Covert Active Sensing of Linear Systems

Dennis Goeckel*, Boulat A. Bash†, Azadeh Sheikholeslami*, Saikat Guha‡, Don Towsley‡
*Dept. of Elec. and Comp. Engineering (ECE), Univ. of Massachusetts, Amherst
†Raytheon BBN Technologies, Cambridge, MA
‡The Univ. of Arizona, College of Optical Sciences, Tucson, AZ
‡College of Info. and Comp. Sciences (CICS), Univ. of Massachusetts, Amherst

Abstract—There has been significant recent work on the limits of covert communications, where the presence of a transmitted signal is kept hidden from an attentive adversary. In this paper, we turn our attention to the covert sensing problem. In particular, one of the main drawbacks of active sensing is that it can reveal the presence and/or location of the sensor. But active sensing is necessary or desirable in many applications, spanning areas from range finding in radar systems to channel estimation in wireless communication systems. Hence, we consider the active sensing of linear systems operating on the transmitted signal, but with the requirement that the probing signal is not detectable by a capable and attentive adversary. We first consider the frequency-nonselective standard “block fading” channel from wireless communications and identify the conditions under which covert active sensing is possible. We then turn our attention to the covert sensing of linear systems modeled as bandlimited wide-sense stationary random processes, where we present initial but not conclusive results. Finally, we give thoughts on the extension of the results to frequency-selective models and present ideas for future work.

I. INTRODUCTION

Passive sensing naturally provides covertness against an attentive adversary attempting to detect the presence of the sensor by measuring electromagnetic transmissions. In fact, one of the main drawbacks of active sensing is that such covertness may be compromised. However, passive sensing is not suitable for some applications, including various forms of channel estimation in wireless communication systems and radar range finding. Hence, we consider here the fundamental limits on the ability to perform active sensing of linear systems while remaining covert from an adversary.

Recently, the fundamental limits of covert communications have been considered in detail, and these results will be leveraged here. In particular, [1] considered the following problem: transmitter Alice wishes to communicate information reliably to legitimate receiver Bob, with whom Alice shares a secret, without detection of the presence of the transmitted signal by an attentive and powerful adversary Willie. When the receivers at Bob and Willie each experience additive white Gaussian noise (AWGN), [1] establishes that $O(\sqrt{n})$ bits (and no more) can be reliably and covertly transmitted, because Alice can remain covert only if she limits her energy to $O\left(\frac{1}{\sqrt{n}}\right)$ per channel use as $n \to \infty$. This power restriction, which will also be in evidence here, is explained intuitively, as follows: the standard deviation in the total noise power observed at Willie in $n$ observations scales as $\sqrt{n}$, the square root of the number of observations, and Alice can (and must) hide her signal in this uncertainty.

Further work [2], [3] has extended the consideration of covert communications to discrete memoryless channels and the consideration of the constants not given in the original scaling work of [1]. Work of Lee and Baxley [4], [5] and Sobers et al [6], [7] has demonstrated that uncertainty in the noise statistics at Willie or the presence of a jammer, respectively, can lead to the ability for Alice to transmit covertly with power not decreasing in the block length and hence transmit $O(n)$ bits in $n$ channel uses reliably; however, here we will assume no jammer is present and implicitly a model similar to [8], where there are multiple observation periods for Willie, and not all of these observation periods contain a sensing signal. Hence, the power limitations are similar to [1], which are also of more interest than constant power transmission, where the sensing results are easily established.

To get some insight into the problem, consider first a standard “block fading” channel, which changes $M$ times per second. With the help of a shared secret between Alice and Bob, we wish to consider Bob’s ability to estimate this channel accurately over a period $[0, T], T \to \infty$, based on pilot signals sent through the channel by Alice. At first glance, accurate sensing seems impossible. In particular, the results of [1] suggest that the total energy in each fading block, with duration $1/M$, must decrease to zero as $T \to \infty$, thus making accurate sensing of the fading value in each block impossible. But this ignores a key dimension of the problem. In particular, if the channel has bandwidth $W$, there are roughly $2WT$ degrees of freedom available to Alice and, roughly speaking, Alice should be able to employ $N = O(W)$ pilot pulses per second and use energy $O\left(\frac{1}{\sqrt{2WT}}\right)$ per each of the pilot pulses while remaining covert. This suggests that, if $W$ (and thus $N$) scales linearly with $T$, Alice can place pilots with aggregate constant energy $\left(\frac{N}{M} \cdot \frac{1}{\sqrt{2WT}} = c\right)$ in each fading block and thereby facilitate accurate channel estimation. It is this fundamental question that we seek to answer: what conditions on the scaling of $N$ with respect to $T$ allow or prevent accurate channel estimation for various assumptions on the process to be estimated?

