ADAPTIVE SPECTRUM SENSING AND ESTIMATION

Dennis Wei and Alfred O. Hero, III

University of Michigan, Department of Electrical Engineering and Computer Science
1301 Beal Avenue, Ann Arbor, MI 48109, USA

ABSTRACT

We propose a multistage adaptive approach to spectrum sensing and estimation with the goal of concentrating more sensing resources on spectral components of interest. The allocation of resources to minimize the mean squared estimation error is formulated as a dynamic program. An optimal policy is given for the case of two sensing stages. For more than two stages, tractable approximate policies are developed based on open-loop feedback control (OLFC). These policies improve monotonically with the number of stages, and in particular upon the optimal two-stage policy. A spectrum sensing simulation shows substantial reductions in mean squared error compared to non-adaptive sensing and a recently proposed adaptive method. Performance gains in detecting unoccupied channels are also shown.

Index Terms— Adaptive sensing, resource allocation, spectrum sensing, spectral estimation, cognitive radio/radar

1. INTRODUCTION

Spectral estimation is a classical problem relevant to such areas as array processing, music and acoustics, and the experimental sciences. Recently, much attention has been focused in particular on spectrum sensing for cognitive radio [1] and cognitive radar [2], where the objective is to identify unused portions of spectrum for opportunistic communication and remote sensing by secondary users while minimizing interference to the licensed primary users.

This paper discusses a multistage adaptive approach to spectrum sensing and estimation. We assume that spectral components are drawn from two classes, only one of which is of interest. By observing the spectrum in stages and adapting in response to previous observations, more sensing resources can be concentrated on components of interest, thus improving estimation accuracy. The proposed method can be used for example to better characterize spectral peaks against a noise background. In cognitive radio, the noise power in channels unoccupied by primary users, i.e., spectral holes, can be measured more accurately. Similarly in cognitive radar, bands with low noise or clutter can be identified and estimates of the power therein can improve target detection and tracking. Noise power estimation has not received much attention in the cognitive radio/radar literature, which has mostly focused on hole detection [1, 2].

Several previous works on adaptive spectrum sensing have adopted a multiple hypothesis testing viewpoint for hole detection with an overall resource constraint coupling the channels together. Sequential thresholding policies inspired by distilled sensing [3] are proposed and analyzed in [4–6], showing substantial decreases in sampling and signal-to-noise ratio (SNR) requirements compared to non-adaptive sensing. The gamma observation model in [4] is closest to the one assumed in the present work; however, [4–6] assume

knowledge of the primary user and noise powers (or a worst-case bound) while we assume only statistical knowledge and attempt to learn the parameters. Statistical modelling and learning is also considered in [7], which describes a sensing and access policy for channel occupancies with Markov time evolution. The policy in [7] however senses a single channel at a time whereas we consider the allocation of a sensing budget over multiple channels and stages.

In a somewhat different approach, sequential hypothesis tests are performed on each channel without an overall budget constraint [8, 9]. The test in [8] allows for early declaration of primary user presence under a hard limit on the number of observations. In [9], the channels are considered in sequence until a single hole is identified. Beyond adaptive or dynamic methods, an excellent survey of the main techniques for spectrum sensing can be found in [1].

In this paper, we focus on spectral estimation as opposed to hypothesis testing and consider multistage adaptive sensing under a constraint on the total effort. The problem is formulated as a dynamic program in Section 2. In Section 3.1, an exact dynamic programming policy is presented for the case of two stages. In Section 3.2, a tractable extension to more than two stages is developed based on the open-loop feedback control (OLFC) approximation [10]. The same approach is taken in [11] for a Gaussian observation model; the current work extends [11] to a gamma model. The proposed OLFC policy improves monotonically with the number of stages, and in particular upon the optimal two-stage policy. A spectrum sensing simulation in Section 4 shows substantial reductions in mean squared error (MSE) compared to non-adaptive estimation and the adaptive method of [4]. Moreover, the OLFC policy is competitive with [4] in terms of hole detection even though it is not specifically designed for that purpose. The paper concludes in Section 5.

2. PROBLEM FORMULATION

We consider a frequency band divided into N channels (not necessarily contiguous) indexed by i. Channels belong to one of two classes as indicated by independent Bernoulli random variables I_i. The prior probabilities of having $I_i = 1$ are denoted as $p_i(0)$ and are taken to be uniform, i.e., $p_i(0) = p(0)$ for all i, in the absence of further knowledge. Adaptivity yields larger gains when one of the classes (class 1 without loss of generality) is rare, i.e., $p(0)$ is small and the vector $\mathbf{I}$ is sparse, although we do not assume this.

