Adaptive Block Sampling for Spectrum Sensing

Ali Tajer
ECE Department
Wayne State University

H. Vincent Poor
EE Department
Princeton University

***Abstract*—This paper considers spectrum access in wideband channels and its objective is to design an agile and reliable mechanism for identifying spectrum opportunities. Driven by the need for reducing the time required for identifying spectrum opportunities, the idea of data-adaptive and sequential *block* sampling is proposed, through which instead of examining each channel individually, a cognitive user takes samples that are linear combinations of simultaneous measurements from multiple channels. If such *coarse samples* indicate that the block of channels contains at least a vacant (unused) channel, then the channels are examined individually in order to accumulate more information about their spectral occupancy states, and otherwise, the entire block is discarded and the process resumes sequentially by examining the next block of channels.**

I. Background

Current statistics about radio-frequency spectrum occupancy patterns indicate that a considerable fraction of the frequency spectrum is under-utilized. As wireless networks are constantly growing in scale and data traffic, in future networks it is anticipated that the occupancy states of the under-utilized segments of the spectrum will vary rapidly and the spectrum holes might not remain unoccupied for long durations. Therefore, it is of paramount importance to identify the spectrum holes *quickly* and consequently the notion of agile spectrum sensing has received extensive research attention.

A notable direction is the quickest sequential search approach of [1], in which a wideband spectrum is split into smaller narrowband channels and the cognitive users scan them sequentially one-at-a-time. Upon scanning and accumulating enough information about each channel a cognitive user decides whether the channel is vacant (unused) or is occupied. If the channel is determined to be a hole, the search is terminated and otherwise the process is carried on until a hole is detected. This approach is very effective when the vacant channels are not rare. In case of rarity of vacant channels the quick search approaches of [2], [3] and distilled sensing [4] can further reduce the time required for identifying the spectrum opportunities. In these approaches adaptive and sequential experimental designs are proposed in order to gather information from the *entire* spectrum and with the purpose of effectively focusing the sampling resources on the more promising segments of the spectrum.

The existing sequential and data-adaptive methods, irrespective of their discrepancies, all conform in the fact that they sequentially take one sample from each sequence, update their decisions, and decide about their next actions. In contrast, this paper proposes to perform *block* sampling in which a cognitive users first takes *coarse* samples, which are linear combinations of simultaneous measurements of a block of channels, and when the block is deemed to contain spectrum opportunities, then *fine* samples are taken to further refine the information about the spectral occupancy states of the channels within the block.

We remark that there exists a different direction in wideband spectrum sensing in which it is assumed that the wideband channel is heavily under-utilized and used only sparsely. In this approach the cognitive radios exploit the sparsity structure of the wideband channel and use a mixed analog-digital sampling strategy that requires sub-Nyquist sampling in order to perfectly recover the spectral occupancy status of the channel [5] [6]. In another relevant direction, the cognitive radios deploy a compressed sensing-based machinery (which is not data-adaptive) for estimating the power spectral density (PSD) of the wideband channel [7]–[12]. These approaches are further extended in [13] to also track temporal variations of the spectral occupancy during sensing. 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. Finally, the third direction pertains to collaborative sensing in cognitive *networks*, where a group of cognitive radios are clustered to collaboratively perform spectrum sensing [14].

II. Preliminaries

A. Sensing Model

Consider a *wideband* spectrum shared by license-holding (primary) users and cognitive (secondary) users with interference-avoiding spectrum access. The cognitive users are allowed to opportunistically seek the portions of the spectrum under-utilized by the active users (either primary users or secondary ones currently using the spectrum) and access them. We assume that the wideband channel is comprised of n non-overlapping narrowband channels indexed by $\{1,\dots,n\}$ that have independent spectral occupancy states.

We assume that the occupancy status of the spectrum remains unchanged during the spectrum sensing process and consider a dichotomous statistical model for the occupancy of the channels. Let the Bernoulli random variable S_i, for $i \in \{1,\dots,n\}$, indicate the occupancy state of the i^{th} channel, where $S_i = 1$ means that the i^{th} channel is occupied and $S_i = 0$ means that the i^{th} channel is vacant. By assuming that each channel is vacant with probability ϵ we have

$$S_i \overset{\text{i.i.d.}}{\sim} \text{Bernouli}\,(1-\epsilon)\ . \tag{1}$$

This work was supported in part by the National Science Foundation under Grants DMS-1118605, ECCS-1343326, and ECCS-1343210.