In Section II, we present the system model and discuss its fit to applications of interest. Section III considers the limitations on the sensing waveform for covertness to be maintained.

This work was supported by the National Science Foundation under grants ECCS-1309573 and CNS-1564057, and by DARPA under contract number HR0011-16-C-0111.
Fig. 1: Covert Sensing Scenario: Without detection by the attentive and capable adversary Willie, Alice wishes to exploit a shared secret with Bob to transmit pilots so as to allow Bob to estimate a linear filtering channel. In the case of a single sensor (e.g. radar), note that Alice and Bob would be the same entity; the same models from the paper still roughly apply, as described in Section II-B.

Sections IV-A and IV-B then present theoretical results on the limits of covert sensing under assumptions of a block random process or a stationary bandlimited random process, respectively. Finally, Section V provides the conclusions.

II. SYSTEM MODEL AND METRICS

The scenario is shown in Figure 1. In this section, we precisely define the channel models, signaling schemes, and metrics of performance.

A. Mathematical Model and Metrics

1) Channel Model: Consider the estimation of a two-dimensional zero-mean, random process $h(\tau, t)$. Employing the standard model from [9], we assume that the process has response $h(\tau, t)$ at time $t$ due to an impulse at time $t - \tau$ in the input signal $x(t)$ from Alice and produces the output signal $y(t)$ at Bob given by

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t)x(t - \tau)d\tau + n^{(b)}(t)$$

where $n^{(b)}(t)$ is a white Gaussian noise process with (two-sided) power spectral density $S_{n^{(b)}(\cdot)} = \frac{N_0}{2}$.

Also present in the environment is a watchful adversary Willie, who we assume observes the transmitted signal without filtering but in the presence of additive white Gaussian noise (AWGN):

$$z(t) = x(t) + n^{(w)}(t)$$

where $n^{(w)}(t)$ is a white Gaussian noise process with (two-sided) power spectral density $S_{n^{(w)}(\cdot)} = \frac{N_0}{2}$.

We will primarily consider a frequency-nonselective slowly-fading model, which can be employed to demonstrate the key points with the extension to the full frequency-selective model discussed in Section V. The frequency-nonselective assumption implies $h(\tau, t) = h(t)\delta(\tau)$, thus yielding

$$y(t) = h(t)x(t) + n^{(b)}(t).$$

Per Section I, we will consider two different random process models for $h(t)$, as described in Section IV.

2) Signaling Model: Alice transmits, and Bob and Willie observe, the signal over a time period $[0, T]$. In general, Alice would be able to use an arbitrary probing waveform $x(t)$. However, here we restrict her probing waveform to consist of equally spaced binary probing pulses. Although this limits Alice’s options, thus providing a limitation on the converse results provided in this paper, as $n \to \infty$ the equally spaced pulses become arbitrarily dense, and thus the ability to employ an arbitrary power across these pulses roughly captures the degrees of freedom in signaling available to Alice.

Assume that we transmit $T/T_s$ pulses at pulse repetition interval $T_s$, with the $k^{th}$ pulse coded with coefficient $n \sqrt{P^{(k)}_a T_s a_k}$ to form:

$$x(t) = \sum_{k=0}^{T/T_s} n \sqrt{P^{(k)}_a T_s a_k} p(t - kT_s)$$

where $P^{(k)}_a$ is the power allocated to that pulse, $a_k \in \{-1, +1\}$ is an i.i.d. sequence, with each element equally likely to be $a_k = +1$ or $a_k = -1$, forming the shared secret between Alice and Bob, and $p(t)$ is a unit-energy pulse such that

$$\int_{-\infty}^{\infty} p(t)(p(t - kT_s)dt = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

The goal is to estimate the random process over the interval $[0, T]$. The pertinent question will be: what pulse repetition interval $T_s$ (or, roughly, bandwidth $1/T_s$) is required to estimate $h(t)$ accurately?

