Abstract—Cognitive radios process their sensed information collectively in order to opportunistically identify and access underutilized spectrum segments (spectrum holes). Due to the transient and rapidly varying nature of the spectrum occupancy, the cognitive radios (secondary users) must be agile in identifying the spectrum holes in order to enhance their spectral efficiency. We propose a novel adaptive procedure to reinforce the agility of the secondary users for identifying multiple spectrum holes simultaneously over a wide spectrum band. This is accomplished by successively exploring the set of potential spectrum holes and progressively allocating the sensing resources to the most promising areas of the spectrum. Such exploration and resource allocation results in conservative spending of the sensing resources and translates into very agile spectrum monitoring. The proposed successive and adaptive sensing procedure is in contrast to the more conventional approaches that distribute the sampling resources equally over the entire spectrum. Besides improved agility, the adaptive procedure requires less-stringent constraints on the power of the primary users to guarantee that they remain distinguishable from the environment noise and renders more reliable spectrum hole detection.

Index Terms—Adaptive, agility, sparsity, spectrum sensing, wideband.

I. INTRODUCTION

A. Background

The notion of cognitive radio has emerged as a mean to alleviate the scarcity of available frequency spectrum through accommodating unlicensed (secondary) users within the underutilized spectrum bands licensed to the legacy (primary) users. For this purpose, secondary users monitor the spectrum in order to identify and exploit the underutilized segments. Such coexistence of the primary and secondary users is constrained by the necessary warranties on the quality of service provided for the primary users.

For detecting an underutilized segment in the wideband spectrum, which we hereinafter call a spectrum hole, secondary users face two major challenges. First, the spectrum holes are spread across the wideband and often their occupancy status changes rapidly, i.e., the holes might not remain unoccupied for a long time. Therefore, the amount of time spent searching for a spectrum hole could potentially constitute a considerable fraction of the time that a spectrum hole remains vacant. Such rapid changes are governed mainly by the spectral activities of the active (primary and other potential secondary) users seeking access to the same spectrum. Therefore, the secondary users should be agile enough to detect a spectrum hole in substantially less time than the period in which the hole remains vacant. Agility becomes even more prominent in wideband systems where the secondary users have to monitor a larger number of channels with a sparse distribution of spectrum holes. The current statistics about the spectrum occupancy patterns provide that a considerable portion of the spectrum is underutilized, hence promoting the notion of cognitive communication. Nevertheless, by envisioning the futuristic scenario of cognitive networks, enabling cognitive communication allows many unlicensed users to compete for the same spectrum resources. Under this envisioned scenario and by taking into account the ever-increasing demand for data communication, the vacant spectrum bands will not be as abundant as they are, which makes the sparsity assumption quite reasonable. We would like to also remark that the same futuristic sparsity assumption on spectrum vacancy is also envisioned in [5] and [6]. The second challenge is to avoid harming the communication of the primary users. For this purpose, the secondary users must reliably distinguish the holes from the spectrum segments accessed by the primary users, regardless of how weak the primary users are.

Sensing the spectrum, being a major task of the secondary users, has received a considerable amount of research interest (see, e.g., [4] and [7]–[11]). We review some of the related works in the following section.

B. Related Works

There exist, broadly, three approaches for identifying spectrum vacancies within a wideband spectrum. A notable approach on agile spectrum sensing in a wideband spectrum is the quickest sequential search scheme of [4]. In this scheme, the spectrum is split into smaller narrowband channels and the secondary user senses one channel at a time. After accumulating enough information about each channel, it decides whether the channel is a hole or is occupied. If the channel is decided to be a hole, the search is terminated and otherwise the process is carried on until a hole is spotted. As the spectrum holes become less frequent (sparser) enhancing the agility of the secondary users becomes more significant. It is noteworthy that when the spectrum holes are not sparse, the quickest detection approach [4] is very effective. In such scenarios, the expected density of the spectrum holes is high and it becomes very likely to encounter a spectrum hole after sensing a few channels.
Therefore, by deploying the sequential quickest detection approach, the expected amount of sampling resources required for identifying multiple holes is small and thereof the overall detection process will be very agile. For a sparse distribution of spectrum holes, however, sensing the channels sequentially for identification of multiple holes will substantially lengthen the expected sensing time. Therefore, the objective of our proposed adaptive procedure is to address the issue of agility for a sparse distribution of spectrum holes and when we are interested in identifying multiple holes, as it is more challenging and realistic in practice. We also remark on the recent results on sequential binary hypothesis testing of multiple sequences [12] where a sparse number of the sequences are generated according to the null distribution and the majority of them according to an alternative distribution. It develops a sequential thresholding-based test for identifying almost all the sequences generated according to the null distribution. These tests tend to identify almost all of the sequences generated by the null distribution.

In another approach, it is assumed that the wideband channel is heavily underutilized and it is used only rarely and sparsely. In this approach, the cognitive radios exploit the sparsity structure of the wideband channel and construct compressed sensing-based machineries for estimating the power spectral density (PSD) of the wideband channel [13]–[19]. These approaches are further extended to also track temporal variations of the spectral occupancy during sensing [20], [21]. Exploiting the sparsity empowers the cognitive radios to sample the signal activity over the channel at a sub-Nyquist rate, which expedites the process of estimating the PSD. While sampling at a sub-Nyquist rate substantially reduces the number of samples required for reconstructing the PSD, the number of samples required yet scales linearly with the bandwidth of signals of the active users. Therefore, the amount of required samples for reconstructing the PSD of large-bandwidth signals (e.g., video signals) is substantially large. When each of the channels is narrow band, however, estimating the PSD encompasses much redundancy as the ultimate objective of the cognitive user is to make a 1-bit decision about a channel (hole versus occupied). By taking this fact into account, our proposed sensing procedure further reduces the sampling resources below the level required by the compressed-sensing-based methods.

In another approach tailored for cognitive networks, a group of cognitive radios are clustered to collaboratively perform spectrum sensing [5], [11], [16], [22], [23]. At the cost of extra communication among the cognitive radios, they can share their individual perception about the spectral occupancy and reach a more reliable decision through exchanging their individual perceptions (coordination). Such information exchange is viable in cognitive networks with a fusion center or an established infrastructure through which the autonomous users can be coordinated. In this paper, we propose a novel adaptive spectrum sensing scheme for ad hoc cognitive networks. The underlying premise is that one can adaptively decide how to spend a given sampling budget in the course of the spectrum sensing process, focusing the sensing resources on more promising segments of the spectrum. This is in contrast to the conventional spectrum sensing schemes with a pre-defined sensing strategy. Our proposed adaptive procedure consists of two phases, namely the exploration and detection phases. The errand of the exploration phase is to swiftly eliminate a substantial portion of the occupied channels and at the same time retain most of the holes. Exploration is carried out through an iterative process, where in each iteration a group of less promising segments of the spectrum are identified and eliminated. The fundamental basis of the exploration phase is that it can be done accurately even if the sensed information (measurements) is very rough and noisy. During exploration, each iteration further monitors the spectrum and improves on the outcome of the previous iteration. The exploration phase is subsequently followed by a detection phase for identifying multiple spectrum holes among the reduced number of channels retained.

The analysis reveals that if we perform multiple cycles of the proposed explorations, for achieving a certain level of reliability, the ratio of the sampling budget required by the proposed adaptive procedure to that required by the nonadaptive procedure is approximately \(\frac{2}{M}\), where \(M\) is the sampling budget per channel in the nonadaptive procedure. Furthermore, the adaptive procedure requires less-stringent constraints on the power of the active users to guarantee a successful hole detection. In particular, let \(\epsilon_n\) be the probability that each channel is a spectrum hole. Then, the requirement for identifying spectrum holes by the nonadaptive procedure is that the power of the active users must scale as \(\omega \left(\sqrt{\frac{1}{\epsilon_n}}\right)\), where \(M\) is the sampling budget per channel. By deploying the adaptive procedure, this requirement reduces to approximately \(\omega \left(\sqrt{\frac{1}{\epsilon_n}}\right)\), where \(M'\) is an integer substantially larger than \(M\).

The exploration phase and the idea of successively searching for the spectrum holes is inspired by the notion of distilled sensing developed in [1]–[3]. In distilled sensing, it is shown that given a fixed sampling budget, certain signals that are detectable/estimable using adaptive measurements cannot be recovered using nonadaptive strategies. The results show that closing the loop between the data analysis and collection processes can yield very significant gains. Despite the similarities between our exploration phase and distilled sensing in [1]–[3], there exist also significant differences mainly in two aspects.

1) We desire to identify any arbitrary number (and not necessarily all) of the spectrum holes where distilled sensing

\[1\] For convenience, we will use the following asymptotic notations: for two sequences \(a_n \geq b_n \geq 0\) we say that \(a_n = o(b_n)\) if \(\lim_{n \to \infty} \frac{a_n}{b_n} = 0\) and \(a_n = \omega(b_n)\) if \(\lim_{n \to \infty} \frac{b_n}{a_n} = 0\). Also \(a_n = o(b_n)\) indicates asymptotic equality of \(a_n\) and \(b_n\), i.e., \(\lim_{n \to \infty} \frac{a_n}{b_n} = 1\).

\[2\] The precise characterization is given in Section IV.
aims to identify all data points that bear a certain sparsity structure.

2) The observation models are different. The works in [1]–[3] consider a Gaussian observation model and the theoretical guarantees depend on the Gaussian assumption. On the other hand, in the spectrum sensing problem in hand, we deal with data drawn from Gamma distributions. The analysis for the Gamma distribution and Gaussian distributions are significantly different. This is due to the fact that the experimental design for distilled sensing is primarily shaped up by the statistical model of the data samples, which in turn necessitates independent analysis for each class of distributions.