In order to distinguish between the vacant and occupied channels we define the sets of vacant and busy channels as

$$\mathcal{H}_0 \triangleq \Big\{i \in \{1,\dots,n\} : S_i = 0\Big\} \ , \tag{2}$$

$$\text{and} \quad \mathcal{H}_1 \triangleq \Big\{i \in \{1,\dots,n\} : S_i = 1\Big\} \ . \tag{3}$$

A cognitive user continuously monitors the channels and takes observations for identifying a vacant channel. Specifically, we denote the set of observations made from channel i by $\mathcal{X}_i \triangleq \{X_i^1, X_i^2, \dots\}$, which are related to the occupancy states of the channels according to

$$X_i^j \ = \ g_i(S_i) \ + \ N_i^j \ , \quad j = 1, 2, \dots \ , \tag{4}$$

where g_i captures the dynamics of sensing (e.g., the fading channel between the user active on channel S_i and the cognitive user) and N_i^j accounts for the observation noise. We assume that the measurements are statistically independent over time and channels. More specifically, conditionally on S_i, the elements of elements of $\mathcal{X}_i$ are independent and identically distributed (i.i.d.) obeying

$$X_i^j \mid S_i \ \sim \ F_i \ , \quad j = 1, 2, \dots \tag{5}$$

where F_0 and F_1 denote the cumulative distribution functions (cdfs) of two distinct distributions on $\mathbb{R}$. The distributions F_0 and F_1 capture the underlying statistical models of the observations taken from the vacant and occupied channels, respectively. For convenience, we assume that F_0 and F_1 have probability density functions (pdfs) f_0 and f_1, respectively.

B. Coarse vs. Fine Sampling

1) Background: Sequential block sampling is closely related to the sequential detection literature, which aims to optimize a balance between decision reliability and the average number of samples taken. The majority of the existing approaches aim to identify *all* of the sequences $\mathcal{X}^i$ that are distributed according to F_0. When the objective is to identify *all* vacant channels, minimizing the *average* number of observations made with constraints on decision quality can be decomposed into minimizing the *average* number of observations necessary for deciding about the spectral state of each individual channel with the same reliability constraints [15]. The optimal test for each channel, which is the test that requires the smallest number of observations and satisfies the reliability constraints, is the sequential probability ratio test (SPRT) [16].

The SPRT approach for the spectrum sensing problem at hand is too conservative in the sense that it tends to make a reliable decision about the states of *all* channels. In contrast, it is possible to skip the channels that can be deemed as weak candidates after some *rough* observations in favor of saving the sensing resources, and consequently, reduce the average time required for identifying the vacant channels. Based on this premise, the quickest detection method proposed and analyzed in [1] aims to find *one* sequence generated according to F_0 (i.e., a vacant channel) and formalizes a detection rule that minimizes the average delay in reaching a decision given that a satisfactory guarantee for the decision is ensured. The optimal test in this setting turns out to be the classical cumulative sum (CUSUM) test.

2) Proposed Sampling Model: As noted above, the existing methods, irrespective of their discrepancies, all conform in the fact that they sequentially take one sample from each sequence, update their decisions, and decide about the next action. In contrast, we propose a data-adaptive sampling model which is a combination of the following two types of sampling strategies:

1) **Coarse sampling:** Instead of taking one sample from each channel at each time, we divide the channels into blocks of size ℓ and take one sample that is a linear combination of ℓ measurements where one measurement is taken from each channel. Such block sampling has, broadly, a two-fold effect. On one hand it takes only one sample for accumulating information about ℓ channels and is substantially smaller than the resources needed by the existing approaches which devote at least one sample to each channel. On the other hand, one combined and aggregated sample is potentially far less informative about the spectral activities of the individual channels in comparison to having ℓ different samples. In order to benefit from the advantage (reduction in sampling rate) and avoid its undesired effects (inaccurate information) these combined samples are used only to obtain some rough confidence about whether the block of channels includes a vacant channel. When a block is deemed to include only busy channels the entire block of channels is discarded and a combined sample is taken from the next block. If the block is deemed to include a vacant channel, then the block is retained for further scrutiny through more refined observations as explained next. The motivation behind such block sampling is that while such rough observations are insufficient for identifying a vacant channel, they might be sufficient for discarding the channels that are very likely to be occupied.
2) **Fine sampling:** Once a block of channels is deemed to contain a vacant channel and is retained for more accurate scrutiny, we perform an SPRT on each of the channels sequentially until a vacant channel is identified. If the block is found not to contain a vacant channel after performing SPRTs on all the channels, the entire block is discarded permanently, and the sampling procedure resumes by taking coarse samples from the next block.