3) Metrics: Willie is faced with a binary detection problem and thus needs to perform a hypothesis test. In particular:

$$H_0 : \text{Alice is not actively sensing}$$
$$H_1 : \text{Alice is actively sensing}$$

The performance of such a hypothesis test is determined by its probability of false alarm $P_{FA} = P(\text{choose } H_1 | H_0)$ and probability of missed detection $P_{MD} = P(\text{choose } H_0 | H_1)$. In defining covertness, we follow prior work [1], [7] and define that a system is covert if, for any $\epsilon > 0$, $\lim_{T \to \infty} \frac{P_{FA} + P_{MD}}{2} > \frac{1}{2} - \epsilon$; that is, in the limit of a large sensing interval, Willie’s probability of error $P^{(W)}_e = P_{FA} + P_{MD}$ can be made arbitrarily close to a random selection of $H_0$ or $H_1$. The appropriateness of this criterion for hypotheses with unequal probabilities, as would generally be the case in practice, is discussed in detail in [7].

Given that the system is covert, we want to minimize the mean squared error in the estimation of $h(t)$ over the interval $[0, T]$, as $T \to \infty$.

B. Mapping to Applications

The paper will focus throughout on the problem given in the previous section: estimation of the function $h(t)$ over $[0, T]$ from the response to pulses put into that system observed in noise. In this section, we talk about how the work applies with a slight change of notation to other scenarios.
1) Channel Estimation in OFDM Systems: Many current wireless standards are built upon an orthogonal frequency division multiplexing (OFDM) framework [10], where each data symbol is effectively coded into a narrow frequency band, and the effective gain applied to that channel at time $i$ is given by $H(t, f_0)$, where $H(t, f)$ is the Fourier transform of $h(t, \tau)$ in the $\tau$ variable, and $f_0$ is the center frequency of the band. If we make the standard assumption that $H(t, f)$ does not change over a given OFDM symbol, we can suppress the time dependence and simply write $H(f)$ for the channel response impacting that symbol. Channel estimation for that OFDM symbol is done by inserting pilot symbols into the data stream; for these symbols, the effective received signal at Bob on the channel centered at $f_0$ after matched filtering and sampling is given by:

$$Y(f_0) = H(f_0)X(f_0) + N(f_0),$$

where $X(f_0)$ is the pilot symbol and $N(f_0)$ is the noise on the observation. Thus, we obtain a frequency domain version of (3). This can be treated exactly as we treat (3) below, with the exception that the time and frequency variables are interchanged. For example, when we seek the bandwidth $W$ of probing required to estimate $h(t)$ over $[0, T]$ accurately based on (3), this is equivalent to asking for what time duration must the channel $H(f)$ be constant so that we can estimate $H(f)$ over a bandwidth $[0, W]$ in (6). The random process models for $h(t)$ also are applicable to the random process $H(f)$.

2) Range Fiding in Radar Systems: The model can also be applied approximately to the use of a pulsed radar to attempt to estimate the range $h(t)$ of a target over $[0, T]$, as follows. Suppose that a pulsed radar with energy $P_a T_s$ employed in the $k$th pulse is probing the target. The mean squared error in the location from a pulsed radar is inversely proportional to the power of the signal [11], and thus we can write a model for the range measurements as:

$$y_k = h(kT_s) + \sqrt{P_a T_s} n_k^{(b)} + a_k n_k^{(w)}, \quad k = 0, 1, \ldots, T/T_s.$$  

where $n_k^{(b)}$ is a zero-mean random variable with variance equal to the mean squared error of the radar when $\sqrt{P_a T_s} = 1$. Noting the similarity between (7) and (18), our model will be approximately applicable to the radar problem. The match is approximate because, to our knowledge, the error variable $n_k^{(b)}$ in (7) is not necessarily Gaussian.

III. DETECTION AT WILLIE

As will be justified below, assume for this analysis that the power $P_a^{(b)}$ does not vary with symbol index $k$, and hence we denote it by $P_a$. Willie is faced with the hypothesis test:

$$H_0 : z(t) = n^{(w)}(t),$$

$$H_1 : z(t) = \sum_{k=0}^{T/T_s} \sqrt{P_a T_s} a_k p(t - kT_s) + n^{(w)}(t)$$

From (5), the set \{p(t - kT_s), k = 0, 1, \ldots, T/T_s\} forms a complete orthonormal basis for the waveform transmitted from Alice. Hence, it is sufficient for Willie to form $Z_k = \int_{-\infty}^{\infty} z(t)p(t - kT_s)dt$ for $k = 0, 1, \ldots, T/T_s$ and base his decision upon $\{Z_k, k = 0, 1, \ldots, T/T_s\}$, thus yielding the hypothesis test:

$$H_0 : Z_k = N_k^{(w)}$$

$$H_1 : Z_k = \sqrt{P_a T_s} a_k + N_k^{(w)}$$

where $N_k^{(w)}$ is an i.i.d. sequence where $N_k^{(w)} \sim N(0, N_k^{(w)})$. Hence, Willie is faced with an identical hypothesis test to that of Theorem 1.2 in [1] with pulse amplitude $a = \sqrt{P_a T_s}$.