The remainder of this paper is organized as follows. In Section II, the system model and the formal statement of the problem are provided. Section III introduces the nonadaptive hole detection scheme, which serves as a baseline for assessing the performance of the proposed adaptive procedure. The nonadaptive hole detection scheme also represents the structure of the detection phase that we deploy in the adaptive procedure. In Section IV, we describe the proposed adaptive hole detection procedure and its related analysis. Simulation results are provided in Sections V, and Section VI provides some concluding remarks.

II. SYSTEM DESCRIPTIONS AND PROBLEM STATEMENT

A. System Setup

Consider a wideband spectrum to be shared by license-holding (primary) users and cognitive (secondary) users with interference-avoiding spectrum access. The primary users are formally allocated a number of channels in the spectrum and have the right to instantaneously access the channels whenever desired. The secondary users, on the other hand, are allowed to opportunistically seek the portions of the spectrum underutilized by the active users (either primary users or secondary ones currently using the spectrum) and access them. Our objective is to provide an agile and reliable mechanism for identifying such communication opportunities for the secondary users.

We assume that the available spectrum consists of $n$ nonoverlapping channels, indexed by $\{1, \ldots, n\}$. At a given instance, the active users communicate over some of these channels, which we refer to as occupied channels, and underutilize the rest, which we call spectrum holes. We assume that the occupancy status of the spectrum remains unchanged during spectrum sensing process and consider a probabilistic model for the occupancy of the spectrum. Let the Bernoulli random variable $Z_i$, for $i \in \{1, \ldots, n\}$, indicate the occupancy state of the $i$th channel, where $Z_i = 1$ means that the $i$th channel is occupied and $Z_i = 0$ states that the $i$th channel is a spectrum hole. We assume that the channels have statistically independent occupancy states and each channel is a spectrum hole with probability $\epsilon_n$. Therefore

$$Z_i \overset{i.i.d.}{=} \text{Bernoulli}(1 - \epsilon_n).$$

Let us also denote the set of the indices of the spectrum holes and the occupied channels by

$$\mathcal{H}_0 \triangleq \{i \in \{1, \ldots, n\} : Z_i = 0\}$$

$$\mathcal{H}_1 \triangleq \{i \in \{1, \ldots, n\} : Z_i = 1\}.$$  

In order to capitalize on the availability of spectrum holes, the secondary users monitor the spectrum via channel measurements. Each measurement of channel $i$, denoted by $X_i$, is of the form

$$X_i = \sqrt{p_i}H_i S_i Z_i + W_i, \quad \text{for} \quad i \in \{1, \ldots, n\}$$

where $p_i$ accounts for the combined effect of the transmission power of the user active on the $i$th channel and its associated signal attenuation (for $i \in \mathcal{H}_1$), $H_i$ denotes the flat-fading channel from the active user to the secondary user and is distributed as zero-mean unit-variance complex Gaussian. Also, $S_i$ denotes the normalized unit-power complex Gaussian. Finally, $W_i$ denotes the additive white Gaussian channel noise distributed as zero-mean unit-variance complex Gaussian. The measurements are statistically independent over time and channels.

Finally, we assume that the powers of the active users sensed by the secondary user are lower bounded by $\gamma_n$, for some arbitrary $\gamma_n > 0$, i.e., $p_i \geq \gamma_n$ for all $i \in \mathcal{H}_1$. Acquiring $\gamma_n$ is feasible for secondary users if the geographical extent (and subsequently the maximum path loss) and the range of the transmission power of the active users are known to them. Clearly, the likelihood of channel vacancy $\epsilon_n$ and the lower bound $\gamma_n$ influence the precision of successfully detecting a hole; increasing $\gamma_n$ and $\epsilon_n$ enhances the reliability of the spectrum sensing process.

B. Spectral Monitoring Goal

The main objective is to use measurements of form (4) in order to identify $T \in \mathbb{N}$ spectrum holes, i.e., $T$ element of $\mathcal{H}_0$. The conventional nonadaptive spectrum sensing procedures perform some prespecified measurements of the spectrum and possibly locate $T$ spectrum holes based on the information extracted from the measurements. This strategy is nonadaptive, in the sense that the measurement process is fixed a priori and does not change during the experiment. In contrast, we devise an adaptive procedure such that the measurement strategy is adjusted sequentially such that future measurements use information gathered from previous ones. We demonstrate that such measurement adaptation substantially improves the reliability and agility of the secondary users in detecting spectrum holes.

We assess the agility by quantifying the amount of time required for sensing the channel. Agility depends linearly on the aggregate amount of measurements (sampling budget) made over the entire wideband spectrum, assuming that the measurements take the same amount of time. Therefore, increasing the sampling budget translates into more delay (or less agility) in detection. On the other hand, increased sampling budget improves the detection reliability. Therefore, for detecting multiple spectrum holes, there exists a tradeoff between agility and reliability. Incorporating the impact of channel occupancy
likelihood (characterized by $\epsilon_n$) and the power of active users establishes that in both adaptive and nonadaptive procedures, there exists an inherent interplay among the agility, reliability, channel occupancy likelihood, and the power of active users.

We characterize such interplays in both nonadaptive and the proposed adaptive sensing procedures. Comparing these interplays subsequently demonstrates the agility and reliability gains afforded by the adaptive procedure. As we consider wideband channels, the analysis provided is asymptotic with respect to large number of channels, $n$.

III. Nonadaptive Spectrum Detection

In this section, we analyze an optimal nonadaptive detection scheme for locating $T$ spectrum holes. The development of this scheme and the ensuing analysis serve a twofold purpose. On one hand, this detection scheme is also deployed in the detection phase of the adaptive procedure proposed in Section IV. On the other hand, it offers a baseline for assessing the gain yielded by the adaptive procedure.

A. Nonadaptive Sensing Procedure

Constructing a nonadaptive spectrum sensing procedure involves two issues. The first one pertains to the experimental design, which is the design of the information-gathering process. In our setup, the experimental design elucidates the distribution of the sampling budget among the channels across the wideband spectrum. The second issue is to design a detector based on some optimality criterion. The decision on the experimental design hinges on the expected proportion of spectrum holes, i.e., $\epsilon_n$. Prior to the sensing procedure, there is no extra side information about spectral activity in any of the channels and all the channels are equally likely to be spectrum holes. Due to the inherent symmetry and the sparse distribution of the spectrum holes ($\epsilon_n \in (0,1)$), we assume that the experimental design measures all channels equally (note that if $\epsilon_n$ is large, this is not the best strategy).

Given this experimental design, it is straightforward to construct a detector that is optimal in the sense that it maximizes the \textit{a posteriori} probability of successfully detecting $T$ spectrum holes, i.e., the maximum \textit{a posteriori} (MAP) detector. Suppose that the measurement budget is $B \in \mathbb{N}$, meaning that we make a total of $B$ measurements of the form (4). Therefore, the experimental design obtains $M \hat{\Delta} = \lfloor B/n \rfloor$ measurements of the form (4) per channel. Define $X_i(\cdot)$ as the $j$th measurement of the $i$th channel. Recalling the assumption that the occupancy statuses do not change over the course of sensing, the terms $\{p_i\}$ and $\{Z_i\}$ remain constant for all $i \in \{1, \ldots, M\}$, while the fading coefficient $H_i(\cdot)$, the noise term $W_i(\cdot)$, and the transmitted symbols $S_i(\cdot)$ might change from one observation to another.

Defining the vectors $X_i \hat{\Delta} = [X_i(1), \ldots, X_i(T)]$, $H_i \hat{\Delta} = [H_i(1), \ldots, H_i(T)]$, $S_i \hat{\Delta} = [S_i(1), \ldots, S_i(T)]$, and $W_i \hat{\Delta} = [W_i(1), \ldots, W_i(T)]$, based on (4) the observation set is given by

$$D_n \hat{\Delta} = \{ X_i : X_i = \sqrt{p_i} Z_i \cdot H_i \circ S_i + W_i, \text{ for } i = 1, \ldots, n \}$$

where “$\circ$” denotes the Hadamard product. From (4), it follows that given the occupancy status $Z_i$, the observation sample $X_i(\cdot)$ has a complex Gaussian distribution with mean zero and variance $\{1 + p_i Z_i\}$, i.e.,

$$X_i(j) \mid Z_i \sim \mathcal{C}\mathcal{N}(0, 1 + p_i Z_i), \quad \text{for } j \in \{1, \ldots, M\}.$$  \hfill (6)

According to (1), $Z_i$ has a Bernoulli distribution with mean $1 - \epsilon_n$. Therefore, the measurement vector $X_i$ is a sample from a mixture distribution with the probability density function (PDF)

$$P(X_i) = \epsilon_n \cdot \frac{1}{\pi M} \exp \left( -\|X_i\|^2 \right) + (1 - \epsilon_n) \cdot \frac{1}{\pi (1 + p_i)^M} \exp \left( -\frac{\|X_i\|^2}{1 + p_i} \right).$$  \hfill (7)

Given the observation set $D_n$, the MAP criterion for identifying $T$ spectrum holes entails finding the subset of channels with cardinality $T$ that has the highest probability of containing only holes and is formalized in the next remark.

Remark 1: The MAP rule for detecting $T$ holes is given by

$$\hat{U}^\text{MAP}_n = \arg \max_{\{U_i \subseteq H_i \mid |U| = T\}} \min_{\{U_i \subseteq H_i \mid |U| = T\}} P(U \subseteq H_i \mid D_n)$$

where

$$U_i \triangleq \frac{1}{1 + p_i} \exp \left( \frac{p_i}{1 + p_i} |X_i|^2 \right).$$

Proof: See Appendix H.