III. Adaptive Block Sampling

A. Sampling Strategy

We define $r \triangleq n/\ell$ as the number of channel blocks[1] and without loss of generality we define

$$\mathcal{G}_i \triangleq \{(i-1)\ell + 1, \dots, i\ell\} \ , \tag{6}$$

as the set of channels grouped in the i-th block for $i \in \{1,\dots,r\}$. With the ultimate objective of identifying T spectrum holes the proposed sampling procedure is initiated by taking *coarse* measurements from all channel groups $\mathcal{G}_1, \dots, \mathcal{G}_r$.

[1]For convenience of notation we assume here that n is an integer multiplier of ℓ; however, all the analysis can be easily extended to address any arbitrary n.

Based on these coarse observations a fraction of the groups that are least-likely to contain a vacant channel are discarded and the rest are retained for more accurate scrutiny. Repeating this procedure successively refines the search support and progressively focuses the observations on the more promising channel groups. More specifically, at each time the sampling procedure selects a *subset* of channel groups $\{\mathcal{G}_1,\dots,\mathcal{G}_r\}$ and takes one coarse sample from each of these groups. Upon collecting these measurements, it takes one of the following actions:

- A_1: **(Observation):** There is insufficient information to decide which groups are most likely to contain spectrum holes; continue to take one more coarse sample from the same groups.
- A_2: **(Refinement):** There is sufficient confidence that some of the groups are very unlikely to contain a vacant channels; discard a portion of the groups with the highest likelihoods of containing of only busy channels. By denoting the number of groups retained prior to a refinement action by r, the number of groups that this action discards is $(1-\alpha)(r-T)$ for some $\alpha\in(0,1)$. Discarding the groups at this rate ensures that at least T groups will be retained for the final detection action (action A_3). When a group of channels are discarded they will be deemed weak candidates for being vacant channels and will not be observed anymore, while the remaining channels are retained for more scrutiny.
- A_3: **(Coarse sampling termination):** There is sufficient confidence that the blocks retained contain vacant channels; stop coarse sampling and start taking fine samples according the steps B_1-B_3 delineated next.

After terminating coarse sampling, all the groups of channel retained are sorted in the descending order of the likelihoods that they contain vacant channels and then we start taking fine samples by first taking a fine sample from the first channel in the first group. Upon collecting a fine sample, the sampling procedure takes one of the following decisions:

- B_1: There is insufficient information to decide whether the channel is vacant; continue to take one more fine sample from the same channel.
- B_2: There is sufficient confidence that the channel is not vacant and there are more channels to examine; discard the channel and take one fine sample from the next channel.
- B_3: There is sufficient confidence that the channel is vacant; stop the sampling process if T vacant channels are identified and, otherwise, continue by taking one fine sample from the next channel.

B. Formulation

In order to formalize actions A_1-A_3 and B_1-B_3 we define two switching functions ψ_c and ψ_f as follows. By denoting indices of the channel groups measured via coarse measurements (actions A_1-A_3) at time $t\in\mathbb{N}$ by $\mathcal{L}_t$ and defining $\bar{\mathcal{L}}$ as the set of all such possible sets we define $\psi_c:\mathbb{N}\rightarrow\bar{\mathcal{L}}$, which maps the time index t to the indices of the channel groups being coarse-measured at time t. Moreover, By denoting the set of indices of the channels measured via fine measurements (B_1-B_3) at time $t\in\mathbb{N}$ by $\mathcal{K}_t$ and defining $\bar{\mathcal{K}}$ as the set of all such possible sets we define $\psi_f:\mathbb{N}\rightarrow\bar{\mathcal{K}}$, which maps the time indes t to the indices of the channels being fine-measured at time t. Hence, when $\mathcal{G}_i\in\psi_c(t)$ we take a coarse measurement at time t from group $\mathcal{G}_i$, where by denoting the sample taken from $\mathcal{G}_i$ at time t by Y_t^i we have

$$\bar{Y}_t^i \;=\; \sum_{l=1}^{\ell} X_l\ , \tag{7}$$

where X_l is a fresh sample from the l-th element of $\mathcal{G}_i$. On the other hand, when $i\in\psi_f(t)$ we take a fine sample from channel i at time t denoted by

$$Y_t^i \;=\; X\ , \tag{8}$$

where X is a fresh sample from channel i. Furthermore we define two stopping time values τ_c and τ_f as the stopping time of the coarse and fine sampling procedures, respectively. Specifically, τ_c is the instant when coarse sampling is terminated (action A_3) and the sampling strategy proceeds with taking fine samples, and τ_f is the time at which the sampling procedure stops taking any further samples and declares the set U as the set of indices of T channels as vacant channels.

