \rho^{k}E[W_1^{k}] \forall k \).

Since \( E[X^{2k-1}] = 0 \) and \( W_2 := X + W_1 \), we have

\[
E[X^{2k-1}] = 0 \quad \text{and} \quad E[W_2^{k}] \geq (1/\rho^{k})E[W_1^{k}].
\]

Therefore, \( W_2 \notin W_x \).

\[
\Rightarrow E[W_2^{k}] \leq \rho^{k}E[W_1^{k}]
\]

\[
E[X^{2k-1}] = 0 \quad \text{and} \quad E[W_2^{k}] \leq E[X^{2k}] \leq \rho^{k}
\]

\[
\Rightarrow E[W_2^{k}] \leq \rho^{k}
\]

\[
\Rightarrow \sum_{i=0}^{k} \frac{2k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) \frac{\sigma_n^{2i}}{\sigma_n^{2i}} \leq \rho^{k}
\]

\[
\Rightarrow \sum_{i=0}^{k} \frac{2k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) \frac{\sigma_n^{2i}}{\sigma_n^{2i}} \leq \rho^{k}, \quad \forall k = 1, 2, \ldots.
\]

So, we just need to show that there exists an absolute threshold \( SNR^* \) such that the last inequality in (15) is true if \( SNR_{\text{peak}} \leq SNR^* \).

We now construct the threshold \( SNR^* \). For \( k \geq 1 \), define \( f(k, SNR) \) to be the following uni-variate degree \( k \) polynomial

\[
f(k, SNR) = \left[ \frac{2k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) \sigma_n^{2i} \right].
\]

Note that all the coefficients of \( f(k, SNR) \) are nonnegative 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 \( SNR_{\text{peak}}^{(2k)} \). Because \( f(k, 0) = (1/\rho^{k}) < \rho^{k} \) and \( \rho > 1 \), the root \( SNR_{\text{peak}}^{(2k)} > 0 \) for all \( k \). Define \( SNR^* \) to be

\[
SNR^* := \inf_{k=1, 2, \ldots, \infty} SNR_{\text{peak}}^{(2k)}. \tag{17}
\]

By definition of \( SNR_{\text{peak}} \), we have \( \left( E[X^{2i}] / \sigma_n^{2i} \right) \leq (SNR_{\text{peak}})^{2i} \). Now, if \( SNR_{\text{peak}} \leq SNR^* \) by the monotonicity of \( f(k, SNR) \), we have

\[
\sum_{i=0}^{k} \frac{2k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) \frac{\sigma_n^{2i}}{\sigma_n^{2i}} \leq \rho^{k}
\]

\[
\Rightarrow \sum_{i=0}^{k} \frac{2k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) SNR_{\text{peak}}^{(2i)} \leq f(k, SNR_{\text{peak}}) \leq \rho^{k}
\]

which is the last inequality in (15).

In terms of the upper-bound in (7), from (17) it is clear that \( SNR^* \leq SNR_{\text{wall}}^{(2k)} = (\rho^{k} - 1)/\rho \). However, the entire result is vacuous if the infimum in (17) is zero. Getting a nontrivial lower bound is more interesting.

Recall that \( SNR_{\text{Wall}}^{(2k)} \) is the root of the polynomial \( f(k, SNR) - \rho^{k} \). The coefficient of \( SNR^k \) in \( f(k, SNR) - \rho^{k} \) is \( (2k)(1/\rho^{k})((1/\rho^{k} - 1)/1\cdot3\cdot(2k-1)) \). Replace \( (1/\rho^{k}-1) \) by 1 to get a new polynomial, \( \tilde{f}(k, SNR) - \rho^{k} \). Since \( \rho > 1 \), we have \( (1/\rho^{k}-1) \leq 1 \) for all \( i = 1, 2, \ldots, k \); and hence the unique positive root of \( \tilde{f}(k, SNR) - \rho^{k} = 0 \) must be smaller than \( SNR_{\text{wall}}^{(2k)} \). Call this root \( \tilde{SNR}_{\text{wall}}^{(2k)} \). Clearly,

\[
SNR^* \geq \inf_{k=1, 2, \ldots, \infty} \tilde{SNR}_{\text{wall}}^{(2k)} \tag{18}
\]

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

\[
\tilde{f}(k, SNR) := \frac{k}{2i} \left( \frac{1}{\rho^{k}} \right) \left( \frac{1 \cdot 3 \cdot (2k - 2i - 1)}{1 \cdot 3 \cdot (2k - 1)} \right) SNR^{i} \]

\[
= \rho^{k}. \tag{19}
\]
Fix $k \geq 1$ and $SNR \geq 0$. From the definition of $\tilde{f}(k, SNR)$ in (19), we get

$$
\tilde{f}(k, \text{SNR}) = \left[ \sum_{i=0}^{k} \frac{(2k-2i+1) \cdot (3 \cdot 5 \cdots (2i-1))}{(2i-1) \cdot (2i-3) \cdots 1} \right] \text{SNR}^i,
$$

and issue editor A. Sayeed were immensely helpful in shaping the presentation and scope of this paper.

**REFERENCES**


This work is also indebted to ongoing interactions with M. Mishra, D. Cabric, and N. Hoven. The anonymous reviewers

**ACKNOWLEDGMENT**


Rahul Tandra received the B.Tech. degree from the Indian Institute of Technology, Bombay, and the M.S. degree from the University of California, Berkeley, in 2005, both in electrical engineering. He is currently pursuing the Ph.D. degree in electrical engineering at the University of California, Berkeley.

He is a member of the Wireless Foundations Research Center. In 2006, he worked as a Summer Intern in the Corporate R&D division at Qualcomm Inc., developing spectrum sensing algorithms for the IEEE 802.22 standard. His research interests are in wireless communications and signal processing. He is particularly interested in fundamental research questions in dynamic spectrum sharing.

Anant Sahai received the B.S. degree from the University of California, Berkeley, in 1994, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1996 and 2001, respectively.

He joined the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley in 2002 as an Assistant Professor. He is also a member of the Berkeley Wireless Research Center and the Wireless Foundations Research Center. In 2001, he spent a year as a Research Scientist at the wireless startup Enuvis developing adaptive signal processing algorithms for extremely sensitive GPS receivers implemented using software-defined radio. Prior to that, he was a Graduate Student at the Laboratory for Information and Decision Systems, MIT. His research interests are in wireless communications, signal processing, and information theory. He is particularly interested in fundamental questions of spectrum sharing.