Due to the dependence of $\{U_i\}$ on $\{p_i\}$, the posterior probability heavily depends on the power of the active users $\{p_i\}$. The uncertainties about $\{p_i\}$ translates into uncertainty about any performance measure of interest, which in turn prevents us from obtaining a \textit{globally} (over all power realizations) optimal hole detector. As a remedy, by alternatively exploiting the notion of \textit{robustness} in the worst case, we can offer some worst case guarantees. In particular, we look at the realizations of $\{p_i\}$ which yield the \textit{worst case} performance for detecting $T$ holes. For this purpose, define $\Psi(\gamma_n)$ as the class of active users satisfying the power constraint $p_i \geq \gamma_n$ for all $i \in H_i$. Then, the worst case detection performance in the $\Psi(\gamma_n)$ class of active users is given by

$$\min_{\{p_i \leq \gamma_n\}} P(\hat{U}^\text{NAP}_n \subseteq H_i).$$

Note that for $i \in H_1$, the $i$th channel becomes less-distinguishable from a spectrum hole as $p_i$ decreases. Therefore, due to the independence among the distinct channels, the weakest detection performance, which occurs when the nonholes are least distinguishable from the holes, corresponds to the smallest possible choices of $\{p_i\}$, i.e., $p_i = \gamma_n$ for $i \in H_1$. For this setting, finding the indices of the $T$ smallest elements of the set $\{U_1, \ldots, U_n\}$ becomes equivalent to finding the indices of the $T$ smallest elements of the set $\{ |X_1|^2, \ldots, |X_n|^2 \}$. Therefore, in order to locate $T$ spectrum holes, the MAP detector requires only the sufficient statistics $Y_i \hat{\Delta} = |X_i|^2$ for $i = 1, \ldots, n$. Corresponding to the sequence of random variables $\{Y_1, \ldots, Y_n\}$, we define
and distribution of the involved random variables changes simultaneously, as well as the analysis for higher order statistics, for which the existing results are not applicable.

Lemma 1: Let \( \{ Y_i \}_{i=1}^{n} \) be a sequence of i.i.d. random variables distributed as Gamma \( \{ M, \alpha_m \} \) and denote its corresponding sequence of order statistics by \( \{ Y^{(i)} \}_{i=1}^{n} \). Let \( b_m \triangleq \alpha_m \left[ \frac{\Gamma(\frac{m+1}{m})}{m} \right] \), and for some \( T \in \mathbb{N} \) define the sequence of random variables \( W^{(i)}_{T} \) for \( i = 1, \ldots, T \). Then, as \( m \to \infty \), \( W^{(i)}_{T} \) converges in distribution to a random variable \( W^{(i)}_{T} \) with cumulative density function (CDF)

\[
Q_{(i)}(w; m) \triangleq P(W^{(i)}_{T} < w) = 1 - \exp(-w^M) \sum_{k=0}^{i-1} \frac{w^{kM}}{k!}
\]

Proof: See Appendix A.

For the setting of Section II-a, the following theorem characterizes the asymptotic performance of the robust hole detector in the \( \Psi(\gamma_n) \) class of active users. It also establishes the tradeoffs among the spectrum vacancy likelihood \( \epsilon_n \), per channel sampling budget \( M \), and the minimum power of active users.

Theorem 1 (Nonadaptive Tradeoff): For the \( \Psi(\gamma_n) \) class of active users, when \( \epsilon_n = o(1) \) and \( n_T = \omega(1) \), the error probability of the robust detector for identifying \( T \) spectrum holes is given by

\[
P_{NA}(n) = \max_{\{ H_i \} \subseteq \Psi(\gamma_n)} P\left( \hat{H}_n \cap H_i \neq \emptyset \right)
\]

\[
= 1 - \max_{\{ H_i \} \subseteq \Psi(\gamma_n)} P\left( \hat{H}_n \cap H_i = \emptyset \right)
\]

\[
= 1 - \max_{\{ H_i \} \subseteq \Psi(\gamma_n)} P\left( \{ Y_{u1}, \ldots, Y_{uT} \} \subseteq H_0 \right)
\]

Note that having all the \( T \) smallest measured energies belonging to the holes is equivalent to not having the energies of all nonholes greater than the \( T \) smallest measured energies, i.e.,

\[
\{ Y_{u1}, \ldots, Y_{uT} \} \subseteq H_0 \equiv \{ Y_{u1} \leq Y_i \forall i \in \{ 1, \ldots, T \} \text{ and } \forall i \in H_1 \}
\]

\[
\equiv \left\{ \max_{t \in \{ 1, \ldots, T \}} Y_{u_t} \leq \min_{i \in H_1} Y_i \right\}
\]

\[
\equiv \{ Y_{uv} \leq \min_{v \in H_1} Y_v \}
\]

Therefore

\[
P_{NA}(n) = 1 - \max_{\{ H_i \} \subseteq \Psi(\gamma_n)} P\left( Y_{uv} \leq \min_{i \in H_1} Y_i \right)
\]

Assessing \( P_{NA}(n) \) as defined previously relies on the properties of the order statistics of a set of random variables. The following lemma is instrumental for characterizing the distributions of the order statistics and evaluating \( P_{NA}(n) \). This is a generalization of a well-studied problem in the context of extreme value theory that considers the first order statistic \([24]\) for Gamma distributions. In this lemma, we give the corresponding results when distribution evolution is allowed, i.e., the number

3The random variable \( G \) with distribution Gamma \( \{ a, b \} \), where \( a, b > 0 \), has a density with respect to the Lebesgue measure of the form \( f_G(t) = \frac{e^{-t} a^{a-1} b^a}{\Gamma(a)} \) for \( t \geq 0 \).
\[ g_{1}(w; n_1) \overset{d}{=} \frac{\partial}{\partial \nu} Q_{1}(w; n_1) \] (we drop the explicit dependence on \( n_1 \) to avoid notational clutter). Note that \( \{W^n_{i,0}\} \) and \( \{W^n_{i,1}\} \) are statistically independent. In what follows, we are going to be interested in the order statistics of these sequences, in particular \( W^n_{(1),i} : i \in \{1, \ldots, T\} \), the smallest \( T \) elements of \( \{W^n_{i,0}\}_{i \in \mathbb{N}_0} \) sorted in increasing magnitude, and \( W^n_{(1),1} = \min_{i \in \mathbb{N}_0} W^n_{i,1} \). Taking this into account and noting that, by the law of large numbers, the assumptions \( \epsilon_n = o(1) \) and \( n \epsilon_n = o(1) \) imply that \( n_0, n_1 \to \infty \) as \( n \to \infty \), we have according to Lemma 1

\[
Q_{(1);1}(w; n_1) = 1 - \exp(-w^M) 
Q_{(1);0}(w; n_0) = P(W^n_{(i),0} < w) 
= 1 - \exp(-w^M) \sum_{k=0}^{\infty} \frac{w^k}{k!}.
\]

As a result, by invoking (11) when \( n_0 \) and \( n_1 \) approach infinity

\[
\min_{\{\nu_i\} \subseteq \nu_{(\gamma_n)}} P \left( \hat{U}_{\text{rob}}^{\text{NA}} \cap \mathcal{H}_1 \neq \emptyset \mid \{Z_i\} \right)
\]

is characterized in (16), shown at the bottom of the page.

We now remove the conditioning on \( \{Z_i\} \), by computing the expectation of the expression above. Recall that \( n_0 / n_1 \) is a random quantity, and by assumption and the law of large numbers \( n_0 / n_1 \to \epsilon_n \). Using this fact and continuous mapping principles, we conclude that

\[
\min_{\{\nu_i\} \subseteq \nu_{(\gamma_n)}} P \left( \hat{U}_{\text{rob}}^{\text{NA}} \cap \mathcal{H}_1 \neq \emptyset \mid \{Z_i\} \right) 
\]

where \( \mathcal{H}_1 \) is characterized in (16), shown at the bottom of the page.

As expected, there exists a tension between reliability and agility. On one hand, increasing the sampling budget per
channel $M$ favors reliability, as according to (12) it improves the probability of successfully detecting a hole. Increasing the sampling budget, on the other hand, imposes more delay in spectrum sensing as the time required for completing the acquisition of data scales linearly with the sampling budget. Finally, (12) also states that an increase in the minimum power of the active users $\gamma_n$, improves the reliability. Theorem 1 characterizes the tradeoff among reliability, hole detection agility, and the power of active users. By using the result of Theorem 1, we offer a necessary and sufficient condition on the scaling of the power of the active users to guarantee asymptotically error-free multiple hole detection in the nonadaptive sensing setting.

**Corollary 1 (Nonadaptive Power Scaling):** For the $\Psi(\gamma_n)$ class of active users, when $\epsilon_n = o(1)$ and $n\epsilon_n = \omega(1)$, the necessary and sufficient condition for $P_{NA}(n) \to 0$ as $n \to \infty$ is that $\gamma_n$ scales with increasing $n$ as

$$\gamma_n = \omega \left( \frac{M}{\sqrt{\epsilon_n}} \right).$$

In other words, if the power of the “faintest” active user grows faster than $\sqrt{\epsilon_n}$, then a secondary user can reliably identify $T$ holes by employing this nonadaptive procedure. The proof follows in a straightforward way from the characterization of $P_{NA}(n)$ given in (12) by noting that $P_{NA}(n) \to 0$ is equivalent to $(1 + \gamma_n)^M \cdot \epsilon_n \to \infty$.

**IV. ADAPTIVE SPECTRUM DETECTION**

**A. Adaptive Sensing Procedure**

Our proposed adaptive spectrum sensing procedure has two phases, namely the exploration phase and the detection phase. The exploration phase, being an iterative procedure, is intended to purify the set of the channels to be sensed carefully for detecting spectrum holes. This phase is accomplished by successively identifying and eliminating a group of channels deemed to be occupied. The detection phase is performed after the exploration phase in order to identify $T$ holes among the subset of candidate channel retained after exploration. The detection scheme deployed is identical to the robust spectrum detection scheme of Section III.