Characterizing the optimal sampling strategy and decision rules relies on optimizing an interplay between two performance measures, one being the number of samples taken and the other being the frequency of erroneous decisions. For given stopping time values τ_c and τ_f and a given sequence of mapping functions $\bar{\psi}_c(\tau_c)\;\triangleq\;\{\psi_c(1),\psi_c(2),\dots,\psi_c(\tau_c)\}$ and $\bar{\psi}_f(\tau_c)\triangleq\{\psi_f(1),\psi_f(2),\dots,\psi_f(\tau_f)\}$, the probability of erroneous detection, that is the probability that the detected channel is an occupied channel, is

$$\mathsf{P}(\tau_c,\bar{\psi}_c(\tau_c),\tau_f,\bar{\psi}_f(\tau_f)) \;\triangleq\; \mathsf{P}\big(|U\cap\mathcal{H}_1|\neq 0\big)\ . \tag{9}$$

Our objective is to minimize the frequency of erroneous decisions subject to a hard constraint on the number of samples taken over all possible stopping times and switching rules. Hence, the optimal sampling strategy is the solution to the following optimization problem:

$$\left\{\begin{array}{ll}\min & \mathsf{P}(\tau_c,\bar{\psi}_c(\tau_c),\tau_f,\bar{\psi}_f(\tau_f))\\ \text{s.t.} & \sum_{t=1}^{\tau_c}|\mathcal{L}_t|+\sum_{t=\tau_c+1}^{\tau_c+\tau_f}|\mathcal{K}_t|\;\leq\;S\end{array}\right. \tag{10}$$

where S controls the aggregate number of coarse and fine measurements.

C. Decision Rules

In order to proceed with analyzing the decision rules, we need to find the statistical distributions of the samples Y_t^i. According to the definitions of Y_t^i in (8), it obeys one of the two following hypotheses:

$$\begin{array}{ll}\mathsf{H}_{\text{A},0}: & Y_t^i\;\sim\;F_0\\ \mathsf{H}_{\text{A},1}: & Y_t^i\;\sim\;F_1\end{array}\ . \tag{11}$$

On the other hand, when $\bar{Y}_t^i$ is a coarse sample as defined in (7) it has a mixed distribution, which is characterized by the the number of busy and vacant constituent channels of the corresponding channel block. We define Q_0 and Q_1 as the cdfs of $\bar{Y}_t^i$ when the corresponding block includes *at least one* vacant channel and no vacant channels, respectively. Hence, when $\bar{Y}_t^i$ is a coarse sample it obeys one of the following two hypotheses:

$$\begin{array}{lll} \mathsf{H}_{\mathrm{B},0}: & \bar{Y}_t^i & \sim Q_0 \\ \mathsf{H}_{\mathrm{B},1}: & \bar{Y}_t^i & \sim Q_1 \end{array}, \tag{12}$$

and we denote the pdfs of Q_0 and Q_1 by q_0 and q_1. Based on these definitions and defining the sequences of likelihood ratio values $\{\Lambda_t^{\mathrm{A}}(i)\}$ and $\{\Lambda_t^{\mathrm{B}}(i)\}$, which are initialized as $\Lambda_0^{\mathrm{A}}(i) = \Lambda_0^{\mathrm{B}}(i) = 1$ and track the likelihood ratios pertinent the hypothesis testing problems in (11) and (12), respectively, the actions $\{\mathrm{A}_i, \mathrm{B}_i\}$ at time $t+1$ for $t \in \mathbb{N} \cup \{0\}$ can be formalized as follows.

A_1 : The mapping function is updated as $\psi_c(t+1) = \psi_c(t)$ and for $i \in \psi_c(t+1)$ the likelihood ratio is updated as

$$\Lambda_{t+1}^{\mathrm{A}}(i) = \Lambda_t^{\mathrm{A}}(i) \cdot \frac{q_0(Y_{t+1}^i)}{q_1(Y_{t+1}^i)} . \tag{13}$$

A_2 : The mapping function is updated as $\psi_c(t+1) \subset \psi_c(t)$ and for $i \in \psi_c(t+1)$ the likelihood ratio is updated as

$$\Lambda_{t+1}^{\mathrm{A}}(i) = \Lambda_t^{\mathrm{A}}(i) \cdot \frac{q_0(Y_{t+1}^i)}{q_1(Y_{t+1}^i)} . \tag{14}$$

A_3 : The stopping time is set as $\tau_c = t$ and all the channels in blocks included in $\psi_c(t)$ are retained for further processing under actions $\mathrm{B}_1 - \mathrm{B}_3$.