From [1], $a^2 \leq 2cN_0^{(w)}$, where $c$ is a constant proportional to $\epsilon$, implies that the communication is covert. Hence, there exists a constant $c_1$ such that:

$$P_a \leq c_1 N_0^{(w)} \sqrt{T/T_s}.$$  

is sufficient for covertness to be obtained. If we let the number of pulses per second be $N = 1/T_s$ and look at the energy per pulse, we get the sufficient condition $P_a T_s \leq c_1 N_0^{(w)} / \sqrt{NT_s}$, which is consistent with [1].

Conversely, if

$$P_a = \omega \left( \frac{1}{\sqrt{T/T_s}} \right),$$

then Willie can detect Alice’s transmission with arbitrarily low error probability.

IV. ESTIMATION ERROR

Section III establishes the power that can be employed while covertness is maintained. Here, given these power constraints, we consider the ability to estimate the linear system of interest. Per Section I, we consider two channel models: a “block fading” model, as commonly used in wireless communications [12], and a bandlimited wide-sense stationary (WSS) random process model.

A. M-Block Fading Model

Define the “rectangle” function as:

$$\text{rect}(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 0, & \text{else} \end{cases}$$

The block fading model is given by:

$$h(t) = \sum_{i=0}^{MT-1} h_i \text{rect}(M(t - i/M))$$

where $h_i, i = 0, 1, 2, \ldots, MT - 1$, is a collection of independent and identically distributed zero-mean random variables of unknown distribution. Thus, Bob seeks to estimate $\{h_i : i = 0, 1, \ldots, MT - 1\}$ from his observations. For fixed $M$ and sufficiently small $T_s$, the duration of $p(t - t_0)$ is sufficiently short such that it is affected only by $h(t)$ at $h(t_0)$. Thus, Bob observes:

$$y(t) = \sum_{k=0}^{T/T_s} \sqrt{P_a T_s} a_k h(kT_s) p(t - kT_s) + n^{(b)}(t)$$
1) Achievability: Suppose Bob forms
\[ y_k = \frac{a_k}{\sqrt{P_a T_s}} \int_{-\infty}^{\infty} y(t)p(t - kT_s)dt \]
where \( k = 0, 1, 2, \ldots, T/T_s \) \( (17) \)
Recall that \( N = 1/T_s \) is the number of pulses per second. Thus, Bob observes the \( NT \) samples:
\[ y_k = h_k + n_k^{(b)} + \sqrt{P_a T_s} \]
where \( \{n_k^{(b)}, k = 0, 1, \ldots, NT\} \) is a collection of i.i.d. Gaussian random variables, each with variance \( N_0^{(b)}/2 \). Thus, for any \( m, m = 0, 1, \ldots, MT - 1 \), Bob has \( N/M \) samples to estimate \( h_m \), with each sample corrupted by independent noise of variance \( N_0^{(b)}/2P_a T_s \). If he employs the estimator
\[ \hat{h}_m = \frac{1}{N/M} \sum_{i=0}^{N/M-1} y_m \cdot \frac{n_i^{(b)}}{\sqrt{P_a T_s}} \] (19)
the estimator is unbiased and has variance:
\[ \text{Var}(\hat{h}_m) = \frac{1}{N/M} \cdot \frac{N_0^{(b)}}{2P_a T_s} \]
\[ = \frac{N_0^{(b)}}{c_1 N_0^{(w)}} \cdot \frac{\sqrt{T_s}}{M} \cdot \frac{1}{\sqrt{N}} \] (20)
where the covariance requirement from (12) has been enforced in the second line. Thus, scaling \( N \) (roughly the bandwidth of the probing pulse) as \( O(M^2T) \) can yield a small mean-squared error per \( h_m \).

**Discussion:** A larger bandwidth gives us more degrees of freedom (and requires a larger shared secret) that Willie must monitor for Alice’s signal per coherence time of \( h(t) \). Hence, as bandwidth increases, with associated cost in hardware and shared secret, Alice and Bob’s advantage increases. Note that assuming a frequency-nonselective model for an arbitrarily large bandwidth is physically problematic; this will be addressed in our extension to the frequency-selective model in Section V. Without a covertness requirement, a system could use constant energy per pulse, and hence would only require \( O(1) \) pulses per fading block; hence, accurate per-block sensing could be done with \( N = O(M) \) pulses per second.