The exploration phase proceeds in an iterative way. In each iteration, it further monitors the channels retained by the previous iteration and eliminates those deemed to be spectrum holes least likely. The core idea is that it is relatively easy to identify occupied channels with low-quality measurements since there are few holes available (recall that $\epsilon_n$ is small). Each iteration carries on by thresholding the observed energy on each channel retained by the previous iteration. The threshold level depends only on $\gamma_n$, and is designed such that at each iteration roughly half of the existing occupied channels are eliminated, while almost all the of spectrum holes are preserved. The output of each exploration phase will have a more condensed proportion of spectrum holes to occupied channels. Subsequently, the detector developed for the nonadaptive procedure is applied on this refined set of channels in order to identify $T$ spectrum holes. This entire procedure bears similarities with Distilled Sensing [1]–[3]; however, the analysis is substantially different. This is due to the different sensing objective (identifying any arbitrary number of holes as opposed to [1]–[3] that aims to identify almost all) as well as the underlying statistical model.

We show that the gains yielded by this adaptive procedure can be interpreted in two ways. First, we demonstrate that when targeting the same level of reliability in hole detection, the adaptive procedure requires substantially less sampling budget, or equivalently, it is substantially more agile. Second, we show that under the same sampling budget, and targeting identical hole detection reliability, the adaptive procedure imposes less-stringent conditions on how fast the power of the active users $\gamma_n$ must scale with increasing $n$. This essentially indicates that for some choices of $\gamma_n$, the adaptive procedure can guarantee successful hole detection while the best nonadaptive procedure fails to do so.

Let us define $K$ as the number of exploration cycles (iterations) in the exploration phase. Also denote the sampling budget per channel in the $k$th exploration cycle by $M_k$. The exploration phase is initialized by including all channels for sensing and resumes as follows. In the first iteration, all channels are allocated the identical sampling budget of $M_1$. The energy levels of all channels are compared against $\lambda_1(1 + \gamma_n)$, where $\lambda_1$ is the median of the distribution $\Gamma M n a (M_1, 1)$. The channels for which the measured energy exceeds this threshold are discarded and the rest are carried over to the second iteration for further sensing. The same procedure is repeated throughout all $K$ cycles. More specifically, in the $k$th cycle, all the channels retained by the $(k-1)$th iteration are allocated the identical sampling budget of $M_k$. The energy levels of these channels are compared against $\lambda_k(1 + \gamma_n)$, where $\lambda_k$ is the median of the distribution $\Gamma M n a (M_k, 1)$ and the exploration is performed via thresholding as in the first iteration. Finally, after the exploration phase, each of the remaining channels is allocated the sampling budget of $M_{K+1}$ and the robust spectrum detection scheme provided in Remark 2 is applied in order to detect $T$ holes.

We set $G_0 = \{1, \ldots, n\}$ and for $k = 1, \ldots, K$, we define $G_k$ as the set of the indices of the channels that are retained by the $k$th exploration cycle. Clearly, we have $G_K \subset \ldots \subset G_1 \subset G_0$ and $G_K$ contains the set of the indices of the candidate channels among which $T$ holes will be detected. The set of measurements defined for the nonadaptive scheme in (5) is extended for the proposed adaptive procedure as follows. We define the set of measurements as

$$\mathcal{D}^K_{\mathbf{r}_{k-1}} \{ \mathbf{X}^k_i : \mathbf{X}^k_i = \sqrt{p_i} Z_i H_i^k \circ S_i^k + W_i^k \quad \text{for} \quad i \in G_{k-1} \} \subseteq \mathcal{D}^K_{n+1}$$

for $k = 1, \ldots, K + 1$. The measurement sets $\mathcal{D}^K_1, \ldots, \mathcal{D}^K_K$ are processed in the exploration phase and the measurement set $\mathcal{D}^{K+1}$ is used in the detection phase. Based on the model of channels dynamics described in Section II-A, only the spectral occupancy $\{Z_i\}$ and the power of active users $\{p_i\}$ remain unchanged during all observations, and fading, channel noise and transmitted signal change to independent states after each observation. Therefore, given the occupancy status $Z_i$, the observation sample $X_i^k(j)$ for $k = 1, \ldots, K + 1$ is distributed as

$$X_i^k(j) \sim N_{\mathbb{C}} \left[ 0, 1 + p_i Z_i \right]$$

for $i \in G_{k-1}$.
for \( i \in \mathcal{G}_{k-1} \) and \( j = 1, \ldots, M \). We also define
\[
Y^k_i = \left\| X^k_i \right\|^2 \quad \text{for} \quad i \in \mathcal{G}_{k-1} \quad \text{and} \quad k = 1, \ldots, K + 1.
\] (19)
Equation (18) and (19) provide that for a given \( Z_i \), \( Y^k_i \) is distributed as
\[
Y^k_i \mid Z_i \sim \text{Gamma}(M_k, 1 + p(Z_i)),
\] (20)
for \( i \in \mathcal{G}_{k-1} \) and \( k = 1, \ldots, K + 1 \). For each \( k = 1, \ldots, K + 1 \), corresponding to the sequence \( \{Y^k_i\}_{i \in \mathcal{G}_{k-1}} \), we define the sequence of order statistics \( \{Y^k_{(i)}\}_{i \in \mathcal{G}_{k-1}} \) in an increasing order such that \( Y^k_{(i)} \) represents the \( i \)-th smallest element of this sequence. The adaptive sensing procedure is formally described in Table I.

### Table I

**Adaptive Robust Spectrum Detection Algorithm**

| Exploration phase |
|-------------------|
| 1. Input \( K \in \mathbb{N} \) and \( \{M_1, \ldots, M_{K+1}\} \) where \( M_k \in \{1, 2, 3, \ldots\} \). |
| 2. Initialize the index set \( \mathcal{G}_0 = \{1, \ldots, n\} \). |
| 3. for \( k = 1, \ldots, K \) do |
| 4. Set \( Y^k_i = \left\| X^k_i \right\|^2 \quad \text{for} \quad i \in \mathcal{G}_{k-1} \). |
| 5. Obtain \( \mathcal{G}_k = \{ i \in \mathcal{G}_{k-1} \mid Y^k_i < \lambda_k (1 + \gamma_n) \} \) where \( \lambda_k \) is the median of \( \text{Gamma}(M_k, 1) \). |
| 6. end for |

| Detection phase |
|-----------------|
| 7. Set \( Y^{k+1}_i = \left\| X^{k+1}_i \right\|^2 \quad \text{for} \quad i \in \mathcal{G}_K \). |
| 8. Identify the indices of a spectrum holes \( \hat{U}^*_{\alpha} \triangleq \{ i \in \mathcal{G}_K : Y^{k+1}_i \leq Y^{k+1}_i (T) \} \). |
| 9. Output \( \hat{U}^*_{\alpha} \). |

Lemma 2: Let \( \ell_0 = |\mathcal{H}_1| \) and for \( k = 1, \ldots, K \) define \( \ell_k \) as the number of the occupied channels retained by the \( k \)-th exploration cycle. Conditionally on \( \ell_{k-1} \) for \( k = 1, \ldots, K \) and for sufficiently large \( n \) the event
\[
\left( c_k - \frac{1}{\log n} \right) \ell_{k-1} \leq \ell_k \leq \left( c_k + \frac{1}{\log n} \right) \ell_{k-1}
\] (22)
holds with probability at least \( 1 - 2 \exp \left( -\frac{2\gamma_n c_k}{\log n} \right) \), where \( c_k \leq \frac{1}{2} \) is a constant. Furthermore, \( c_k \rightarrow 1/2 \) if and only if \( p_t = \gamma_n \) for \( t \in \mathcal{H}_1 \).

Proof: See Appendix C.

A careful use of the above lemmas establishes the performance of the adaptive robust spectrum detection in the following theorem.

Theorem 2 (Adaptive Tradeoff): For the \( \Psi(\gamma_n) \) class of active users, when \( \gamma_n = o(1) \) and \( n\gamma_n = o(1) \), the error probability of the adaptive robust spectrum detection procedure for identifying \( T \) spectrum holes is given by
\[
P_{\Delta}(n) \triangleq \min_{\{T, \ell_t \leq |\mathcal{H}_1|\}} \left\{ P \left( \hat{U}_{\ell_t} \cap \mathcal{H}_1 \neq \emptyset \right) \right\} \leq 1 - \left( 1 + \left[ \left( 1 + \gamma_n \right)^{M_{K+1}} - K \gamma_n \right]^{-1} \right)^{-T}.
\] (23)

Proof: See Appendix D.

The tradeoff above suggests that the asymptotic performance of the adaptive procedure does not impose any constraint on the choice of \( M_k \) for \( k = 1, \ldots, K \), other than the trivial constraint \( M_K > 1 \). Hence, for achieving the best asymptotic performance the best sequential experimental design requires maximizing \( M_{K+1} \), i.e., the sampling budget per channel in the detection phase. More specifically, given a fixed sampling budget, i.e., a “cap” on the maximum number of measurements available, the best asymptotic strategy is to allocate as much as possible sampling budget for the detection phase and to allocate as low as one sample per channel in each exploration cycle. This implies that as low as one sample per channel in each exploration cycle suffices to ensure that for sufficiently large \( n \), the exploration phase retains almost all the holes and discards almost half of
the occupied channels. Therefore, to achieve the best asymptotic behavior, we should set $M_k = 1$ for $k = 1, \ldots, K$.

In order to quantify the gains yielded by the adaptive algorithm, we compare the results for the nonadaptive and adaptive schemes provided in Theorems 1 and 2. In particular, with the aim of attaining identical asymptotic reliability levels in the adaptive and the nonadaptive procedures, i.e., $P_A(n) \equiv P_{NA}(n)$, we characterize the agility gain, which we define as the ratio of the sampling budgets required by the nonadaptive procedure to that required by the adaptive scheme, i.e.,

$$\text{agility gain} = \frac{\Delta \text{ sampling budget of the non-adaptive procedure}}{\text{ sampling budget of the adaptive procedure}}.$$

**Theorem 3 (Agility):** For the $\Psi(\gamma_n)$ class of active users, when $\epsilon_n = o(1)$ and $n\epsilon_n = \omega(1)$, the agility gain of the adaptive robust spectrum detection algorithm with $M_n$ sampling budget is asymptotically lower bounded by $(2^{2K} + \frac{2}{M})^{-1}$, where $K$ is the number of exploration cycles.