The channels are observed in T stages indexed by t with effort levels $\lambda_i(t)$, which may represent sensing time in cognitive radio, transmitted energy in cognitive radar, or other resources depending on the application. For a channel with power s_i, the observation $y_i(t)$ in stage t is conditionally distributed as

$$y_i(t) \mid s_i, \lambda_i(t-1) \sim \Gamma\left(\lambda_i(t-1), \frac{s_i}{\lambda_i(t-1)}\right), \quad t = 1, \ldots, T,$$

(1)

where $\Gamma(k, \theta)$ denotes a gamma distribution with shape parameter k and scale parameter θ. It follows that the ratio of the mean of $y_i(t)$ to the standard deviation is $\lambda_i(t-1)$, and hence higher effort results in higher measurement precision. In non-adaptive sensing, the allocations $\boldsymbol{\lambda}(t)$ are determined without reference to the observations. In contrast, with adaptive sensing, $\boldsymbol{\lambda}(t)$ in stage t can depend causally on the observations $\mathbf{Y}(t) = \{\mathbf{y}(1), \ldots, \mathbf{y}(t)\}$. The mapping from $\mathbf{Y}(t)$ to $\boldsymbol{\lambda}(t)$ is referred to as the effort allocation policy. The spectrum $\mathbf{s}$ is assumed to be constant during the observation period. Given $\mathbf{s}$ and $\boldsymbol{\lambda}(t-1)$, $t = 1, \ldots, T$, the observations $y_i(t)$ are assumed to be conditionally independent across channels and stages.

The observation model in (1) arises if the random process underlying the spectrum $\mathbf{s}$ is Gaussian. Then the periodogram estimate of each spectral component is a chi-squared random variable with two degrees of freedom and mean equal to the true component. The average of $\lambda_i(t-1)$ independent periodogram estimates, where $\lambda_i(t-1)$ is an integer, is then gamma-distributed as in (1). Similarly, (1) corresponds to energy detection performed on $\lambda_i(t-1)$ observations of a complex Gaussian frequency component [4]. For convenience and generality, we allow $\lambda_i(t-1)$ to take on arbitrary non-negative values, which can model continuous-valued observation times.

Unlike in [4–6], we do not assume that the powers s_i are known or can be bounded deterministically. Instead, each s_i is assumed to have a prior distribution that depends on the class I_i:

$$s_i \mid I_i \sim \Gamma^{-1}(\alpha_{I_i i}(0), \beta_{I_i i}(0)), \quad I_i = 0, 1, \quad i = 1, \ldots, N, \tag{2}$$

where the inverse gamma distribution with shape parameter $\alpha_{I_i i}(0)$ and scale parameter $\beta_{I_i i}(0)$ is chosen to be conjugate to the gamma distribution in (1). The components s_i are assumed to be independent a priori. As with the probabilities $p_i(0)$, the parameters $\alpha_{I_i i}(0)$ and $\beta_{I_i i}(0)$ are usually initialized to be uniform over channels, i.e., $\alpha_{I_i i}(0) = \alpha_{I_i}(0)$ and $\beta_{I_i i}(0) = \beta_{I_i}(0)$, but with $(\alpha_0(0), \beta_0(0)) \neq (\alpha_1(0), \beta_1(0))$ so that the classes are statistically different.

Our goal is to estimate the powers s_i assuming that only class 1 is of interest. Specifically, the objective is to minimize the MSE,

$$\mathbf{E}\left\{\sum_{i=1}^{N} I_i (\hat{s}_i - s_i)^2\right\}, \tag{3}$$

corresponding to estimates $\hat{s}_i$ based on all T stages of observations, and subject to a constraint on the total sensing budget,

$$\sum_{t=0}^{T-1} \sum_{i=1}^{N} \lambda_i(t) = \Lambda(0). \tag{4}$$

The expectation in (3) is taken over $\mathbf{I}$, $\mathbf{s}$, and $\mathbf{Y}(T)$.