The values of the sequence of likelihood ratios $\{\Lambda_t^{\mathrm{A}}(i)\}$ are used to dynamically decide what action should be taken at each time. The detailed discussion and the pertinent performance analysis are available in [3].

B_1 : The mapping function is updated as $\psi_f(t+1) = \psi_f(t)$ and for $i \in \psi(t)$ the likelihood ratio is updated as

$$\Lambda_{t+1}^{\mathrm{B}}(i) = \Lambda_t^{\mathrm{B}}(i) \cdot \frac{f_0(\bar{Y}_{t+1}^i)}{f_1(\bar{Y}_{t+1}^i)} . \tag{15}$$

B_2 : The mapping function is updated as $\psi_f(t+1)$= index of the next channel and the likelihood ratio is updated as

$$\Lambda_{t+1}^{\mathrm{B}}(i) = \frac{f_0(\bar{Y}_{t+1}^i)}{f_1(\bar{Y}_{t+1}^i)} . \tag{16}$$

B_3 : Channel i is a vacant channel. If T channels are identified stop further sampling and $\tau_f = t$. Otherwise, set $\psi_f(t+1)$= index of the next channel and the likelihood ratio is updated by resetting $\Lambda_t^{\mathrm{B}}(i) = 1$

$$\Lambda_{t+1}^{\mathrm{B}}(i) = \frac{f_0(\bar{Y}_{t+1}^i)}{f_1(\bar{Y}_{t+1}^i)} . \tag{17}$$

The relationship between the decision rules for identifying the unused channels and the sequence of likelihood ratio tests follow the standard SPRT.

IV. Conclusion

In this paper we have proposed a technique for agile identification of the spectrum holes. The core structure of the sensing strategy is a novel sampling strategy that adaptively and sequentially decides to take either coarse samples, which encompass measurements from multiple channels in one sample, or fine samples, which encompass measurements from only one channel. The data samples collected are processed by sequential probability ratio tests in order to form decisions about the spectral states of the channel blocks and individual channels. This approach is effective in focusing the sampling resources on the promising segments of the spectrum, which are the segments that are more likely to include spectrum holes.

References

[1] L. Lai, H. V. Poor, Y. Xin, and G. Georgiadis, "Quickest search over multiple sequences," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5375–5386, Aug. 2011.

[2] A. Tajer, R. Castro, and X. Wang, "Adaptive sensing of congested spectrum bands," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 6110 – 6125, Sep. 2012.

[3] A. Tajer and H. V. Poor, "Quick search for rare events," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4462 – 4481, Jul. 2013.

[4] J. Haupt, R. Castro, and R. Nowak, "Distilled sensing: Adaptive sampling for sparse detection and estimation," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6222–6235, Sep. 2011.

[5] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics on Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[6] ——, "Wideband spectrum sensing at sub-Nyquist rates," *IEEE Signal Processing Magazine*, vol. 28, no. 4, pp. 102–135, Jul. 2011.

[7] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, Apr. 2007, pp. 1357–1360.

[8] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 2337–2340.

[9] Y. Wang, Z. Tian, and C. Feng, "A two-step compressed spectrum sensing scheme for wideband cognitive radios," in *Proc. IEEE Global Communications Conference*, Miami, FL, Dec. 2010.

[10] S. Hong, "Multi-resolution Bayesian compressive sensing for cognitive radio primary user detection," in *Proc. IEEE Global Communications Conference*, Miami, FL, Dec. 2010.

[11] Z. Yu, S. Hoyos, and B. M. Sadler, "Mixed-signal parallel compressed sensing and reception for cognitive radio," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, NV, Apr. 2008.

[12] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *Proc. IEEE Global Communications Conference*, New Orleans, LA, Dec. 2008.

[13] D. Angelosante and G. B. Giannakis, "RLS-weighted lasso for adaptive estimation of sparse signals," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009.

[14] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 60–73, Nov. 2008.

[15] A. Wald, "Sequential tests of statistical hypotheses," *Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, Jun. 1945.

[16] A. Wald and J. Wolfowitz, "Character of the sequential probability ratio test," *Annals of Mathematical Statistics*, vol. 19, no. 3, pp. 326–339, Sep. 1948.