**Proof:** See Appendix E.

It is noteworthy that while the number of exploration cycles $K$ can be made arbitrarily large (but fixed as a function of $n$), increasing it beyond some point will affect the agility very insignificantly. More specifically, for large $K$, the agility gain lower bound will be dominated by the term $\frac{2}{M}$. This underlines the fundamental limit of the agility gain yielded by the adaptive procedure.

An analogue of Corollary 1 can be derived for the adaptive procedure, providing a necessary and sufficient condition on the scaling of $\gamma_n$ for guaranteeing a reliable hole detection with the proposed algorithm. For comparison purposes, we assume that both adaptive and nonadaptive procedures are granted the same sampling budget.

**Corollary 2 (Adaptive Power Scaling):** For the $\Psi(\gamma_n)$ class of active users, when $\epsilon_n = o(1)$ and $n\epsilon_n = \omega(1)$, given that the sampling budget is $\tilde{M}_n$, a necessary and sufficient condition for $P_A(n) \to 0$ as $n \to \infty$ is that

$$\gamma_n = \omega\left(\frac{M'}{2K\epsilon_n}\right).$$

where $M' \geq 2^K(M - 2) + 2$.

**Proof:** See Appendix F.

Comparing the result above with that of Corollary 1 shows that an adaptive scheme can cope with much weaker active users than a nonadaptive scheme. More specifically, by noting that $M'$ is substantially larger than $M$, the power scaling requirement in the adaptive scenario, which is smaller than $\omega\left(\frac{M}{\sqrt{T}}\right)$ becomes substantially smaller than its counterpart in the nonadaptive scenario $\omega\left(\frac{1}{\epsilon_n}\right)$. As a result, there are scenarios where nonadaptive schemes fail to successfully identify $T$ holes, while the adaptive scheme succeeds.

**C. Optimal Exploration Cycles ($K$)**

The role of the exploration cycles is to feed the detector with a group of channels with a more condensed proportion of spectrum holes to occupied channels. According to the characterization of $P_A(n)$ given in (23), changing $K$ has two effects on the asymptotic behavior of $P_A(n)$. Specifically, increasing $K$ results in a decrease in $P_A(n)$ through its direct impact via the term $2^K$. Increasing the exploration cycles $K$, on the other hand, increases the sampling resources during exploration and leaves less resources for the detection phase, i.e., $M_{K+1}$ decreases. Following the same arguments as in the proof of Theorem 3, it can be readily verified that the superposition of the two effects implies that $P_A(n)$ will be monotonically decreasing with increasing $K$. The results of Theorem 3 and Corollary 2 also substantiate that increasing $K$ enforces a less stringent constraint on power scaling and enhances the agility.

An important feature in the analysis of the proposed algorithm was that the exploration step asymptotically retains all the existing holes, while discarding a large number of occupied channels. This ensures the task of the detection phase is much more effective. As all the results presented so far hint that a large value of $K$ yields higher gains it is natural to ask what is the “optimal” value of $K$ so that the exploration phase asymptotically retains all the holes. This imposes an upper bound on possible values of $K$, and gives rise to the following theorem.

**Theorem 4:** The optimal growth of the number of cycles $K$ so that the exploration phase asymptotically retains all the holes for consideration in the detection phase is given by

$$K^* = \log\log\frac{1}{\epsilon_n}.\quad (25)$$

Furthermore, the outcome of Theorem 2 remains valid with any choice of $K$ not exceeding $K^*$, which is no longer fixed as a function of $n$.

**Proof:** See Appendix G.

**V. SIMULATION RESULTS**

In this section, we provide some simulation results to empirically demonstrate the effectiveness of the proposed adaptive procedure. In Fig. 1, we aim at identifying $T = 2$ spectrum holes and compare the reliability of the nonadaptive scheme with that of the proposed adaptive procedure with $K = 1, \ldots, 4$ exploration cycles. Comparisons are provided over the range of $n = 10^5$ to $10^{100}$ channels. The sampling budget in both schemes is set as $5n$ which means that in the nonadaptive scheme each channel is measured $M = 5$ times. In the adaptive scheme in each cycle of the exploration phase, each channel is measured once. The sampling resources not used in the exploration phase are equally divided among the remaining channels retained by the exploration phase. We set the channel occupancy probability $\epsilon_n = n^{-2/3}$, which clearly satisfies the conditions $\epsilon_n = o(1)$ and $n\epsilon_n = \omega(1)$, and finally assume that the power of active users are $p_i = \gamma_n = n^{1/5}$, $\forall i \in R_1$. As predicted by the analysis, for large values of $n$ the adaptive procedure yields a significant improvement over the nonadaptive scheme, e.g., for $n = 100$, we gain two orders of magnitude in error probability after four cycles of exploration. The improvement attained for small values of $n$ is also considerable, e.g., for $n = 20$, it is one order of magnitude. It is noteworthy that the choices of $\epsilon_n = n^{-2/3}$ and $p_i = \gamma_n = n^{1/5}$ have been arbitrary and ex-
Fig. 1. Comparison of the error probabilities $P_s(n)$ and $P_{s,A}(n)$ for detecting $T = 2$ over a wide range of $n$.

Fig. 2. Average number of observations taken per channel versus the number of channels $n$. All the curves correspond to the target error probability for detecting $T = 2$ holes is $P_s(n) = P_{s,A}(n) = 10^{-4}$.

tensive simulations show that the gains are not very sensitive to these choices.

Next, we investigate the agility gain. Define the normalized sampling budget as the aggregate sampling budget divided by the number of channels. In Fig. 2, we plot this quantity against the number of channels $n$, when requiring both approaches to have the same error probability for detecting $T = 2$ holes. For the adaptive scheme, we consider the performance with $K = 1, 4, 8, 16$, and 32 cycles of exploration. Again we consider the choices of $r_n = n^{-2/3}$ and $p_i = \gamma_n = n^{1/3}$. It is seen that the agility is improved by increasing $K$, and for $K = 5$ the adaptive procedure requires about 80% less sampling budget than the nonadaptive procedure. In other words, the adaptive procedure is 5 times faster than the nonadaptive scheme for detecting a hole with error probability is $P_{s,A} = 10^{-4}$.

Fig. 3. Detectability regions for $M = 5$, $K = 4$, $r_n = n^{\alpha-1}$, and $\gamma_n = n^\beta$.

Fig. 4. Average number of observations taken per channel versus sparsity of the holes in the adaptive and the sequential method of [4] for achieving $P_s(n) = 10^{-4}$.

adaptive procedures are $\beta \geq \frac{1-\alpha}{M}$ and $\beta \geq \frac{1-\alpha}{M}$, respectively, where $M^2 \geq 2K(M - 2) + 2$. The dashed lines depict these theoretical boundaries and split the space of $(\alpha, \beta)$ into three separate regions. We also provide the simulation results for some pairs of $(\alpha, \beta)$ in each of these regions. It is expected that in the lower region, both procedures fail; in the middle region, only the adaptive procedure succeeds and in the upper region both succeed. The simulation results (except for a few points) support the theoretical asymptotic predictions. In the simulations, we have set $M = 5$, $K = 4$, and $n = 1000$. We remark that requiring as low as about five sensing samples per channel is substantially less than the compressed-sensing-based approaches (see, e.g., [13]), that for obtaining a reliable estimate of the PSD take samples at a rate equal to at least half Nyquist rate, which for a wideband channel is large.

In Fig. 4, we compare the agility of our proposed adaptive procedure with that of the quickest search [4] for achieving the target error probability $P_s(n) = 10^{-4}$. We compare the two schemes at different sparsity levels for the distribution of the spectrum holes. Specifically, we consider $r_n = n^{\alpha-1}$ for $\alpha \in [0, 1]$ where smaller values of $\alpha$ correspond to sparser distribution of the holes. The objective is to identify $T = 2, 3$ holes out of $n = 1000$ channels. The powers for the active users are assumed to be $\gamma = n^\beta$ for $\beta = 1/5, 1/10$, and we deploy
Fig. 5. Illustration of the exploration phase in eliminating the nonholes for $n = 1000$ in three different system settings.

Now recall that for a Gamma $\{M, \alpha_m\}$ distribution

$$F(y; m) = \frac{1}{\Gamma(M)\alpha_m} \int_0^y \left(\frac{x}{\alpha_m}\right)^{M-1} \exp\left(-\frac{x}{\alpha_m}\right) dx.$$  

Noting that for $0 \leq x \leq y$ we have $\exp(-y/\alpha_m) \leq \exp(-x/\alpha_m) \leq 1$ we immediately get

$$\frac{\Gamma(M+1)}{\Gamma(M+1)} \left(\frac{y}{\alpha_m}\right)^M < F(y; m) \leq \frac{1}{\Gamma(M+1)} \left(\frac{y}{\alpha_m}\right)^M.$$  

By noting that we had defined $b_m \triangleq \alpha_m \left[\frac{\Gamma(M+1)}{m}\right]^\frac{1}{M}$, after some simplifications we find the following bounds above for $F(w_b; m)$:

$$\frac{w^M}{m} \exp\left(-w^M \cdot \frac{\Gamma(M+1)}{m}\right) \leq F(w_b; m) \leq \frac{w^M}{m}.$$  

On the other hand for every fixed $k$, we get

$$\left(\frac{m}{k}\right) \left[\frac{F(w_b; m)}{1 - F(w_b; m)}\right]^k \leq \prod_{i=0}^{k-1} \left[\frac{m}{k}\right] \left[\frac{m}{k}\right] \left[\frac{w^M}{m}\right] \left[\frac{1 - o(1)}{1 - \frac{w^M}{m}}\right]^k.$$  

Equations (26)–(28) yield

$$Q(i; w; m) \leq 1 - \exp(-w^M) \sum_{k=0}^{i-1} \frac{w^{KM}}{k!}.$$  

VI. CONCLUSION

In this paper, we have presented an adaptive sensing methodology for identifying multiple spectral holes in a wideband cognitive radio system. By gradually adjusting the measurement process using information gleaned from the previous measurements, we are able to significantly improve the probability of correctly detecting multiple spectrum holes. This dramatic gain is patent both in the theoretical analysis and in the simulation results. More importantly, in the cognitive radio setting, this improvement translates into a significant increase in agility, allowing spectrum holes to be identified quickly and therefore widening the window of time for opportunistic data transmission. Such dramatic improvements are not possible without the use of adaptive sampling techniques, as demonstrated.