2.1. Formulation as a dynamic program

The selection of an effort allocation policy to minimize the MSE in (3) subject to (4) can be formulated as a dynamic program. This involves the definition of a state $\mathbf{x}(t)$ that allows the cost (3) to be rewritten as a sum of terms indexed by t, each depending only on $\mathbf{x}(t)$ and the control $\boldsymbol{\lambda}(t)$. In the present case, $\mathbf{x}(t)$ is a belief state composed of the parameters of relevant posterior distributions. The first such distribution is of the powers s_i given I_i and observations $\mathbf{Y}(t)$. It can be shown [12] that $s_i \mid I_i, \mathbf{Y}(t)$ retains an inverse gamma distribution as in (2) and is independent of $s_j, j \neq i$ for all t. The parameters of the distribution evolve according to

$$\alpha_{I_i i}(t) = \alpha_{I_i i}(t-1) + \lambda_i(t-1), \tag{5a}$$
$$\beta_{I_i i}(t) = \beta_{I_i i}(t-1) + \lambda_i(t-1)y_i(t), \quad I_i = 0, 1. \tag{5b}$$

Next we consider the posterior distribution for I_i conditioned on $\mathbf{Y}(t)$. As with the variables $s_i \mid I_i, \mathbf{Y}(t)$, it can be shown that $I_i \mid \mathbf{Y}(t)$ remains independent Bernoulli. Defining $p_i(t) = \Pr(I_i = 1 \mid \mathbf{Y}(t))$, we have the following recursion [12]:

$$p_i(t) = \frac{p_i(t-1)f_1}{p_i(t-1)f_1 + (1 - p_i(t-1))f_0} \tag{6}$$

where f_{I_i} is the probability density function of $y_i(t) \mid I_i, \mathbf{Y}(t-1)$. This last quantity follows a beta prime distribution [13],

$$y_i(t) \mid I_i, \mathbf{Y}(t-1) \sim \beta'\left(\lambda_i(t-1), \alpha_{I_i i}(t-1), \frac{\beta_{I_i i}(t-1)}{\lambda_i(t-1)}\right), \tag{7}$$

where $\lambda_i(t-1)$ and $\alpha_{I_i i}(t-1)$ are shape parameters and the third parameter is a scale parameter.

Based on the posterior distributions for s_i and I_i, we define the state to be $\mathbf{x}(t) = (\mathbf{p}(t), \boldsymbol{\alpha}(t), \boldsymbol{\beta}(t), \Lambda(t))$, where $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\beta}(t)$ include all components $\alpha_{I_i i}(t)$ and $\beta_{I_i i}(t)$, and $\Lambda(t)$ represents the budget remaining in stage t. It can be seen from (3) that the estimate $\hat{s}_i$ should be chosen as the conditional mean $\hat{s}_i = \mathbf{E}\{s_i \mid I_i = 1, \mathbf{Y}(T)\} = \beta_{1i}(T)/(\alpha_{1i}(T) - 1)$. The cost can then be expressed as

$$\mathbf{E}_{\mathbf{Y}(T), \mathbf{I}}\left\{\sum_{i=1}^{N} \frac{I_i \beta_{1i}(T)^2}{(\alpha_{1i}(T) - 1)^2(\alpha_{1i}(T) - 2)}\right\}. \tag{8}$$

Using (5) and (7) and noting that $\boldsymbol{\lambda}(T-1)$ does not depend on $\mathbf{y}(T)$, the expectations over $\mathbf{y}(T)$ and $\mathbf{I}$ in (8) can be evaluated to yield [12]

$$\mathbf{E}_{\mathbf{Y}(T-1)}\left\{\sum_{i=1}^{N} \frac{w_i(T-1)}{\alpha_{1i}(T-1) - 1 + \lambda_i(T-1)}\right\}, \tag{9}$$

where

$$w_i(t) = \frac{p_i(t)\beta_{1i}(t)^2}{(\alpha_{1i}(t) - 1)(\alpha_{1i}(t) - 2)}. \tag{10}$$

The quantity in braces in (9) depends only on $\mathbf{x}(T-1)$ and $\boldsymbol{\lambda}(T-1)$ and hence is of the required form. Thus the effort allocation problem is to minimize (9) with respect to $\boldsymbol{\lambda}(0), \ldots, \boldsymbol{\lambda}(T-1)$ subject to the budget constraint (4). The dependence of (9) on $\boldsymbol{\lambda}(0), \ldots, \boldsymbol{\lambda}(T-2)$ is implicit through the distribution of $\mathbf{Y}(T-1)$. Interestingly, the functional form of the cost function (9) is the same as for the case of Gaussian-distributed observations and MSE in [11], with the main difference being the definition of $w_i(t)$ in (10).