2) Converse: It is straightforward to show that the power in a block should be divided equally between the pulses in that block. What remains is to show that the power should be divided equally among the blocks. If we generalize (20) to the case when the power varies per block as \( P_a^{(m)} \), the mean squared error is:
\[ \text{Var}(\hat{h}_m) = \frac{1}{(N/M)} \cdot \frac{1}{M} \cdot \sum_{m=0}^{M-1} \frac{N_0^{(b)}}{2P_a^{(m)} T_s} \] (22)
which is minimized when \( P_a^{(m)} = P_a \); in other words, the power (not unexpectedly) should be fixed across the blocks. This allows us to employ the converse result of Section III, from which it follows that \( O(M^2T) \) pulses per second is necessary for arbitrarily small error per block fading coefficient.

**B. Bandlimited Random Process**

Let \( h(t) \) be a lowpass stationary random process (i.e. power spectral density non-zero only in the frequency range \([-B, B]\)] with autocorrelation function \( R_h(\tau) = E[h(t)h^*(t + \tau)] \). Because the bandwidth is limited, we can assume that the appreciable duration of \( p(t - t_0) \) is sufficiently short such that it is affected only by \( h(t) \) at \( h(t_0) \) as \( N \to \infty \) (and thus \( T_s \to 0 \)). Thus, Bob observes:
\[ y(t) = \sum_{k=0}^{T/T_s} \sqrt{P_a^{(k)} T_s} a_k h(kT_s) p(t - kT_s) + n^{(b)}(t) \] (23)
To estimate \( h(t) \) over \([0, T]\) Bob forms:
\[ y_k = \frac{a_k}{\sqrt{P_a T_s}} \int_{-\infty}^{\infty} y(t)p(t - kT_s)dt \]
where \( \{n_k^{(b)}, k = 0, 1, \ldots, NT\} \) is a collection of i.i.d. Gaussian random variables, each with variance \( N_0^{(b)}/2 \), and \( \{P_a^{(k)} T_s, k = 0, 1, \ldots, T/T_s\} \) is the sequence of energies of the probing symbols.

Recall that \( N = 1/T_s \) is the number of pulses per second, and consider the estimation of \( h(kT_s), k = 0, 1, \ldots, TN \). Define the \( NT \) by \( 1 \) vectors:
\[ y = [y_0, y_1, \ldots, y_{TN}]^T \]
\[ h = [h(0), h(T_s), \ldots, h(T \cdot N \cdot T_s)]^T \]
\[ = [h(0), h(T_s), \ldots, h(T)]^T \]
\[ n = \left[ \frac{n_0^{(b)}}{\sqrt{P_a^{(0)} T_s}}, \frac{n_1^{(b)}}{\sqrt{P_a^{(1)} T_s}}, \ldots, \frac{n_{TN}^{(b)}}{\sqrt{P_a^{(TN)} T_s}} \right]^T \] (28)
First, note that using equal-power pilots, as assumed in Section III to obtain the covertness constraint (12), is optimal. To argue such, we will employ the result of [13]. Unfortunately, the pilot powers considered in [13] are constant while the covertness constraint imposes a power per pilot that must necessarily decrease as the number of pilots increases. Fortunately, using Weyl’s inequality it is straightforward to show that when the power of each pilot changes with the number of pilots \( NT \) as in (12), the result obtained in [13] still holds. Hence, we can use pilots with equal powers \( P_a = P_a^{(0)} = P_a^{(1)} = \cdots = P_a^{(TN)} \), and thus the vector \( n \) can be written as:
\[ n = \left[ \frac{n_0^{(b)}}{\sqrt{P_a^{(0)} T_s}}, \frac{n_1^{(b)}}{\sqrt{P_a^{(1)} T_s}}, \ldots, \frac{n_{TN}^{(b)}}{\sqrt{P_a^{(TN)} T_s}} \right]^T \] (29)
Now suppose Bob forms a minimum mean squared error (MMSE) estimate of \( h \). Let \( R_{TN} = E[hh^H] \) be the covariance matrix of \( h \), and \( D_{TN} = E[nn^H] \) be the covariance matrix.
of \( n \). The covariance matrix of the estimation error \( \mathbf{e} = \mathbf{h} - \hat{\mathbf{h}} \) is [13]:

\[
M_{TN} = R_{TN} - R_{TN}(R_{TN} + D_{TN})^{-1}R_{TN}
\]

where \( (31) \) follows from [14, Section 0.7.4]. Each diagonal element of \( M_{TN} \) is the mean squared error of an observation. Hence, the average mean squared error per sample in estimating \( \mathbf{h} \) is [13]:

\[
\mathcal{E} = \frac{1}{TN} \text{Tr} \left[ (R_{TN}^{-1} + D_{TN}^{-1})^{-1} \right] \quad (32)
\]