APPENDIX A

PROOF OF LEMMA 1

Let $F(t; m) \triangleq P(Y \leq t)$ denote the CDF of the Gamma $\{M, \alpha_m\}$ random variables $\{Y_i\}_{i=1}^m$. Also let $Q(i; w; m) \triangleq P(W^{(i)} \leq w)$ denote the CDF of $W^{(i)}$. Therefore

$$Q(i; w; m) = P(W^{(i)} \leq w) = P(Y^{(i)} \leq w) = 1 - \left[1 - F(w_b; m)\right]^m = 1 - \left[1 - F(w_b; m)\right]^m \sum_{k=0}^{i-1} \left[\frac{m}{k}\right] \left[\frac{F(w_b; m)}{1 - F(w_b; m)}\right]^k.$$  

(26)
which is the desired result.

APPENDIX B
PROOF OF LEMMA 2

For the nonadaptive procedure by invoking the assumption \( \epsilon_n = o(1) \), we know that \( \gamma_n = \omega(1) \) from Corollary 1. Throughout the analysis for the adaptive procedure we also assume that \( \gamma_n = \omega(1) \) and characterize the precise necessary and sufficient condition for the scaling rate of \( \gamma_n \) at the end of Section IV. We start with the following remarks.

Remark 3: Let \( X \sim \text{Gamma}(M,a) \). Then for any \( \lambda > 0 \) we have \( P(X < \lambda) \geq 1 - \exp\left(-\frac{\lambda}{a}\right) \).

Remark 4: For \( B \) distributed as \( \text{Binomial}(m,a) \) and \( b < E[B] \), we have
\[
P(B \leq b) \leq \left(\frac{m - ma}{m - b}\right)^b \left(\frac{ma}{b}\right)^b.
\]

Let us define \( \mathcal{J}_0 = \mathcal{G}_0 \cap \mathcal{H}_0 \) and for \( k = 1, \ldots, K \) let \( \mathcal{J}_k = \mathcal{G}_k \cap \mathcal{H}_k \) be the set containing the indices of the spectrum holes retained by the \( k \)th exploration cycle. By the definition of \( m_k \) we have \( |\mathcal{J}_k| = m_k \). Also, for any \( i \in \mathcal{G}_{k-1} \) define
\[
Y_i^{k} \triangleq \{Y_i < \lambda_k(1 + \gamma_n)\}
\]
where
\[
1\{A\} = \begin{cases} 1, & \text{if } A \text{ is true} \\ 0, & \text{if } A \text{ is false} \end{cases} \quad \text{for } k = 1, \ldots, K.
\]

According to the algorithm in Table I for the \( i \)th channel, which can be either a hole or a nonhole, \( T_i^{k} \) is 1 if \( i \) is retained by the \( k \)th exploration cycle and is 0 otherwise. Therefore, the number of holes retained by the \( k \)th exploration cycle is \( m_k = \sum_{i \in \mathcal{J}_{k-1}} T_i^{k} \). By noting that \( |\mathcal{J}_{k-1}| = m_{k-1} \), we find that
\[
m_k \sim \text{Binomial}(m_{k-1}, a_k), \quad \text{for } k = 1, \ldots, K
\]
where
\[
a_k \Delta P(T_i^{k} = 1 \mid i \in \mathcal{J}_{k-1})
\]

\[
P(m_k \leq b) = P \left( m_k \leq \frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1} \right)
\]
\[
\leq \left[ \left( 1 - a_k(1 + \gamma_n) \right) \frac{\lambda_k(1 + \gamma_n)}{\gamma_n} \right]^{m_{k-1}} \left( \frac{\lambda_k(1 + \gamma_n)}{\gamma_n} \right)^{m_{k-1}}
\]
\[
\leq \left[ \left( 1 + \gamma_n \right) \exp\left(-\lambda_k(1 + \gamma_n)\right) \right]^{\frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1}}
\]
\[
= \exp \left( \frac{1}{1 + \gamma_n} \left( \log(1 + \gamma_n) - \lambda_k(1 + \gamma_n) + \gamma_n \log \left( \frac{1}{1 + \gamma_n} \right) \right) \right) \leq 1
\]
\[
\leq \exp \left( - \frac{m_{k-1}}{\mu^n} \right)
\]

(30)

Thus, based on (20), \( Y_i^{k} \) is distributed as \( \text{Gamma}(M_k, 1) \) for \( i \in \mathcal{J}_{k-1} \). Therefore, by using Remark 3, we find that
\[
a_k \geq 1 - \exp\left(-\lambda_k(1 + \gamma_n)\right).
\]
\[
P \left( \frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1} \leq m_k \leq m_{k-1} \right)
\]
\[
= P \left( m_k \leq \frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1} \right)
\]
\[
\geq 1 - \exp\left(-\frac{m_{k-1}}{\mu^n} \right)
\]

(29)

which concludes the result.

For sufficiently large \( n \) (and thereof sufficiently large \( \gamma_n \)) we have \( \frac{1}{1 + \gamma_n} \geq \exp\left(-\lambda_k(1 + \gamma_n)\right) \) which in conjunction with (29) provides that for sufficiently large \( n \)
\[
\frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1} \leq a_k m_{k-1}.
\]

Therefore, by setting \( b \triangleq \frac{\gamma_n}{1 + \gamma_n} \cdot m_{k-1} \), for the random variable \( m_k \sim \text{Binomial}(m_{k-1}, a_k) \) we have \( b \leq E[m_k] = a_k m_{k-1} \). As a result, for this choice of \( b \) the condition of Remark 4 is satisfied and, for sufficiently large \( n \) we find that \( P(m_k \leq b) \) is lower bounded according to (32). Note that (30) [see (30–32), shown at the bottom of the page] holds according to Remark 4. Equation (31) is obtained by invoking the inequality given in (29), and noting that \( a_k \leq 1 \) as \( a_k \) is a probability. Finally, the inequality in (32) is justified by noting that for sufficiently large \( n \) and any \( \alpha > 0 \)
\[
\frac{\log(1 + \gamma_n)}{1 + \gamma_n} - \lambda_k + \frac{1}{1 + \gamma_n} \leq -\frac{1}{n^\alpha}
\]
as all terms in the LHS and RHS except the constant \(-\lambda_k\) are diminishing in the asymptote of large \( n \). Recall that \( \mathcal{J}_k \subseteq \mathcal{J}_{k-1} \) and consequently, \( m_k \leq m_{k-1} \). This fact in conjunction with (32) provides that.
APPENDIX C
PROOF OF LEMMA 3

Let $L_0 = g_0 \cap H_1$ and for $k = 1, \ldots, K$ define $L_k^\Delta = g_k \cap H_1$ which contains the set of the indices of the nonholes contained retained by the $k$th exploration cycle. By the definition of $\ell_k$, we have and $|L_k| = \ell_k$. By recalling the definition of $T^k_i$ and taking into account their independence, the number of nonholes that the $k$th exploration cycle retains is $\ell_k = \sum_{i \in L_k^\Delta} T^k_i$. Applying Hoeffding’s inequality provides

$$P\left( \ell_k - E[\ell_k] > \frac{\ell_{k-1}}{\log n} \right) \leq 2 \exp\left( -\frac{2\ell_{k-1}}{(\log n)^2} \right). \tag{33}$$

On the other hand

$$E[\ell_k] = \sum_{i \in L_k^\Delta} E[T^k_i] = \sum_{i \in L_k^\Delta} P(T^k_i = 1) = 1 \in L_k^\Delta). \tag{34}$$

According to (20), for $i \in L_k^\Delta$, $Y^k_i$ is distributed as Gamma $(M_k, 1 + p_i)$. For any $i \in L_k^\Delta$, we denote the CDF of Gamma $(M_k, 1 + p_i)$ by $F^k_i(x; \theta)$. Since $p_i > \gamma_n$, we get

$$P(T^k_i = 1 | i \in L_k^\Delta) = F^k_i\left( \lambda_k(1 + \gamma_n) \right) \leq F^k_i\left( \lambda_k(1 + p_i) \right) = \frac{1}{\Gamma(M_k)(1 + p_i)^M} \int_0^{\lambda_k(1 + p_i)} x^{M-1} \exp\left( -\frac{x}{1 + p_i} \right) dx$$

$$= \frac{1}{\Gamma(M_k)} \int_0^{\lambda_k} x^{M-1} \exp\left( -\frac{x}{2} \right) dx = \frac{1}{2}$$

where the last step holds as we had defined $\lambda_k$ as the median of Gamma $(M_k, 1)$. As a result, from (34), we find that $E[\ell_k] \leq \ell_{k-1} = \frac{\ell_{k-1}}{2}$. By selecting $c_k := \frac{E[\ell_k]}{\ell_{k-1}} < \frac{1}{2}$ and recalling (33), we get the desired result. Finally, note that $F^k_i(\lambda_k(1 + \gamma_n))$ is monotonic in $p_i$. Therefore, the necessary and sufficient condition for having $E[\ell_k] = \ell_{k-1} = \frac{\ell_{k-1}}{2}$ is that $F^k_i(\lambda_k(1 + p_i)) = F^k_i(\lambda_k(1 + p_i))$, which in turn requires that $p_i = \gamma_n$ for $i \in H_1$.

APPENDIX D
PROOF OF THEOREM 2

Similar to the proof of Theorem 1, for quantifying the performance of the robust detection scheme, we consider the worst case scenario and set $p_i = \gamma_n$ for all $i \in H_1$. Define the event $A_k$ as in (37), shown at the bottom of the page.

Lemma 2 and 3 characterize the probability of event $A_k$ given $m_{k-1}$ and $\ell_{k-1}$, and state that

$$\zeta_k := P(A_k | m_{k-1}, \ell_{k-1}) \leq \chi\left( \left(1 - \frac{1}{1 + \gamma_n}\right)m_{k-1}, m_{k-1}, \ell_{k-1} \right)$$

$$P\left( \left(1 - \frac{1}{1 + \gamma_n}\right)\ell_{k-1} < \ell_k < \left(1 + \frac{1}{1 + \gamma_n}\right)\ell_{k-1} \right)$$

$$\geq 1 - \exp\left( -\frac{m_{k-1}}{n^\alpha} \right) - 2 \exp\left( -\frac{2\ell_{k-1}}{(\log n)^2} \right). \tag{35}$$

By concatenating all the bounds, we conclude that the event $(A_1, A_2, \ldots, A_K)$ is a sufficient condition for the event $A$ defined in (38), shown at the bottom of the page.

By noting that given $m_{k-1}$ and $\ell_{k-1}$, the event $A_k$ is independent of events $A_i$ for $i < k$, we immediately conclude that the event $A$ holds with probability at least

$$P(A) > P(A_1, A_2, \ldots, A_K)$$

$$= \prod_{i=1}^{K} P(A_k \text{ given } m_{k-1}, \ell_{k-1}) = \prod_{i=1}^{K} \zeta_i.$$ 

In the asymptote of large $n$, for any fixed $K$, we have (conditions on the choice of $K$ are discussed in Theorem 4)

$$\lim_{n \to \infty} \left(1 - \frac{1}{1 + \gamma_n}\right)^K = 1$$

and

$$\lim_{n \to \infty} \left(1 + \frac{1}{1 + \gamma_n}\right)^K = 1$$

$$\lim_{n \to \infty} \left(1 + \frac{1}{\log n}\right)^K = 1$$

which provide that in the asymptote of large $n$

$$A = \left\{ m_K = m_0 \quad \text{and} \quad \ell_K - \frac{\ell_0}{2K} \right\}$$

with probability at least $\prod_{i=1}^{K-1} \zeta_i$. On the other hand, by recalling the assumptions $c_0 = o(1)$ and $n_0 = \omega(1)$, according to the law of large numbers we get $\ell_0 = n(1 - c_0) = n$ and $m_0 = n c_0 = \omega(1)$. Hence, we can ensure that for $k -$
1, . . . , K, \zeta_k \rightarrow 1 as n \rightarrow \infty. More specifically, by selecting
\alpha = \lim_{n \rightarrow \infty} \frac{\log m_n}{\log n} for \zeta_1 we have
\lim_{n \rightarrow \infty} \frac{m_0}{n^\alpha} = \lim_{n \rightarrow \infty} \frac{n \epsilon_n}{n^\alpha} = +\infty
which shows that
\lim_{n \rightarrow \infty} \exp\left(-\frac{m_0}{n^\alpha}\right) = 0 \quad (39)
and
\lim_{n \rightarrow \infty} \frac{2\epsilon_n}{(\log n)^2} = \lim_{n \rightarrow \infty} \frac{2n}{(\log n)^2} = +\infty
which shows that
\lim_{n \rightarrow \infty} 2 \exp\left(-\frac{2\epsilon_{n-1}}{(\log n)^2}\right) = 0. \quad (40)
Equations (39)–(40) and (35) yield that \lim_{n \rightarrow \infty} \zeta_1 = 1. By following the same procedure, the same result for all \zeta_k can be recovered. As a result, for sufficiently large n, with probability 1 we have \( m_K = m_0 \) and \( \epsilon_K = \epsilon_0/2^K \).

Now let us define \( \tilde{n}_0 \) and \( \tilde{n}_1 \) as the number of holes and nonholes retained after the exploration phase. By following the same line of argument as in the proof of Theorem 1 and by appropriately applying Lemma 1, we can show that
\[ P_A(n) = 1 - \left(1 + \left[\frac{1}{\gamma_n}\right]^M \cdot \frac{\tilde{n}_0}{\tilde{n}_1} \epsilon_n\right)^{-T}. \quad (41) \]
Note that for sufficiently large n
\[ \frac{\tilde{n}_0}{\tilde{n}_1} \rightarrow \frac{m_K}{\epsilon_K} \rightarrow \frac{m_0}{\epsilon_0} \cdot 2^K = \frac{n_0}{n_1} \cdot 2^K = 2^K \epsilon_n. \]
Plugging this asymptotic value of \( \frac{\tilde{n}_0}{\tilde{n}_1} \) into (41) completes the proof.

APPENDIX E
PROOF OF THEOREM 3
Based on the structure of the experimental design, for a given \( M_n \) sampling budget the nonadaptive procedure takes M observations per channel. By equating the error probabilities of hole detection in the nonadaptive and adaptive schemes given in (12) and (23), we find that as \( n \rightarrow \infty \)
\[ (1 + \gamma_n)^M \epsilon_n = (1 + \gamma_n)^{M_K+1} \cdot 2^K \epsilon_n \]
which shows that
\[ M_{K+1} = M \cdot \frac{K}{\log(1 + \gamma_n)} \leq M. \quad (42) \]
Next, let us evaluate the sampling budget required by the nonadaptive procedure given that in the detection phase we take \( M_{K+1} \) samples per channel that has survived the exploration phase. As discussed earlier, the characterization of \( P_A(n) \) conveys that the optimal experimental design is to take 1 sample per channel in each exploration cycle. Therefore, the aggregate amount of samples taken throughout exploration and detection phases is \( \sum_{k=1}^K (m_k + \ell_k) \). On the other hand, from lemmas 2 and 3, we find that in the asymptote of large n we have \( m_k = m_0 \) and \( \ell_k = \frac{1}{2^k} \epsilon_1 \) with probability 1 for \( k = 1, \ldots, K \). Recalling that \( \frac{m_0}{\ell_0} \rightarrow \epsilon_n \rightarrow 0 \) and \( \ell_0 = n \), we find that the sampling budget required by the adaptive procedure is asymptotically equal to
\[ \sum_{k=0}^{K-1} \ell_k + \ell_K M_{K+1} = \sum_{k=0}^{K-1} \frac{n}{2^k} + \frac{n}{2^K} M_{K+1} = \frac{K-1}{n} \frac{n}{2^k} + \frac{n}{2^K} M_{K+1} \leq n \left(2 - \frac{1}{2^k} \frac{M_{K+1}}{2^K}\right). \quad (43) \]
Equations (42) and (43) together show that the sampling budget of the adaptive procedure is asymptotically upper bounded by \( n(2 + M/2^K) \). As a result, for maintaining \( P_A(n) = P_{NA}(n) \), the asymptotic ratio of the sampling budget required by the nonadaptive procedure to that required by the adaptive scheme is lower bounded by
\[ \frac{M_n}{n(2 + M/2^K)} = \left(\frac{2}{M} + \frac{1}{2^K}\right)^{-1}. \]

APPENDIX F
PROOF OF COROLLARY 2
First we find how \( M_K \) and \( M \) should be related such that they give rise to the same sampling budget required by the adaptive and nonadaptive schemes. By following a similar line of argument as in the proof of Theorem 3 and (43), we find that
\[ M_n = \sum_{k=0}^{K-1} \frac{m_k + \ell_k}{2^K} \leq n \left(2^K - \frac{1}{2^K} \frac{1}{M_{K+1}} + \frac{M_{K+1}}{2^K}\right) \]
which provides \( M_{K+1} > 2^K(M - 2^k) + 2 \). On the other hand, from (23) forcing \( P_A(n) \rightarrow 0 \) yields that it is necessary and sufficient to have
\[ \gamma_n = \omega \left(\frac{M_{K+1}}{2^k \epsilon_n}\right). \]
Setting \( M' = M_{K+1} \) establishes the desired result.

APPENDIX G
PROOF OF THEOREM 4
A key constraint that guarantees retaining almost all the holes throughout the exploration phase is cast in (36) as
\[ \lim_{n \rightarrow \infty} \left(1 - \frac{1}{1 + \gamma_n}\right)^K = 1. \quad (44) \]
According to Corollary 2, guaranteeing this equality necessitates that $K$ does not grow faster than $\gamma_n$, i.e.,

$$K \in o(\gamma_n) \quad \text{while} \quad \gamma_n = \omega\left(\sqrt[\nu]{\frac{1}{2K\epsilon_n}}\right). \quad (45)$$

Define

$$\Omega_n \triangleq \{\gamma_n \mid \gamma_n \text{ does not satisfy (24)}\}. \quad (46)$$

From (45)–(46), we obtain that

$$K = \sup_{\Omega_n} \gamma_n \geq \omega\left(\sqrt[\nu]{\frac{1}{2K\epsilon_n}}\right) \Rightarrow M' \log K + 2 \geq \log \frac{1}{\epsilon_n}. \quad (47)$$

By taking into account the characterization of $M'$ in Corollary 2, we find the following upper bound on $K$

$$(2K(M - 2) + 2) \log K + 2 \leq \log \frac{1}{\epsilon_n}. \quad (48)$$

Next, we identify the scaling law of $K$ with respect to $n$ such that the asymptotic inequality above is satisfied. As $K$ scales, the constant summands (i.e., +2) will be dominated by other summands, the inequality above can be stated as

$$\log(M - 2) + K + \log \log K < \log \log \frac{1}{\epsilon_n}. \quad (49)$$

$(M - 2)$ being a constant as well as $\log \log K$ are both dominated by $K$, giving rise to the following upper bound on $K$

$$K \leq \log \log \frac{1}{\epsilon_n}. \quad (50)$$

We also need to examine the constraints on $K$ that guarantee that the exploration phase discards the nonholes as expected in the proof of Theorem 2. More specifically, we need to ensure that the two following equalities hold:

$$\lim_{n \to \infty} \left(1 - \frac{1}{2 \log n}\right)^K = 1 \quad \text{and} \quad \lim_{n \to \infty} \left(1 + \frac{1}{\log n}\right)^K = 1 \quad (51)$$

which indicate that $K$ must satisfy

$$K \in o(\log n). \quad (52)$$

By recalling the assumption $\epsilon_n = \omega(1)$, or its equivalent form $\frac{1}{\epsilon_n} = o(n)$, it can be readily verified that the bound given in (50) is stronger than that in (52). By taking into account the monotonicity of the detection performance with increasing $K$, the optimal choice of $K$ is the upper bound in (50).