Substituting \( D_{TN} = P_aT_nI \) where \( I \) is an \( TN \times TN \) identity matrix yields,

\[
\mathcal{E} = \frac{1}{TN} \text{Tr} \left[ (R_{TN}^{-i} + \left( \frac{1}{P_aT_nI} \right)^{-1})^{-1} \right]. \quad (33)
\]

Recalling the coverness criterion, set \( P_a = c \sqrt{\frac{1}{TN}} \) to yield:

\[
\mathcal{E} = \frac{1}{TN} \text{Tr} \left[ (R_{TN}^{-i} + \left( \frac{\sqrt{TN}}{c} f \right)^{-1})^{-1} \right] = \frac{1}{TN} \sum_{i=0}^{TN} \frac{1}{\lambda_i} \frac{1}{\sqrt{\lambda_i^2 + \frac{c^2}{TN} f^2}} \quad (34)
\]

where \( \lambda_i, i = 0, 1, \ldots, TN \) are the eigenvalues of \( R_{TN} \), and (34) follows from the fact that if \( \{ \alpha_i \} \) are eigenvalues of a matrix \( A \), \( \text{Tr}[A^{-1}] = \sum \alpha_i^{-1} \).

Equation (34) shows how the behavior of the eigenvalues of \( R_{TN} \), as \( T \to \infty \), will determine the ability to covertly sense the bandlimited process. Here, we give some initial but not conclusive thoughts on the sum in (34). Using [15], the eigenvalues are (asymptotically) proportional to samples of the Fourier transform of \( R_h(\tau) \) truncated to \([0, T] \); hence, they are the samples of \( S_Z(f) = S_h(f) * T \text{sinc}(Tf) \). For increasing \( x \), the \( \text{sinc}(x) \) function decays as \( 1/x \), and hence \( S_Z(f) \) only decays at a rate \( 1/f \), thus suggesting that the sum in (34) may not go to zero as \( TN \) becomes large, even if \( N \) is chosen to scale at a rate faster than \( T \). It is possible to see how such a result might occur: even though we are getting fine sampling of the process on \([0, T] \), we do not observe the process outside \((0, T) \); since the heavy tails of the \( \text{sinc}(.) \) mean that those samples can influence \( h(t) \) on \((0, T) \) appreciably, this “blindness” and the limited power required for coverness might lead to a lower bound on the estimation error.

V. CONCLUSIONS

In this paper, we have introduced the covert sensing problem: Alice and Bob desire to sense a channel \( h(t, \tau) \) over \([0, T] \) without being detected by attentive adversary Willie. We consider in detail the frequency-nonelective case. For an \( M \)-block fading channel model, we investigate the bandwidth required for the channel to be accurately and covertly sensed. We find that the bandwidth of the sensing signal must scale linearly with \( T \), the length of the time interval, in order to provide the advantage Alice and Bob require for accurate covert sensing of the block fading channel.

In the case of sensing a bandlimited random process, as considered in Section IV-B, we formulate the problem and provide an expression for the mean-squared error when sensing is performed covertly. However, unlike the case of block fading, we do not have conclusive analytical results on the scaling required for the bandwidth of the sensing signal, and numerical results are inconclusive. Hence, investigating the convergence of (34) for various settings of \( N \) (as a function of \( T \)) is currently under investigation.

These results can be readily extended to the case of frequency-selective fading, in particular in the case of block fading. Suppose that a channel of bandwidth \( W \) is to be estimated over duration \( T \), and the channel can be modeled as a collection of two-dimensional “tiles” in the time-frequency plane, each of which fades independently. From Section IV, it is then apparent that accurate channel estimation is possible if and only if the number of independent fading variables to be estimated is in a bandwidth of \( W \) and time period of \( T \) is on the order of \( O(\sqrt{TW}) \).

REFERENCES