**Appendix H**

**Proof of Remark 1**

The probability that a fixed set of channels $\mathcal{U}$ with cardinality $T$ contains only holes is

$$P(\mathcal{U} \subset \mathcal{H}_0 \mid \mathcal{D}_n) = \prod_{i \in \mathcal{U}} P(i \in \mathcal{H}_0 \mid \mathcal{D}_n)$$

$$= \prod_{i \in \mathcal{U}} P(i \in \mathcal{H}_0) \cdot P(D_n \mid i \in \mathcal{H}_0)$$

$$= \prod_{i \in \mathcal{U}} P(i \in \mathcal{H}_0) \cdot P(X_i \mid i \in \mathcal{H}_0)$$

$$= \prod_{i \in \mathcal{U}} \frac{1}{\sqrt{2\pi} \epsilon_n} \exp\left(-\frac{1}{2} |X_i|^2\right)$$

$$= \prod_{i \in \mathcal{U}} \frac{1 - \epsilon_n}{1 + p_i} \frac{1}{\left(1 + p_i\right)^M} \exp\left(-\frac{1}{1 + p_i} \|X_i\|^2\right). \quad (53)$$

where the above expansion follows from Bayes's rule and the statistical independence of the observations made over different channels. Hence, the MAP detector that maximizes $P(\mathcal{U} \subset \mathcal{H}_0 \mid \mathcal{D}_n)$ among all sets of channels with cardinality $T$ is given by

$$\hat{\mathcal{U}}_{MAP} = \arg \max_{\mathcal{U} \mid |\mathcal{U}| = T} \log P(\mathcal{U} \subset \mathcal{H}_0 \mid \mathcal{D}_n)$$

$$= \arg \min_{\mathcal{U} \mid |\mathcal{U}| = T} \sum_{i \in \mathcal{U}} \log \left[1 + \frac{1}{\epsilon_n} \cdot U_i\right]. \quad (54)$$

where

$$U_i = \frac{1}{\left(1 + p_i\right)^M} \exp\left(-\frac{p_i}{1 + p_i} \|X_i\|^2\right). \quad (55)$$

Therefore, $\hat{\mathcal{U}}_{MAP}$ contains the indices of the $T$ smallest elements of the set $\{U_1, \ldots, U_n\}$.

**References**

[1] J. Haupt, R. Castro, and R. Nowak, “Distilled sensing: Adaptive sampling for sparse detection and estimation,” *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6222–6235, Sep. 2011.

[2] J. Haupt, R. Castro, and R. Nowak, “Adaptive discovery of sparse signals in noise,” presented at the 42nd Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Oct. 2008.

[3] J. Haupt, R. Castro, and R. Nowak, “Distilled sensing: selective sampling for sparse signal recovery,” presented at the 12th Conf. Artif. Intell. Statist., Clearwater Beach, FL, Apr. 2009.

[4] L. Lai, H. V. Poor, Y. Xin, and G. Georgiadis, “Quickest search over multiple sequences,” *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5375–5386, Aug. 2011.

[5] J. A. Bazerque and G. B. Giannakis, “Distributed spectrum sensing for cognitive radio networks by exploiting sparsity,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1847–1862, Mar. 2010.

[6] J. Meng, W. Yin, H. Li, E. Hossain, and Z. Han, “Collaborative spectrum sensing from sparse observations in cognitive radio networks,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 327–337, Feb. 2011.

[7] S. Haykin, D. J. Thomson, and J. H. Reed, “Spectrum sensing for cognitive radio,” *Proc. IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[8] J. Ma, G. Y. Li, and B. H. Juang, “Signal processing in cognitive radio,” *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[9] T. Yucek and H. Arslan, “A survey of spectrum sensing algorithms for cognitive radio applications,” *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, First Quarter 2009.
[10] F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, “Next-generation dynamic spectrum access cognitive radio wireless networks: A survey,” Comput. Netw., vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[11] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, “Collaborative wideband sensing for cognitive radios,” IEEE Signal Process. Mag., vol. 25, no. 6, pp. 60–73, Nov. 2008.

[12] M. Malloy and R. Nowak, “Sequential analysis in high-dimensional multiple testing and sparse recovery,” presented at the IEEE Int. Symp. Inf. Theory, Aug. 2011.

[13] Z. Tian and G. B. Giannakis, “Compressed sensing for wideband cognitive radios,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Honolulu, HI, Apr. 2007, pp. 1357–1360.

[14] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, “Compressive wideband spectrum sensing,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Taipei, Taiwan, Apr. 2009, pp. 2337–2340.

[15] Y. Wang, Z. Tian, and C. Feng, “A two-step compressed spectrum sensing scheme for wideband cognitive radios,” presented at the presented at the IEEE Global Telecommun. Conf., Miami, FL, Dec. 2010.

[16] G. Vazquez-Vilar, R. Lopez-Valcarce, C. Mosquera, and N. Gonzalez-Prelcic, “Wideband spectral estimation from compressed measurements exploiting spectral a priori information in cognitive radio systems,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Dallas, TX, Apr. 2010, pp. 2958–2961.

[17] S. Hong, “Multi-resolution bayesian compressive sensing for cognitive radio primary user detection,” presented at the presented at the IEEE Global Telecommun. Conf., Miami, FL, Dec. 2010.

[18] Z. Yu, S. Hoyos, and B. M. Sadler, “Mixed-signal parallel compressed sensing and reception for cognitive radio,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Las Vegas, NV, Apr. 2008, pp. 3861–3864.

[19] Z. Tian, “Compressed wideband sensing in cooperative cognitive radio networks,” presented at the presented at the IEEE Global Telecommun. Conf., New Orleans, LA, Dec. 2008.

[20] D. Angelosante and G. B. Giannakis, “RLS-weighted lasso for adaptive estimation of sparse signals,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Taipei, Taiwan, Apr. 2009, pp. 3245–3248.

[21] D. Angelosante, G. B. Giannakis, and E. Grossi, “Compressed sensing of time-varying signals,” presented at the presented at the IEEE Int. Conf. Digital Signal Process., Santorini, Greece, Jul. 2009.

[22] Y. Wang, A. Pandharipande, Y. L. Poloy, and G. Leus, “Distributed compressive wide-band spectrum sensing,” presented at the presented at the Inf. Theory Appl. Workshop, San Diego, CA, Feb. 2009.

[23] F. Zeng, Z. Tian, and C. Li, “Distributed compressive wideband spectrum sensing in cooperative multi-hop cognitive networks,” presented at the presented at the IEEE Int. Conf. Commun., Cape Town, South Africa, Jul. 2010.

[24] H. A. David and H. N. Nagaraja, Order Statistics, 3rd ed. New York: Wiley, 2003.

Ali Tajer (S’05–M’10) received the B.Sc. and M.Sc. degrees in Electrical Engineering from Sharif University of Technology, and the M.A. degree in Statistics and the Ph.D. in Electrical Engineering from Columbia University. During 2010-2012 he was a postdoctoral research associate at Princeton University and an adjunct assistant professor at Columbia University. Since August 2012 he has been an assistant professor at Wayne State University. His research interests lie in the general areas of network information theory, wireless communications, and applied statistics. Dr. Tajer received the Jury Award for the Excellence of Research from Columbia University in 2011.

Rui M. Castro received a Licenciatura degree in aerospace engineering in 1998 from the Instituto Superior Tecnico, Technical University of Lisbon, Portugal, and a Ph.D. degree in electrical and computer engineering from Rice University, Houston, Texas in 2008. Between 1998 and 2000, he was a researcher with the Communication Theory and Pattern Recognition Group, Institute of Telecommunications, Lisbon, and in 2002 he held a visiting researcher position at the Mathematics Research Center, Bell Laboratories Research. He was a postdoctoral fellow at the University of Wisconsin in 2007–2008, and between 2008 and 2010 he held an Assistant Professor position in the department of electrical engineering at Columbia University. He is currently an assistant professor in the Department of Mathematics and Computer Science of the Eindhoven University of Technology (TU/e), in the Netherlands. His broad research interests include learning theory, statistical signal and image processing, network inference, and pattern recognition. Mr. Castro received a Rice University Graduate Fellowship in 2000 and a Graduate Student Mentor Award from the University of Wisconsin in 2008.

Xiaodong Wang (S’98–M’98–SM’04–F’08) received the Ph.D. degree in Electrical Engineering from Princeton University. He is a Professor of Electrical Engineering at Columbia University in New York. Dr. Wang’s research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled “Wireless Communication Systems: Advanced Techniques for Signal Reception”, published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON INFORMATION THEORY. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.