3. EFFORT ALLOCATION POLICIES

3.1. Exact dynamic programming

Given the dynamic programming formulation in Section 2.1, an optimal effort allocation policy can be determined through exact dynamic programming. This procedure is tractable for policies of one and two stages. For $T = 1$ (the non-adaptive case), the expectation in (9) is absent and the problem is an explicitly stated optimization constrained by (4). Furthermore, because of the form of (9) and the linearity of the constraint, the optimization problem is convex. For $T = 2$, conditioned on the state $\mathbf{x}(1)$, the second stage $\boldsymbol{\lambda}(1)$ of an optimal policy is determined similarly to the $T = 1$ case. Defining $J_1^*(\mathbf{x}(1))$ as the optimal cost of the second stage, the first stage is determined recursively by solving

$$\min_{\boldsymbol{\lambda}(0)} \quad \mathbf{E}_{\mathbf{y}(1)}\{J_1^*(\mathbf{x}(1)) \mid \mathbf{x}(0), \boldsymbol{\lambda}(0)\} \quad \text{s.t.} \quad \sum_{i=1}^{N} \lambda_i(0) \leq \Lambda(0). \tag{11}$$

Under a prior that is uniform over channels as discussed in Section 2, the initial allocation $\boldsymbol{\lambda}(0)$ is also uniform by symmetry, i.e., $\lambda_i(0) = \rho^{(2)}(0)\Lambda(0)/N$ for all i with $\rho^{(2)}(0) \in [0,1]$. The optimization in (11) then reduces to a line search with respect to $\rho^{(2)}(0)$.

3.2. Open-loop feedback control

For optimal policies with more than two stages, a recursive optimization similar to (11) is required for $t = 0, \ldots, T-2$ and becomes prohibitive for $t > 0$ because of the loss of symmetry. Thus for $T > 2$, we turn to an approximate method known as open-loop feedback control (OLFC) [10]. We consider the selection of the effort allocation $\boldsymbol{\lambda}(t)$ in stage t given the current observations $\mathbf{Y}(t)$, or equivalently the state $\mathbf{x}(t)$. The simplifying assumption in OLFC is that future allocations $\boldsymbol{\lambda}(t+1), \ldots, \boldsymbol{\lambda}(T-1)$ can depend only on $\mathbf{x}(t)$. As shown in [12], this assumption allows the conditional expectation over $\mathbf{y}(t+1), \ldots, \mathbf{y}(T-1)$ in (8) to be evaluated in closed form, in addition to the expectations over $\mathbf{y}(T)$ and $\mathbf{I}$ from before. The result is a cost function similar in form to (9). Consequently, the effort allocation problem under OLFC can be stated explicitly as the following joint optimization over $\boldsymbol{\lambda}(t), \ldots, \boldsymbol{\lambda}(T-1)$:

$$
\min_{\boldsymbol{\lambda}(t),\ldots,\boldsymbol{\lambda}(T-1)} \quad \sum_{i=1}^{N} \frac{w_i(t)}{\alpha_{1i}(t) - 1 + \sum_{\tau=t}^{T-1} \lambda_i(\tau)}
$$
$$
\text{s.t.} \quad \sum_{i=1}^{N} \sum_{\tau=t}^{T-1} \lambda_i(\tau) = \Lambda(t). \tag{12}
$$

Similar to the one-stage effort allocation problem, (12) is a convex optimization and again coincides with the Gaussian case in [11].

As in [11], the OLFC allocation problem (12) yields a water-filling solution for the variables $\overline{\lambda}_i(t) = \sum_{\tau=t}^{T-1} \lambda_i(\tau)$. Define $r_i(t) = \alpha_{1i}(t) - 1$ and re-order the indices i such that the quantities $\sqrt{w_i(t)}/r_i(t)$ are sorted in non-increasing order. Then the optimal value of $\overline{\boldsymbol{\lambda}}(t)$ is given by

$$
\overline{\lambda}_i^*(t) = \left(\Lambda(t) + \sum_{j=1}^{k} r_j(t) \right) \frac{\sqrt{w_i(t)}}{\sum_{j=1}^{k} \sqrt{w_j(t)}} - r_i(t) \tag{13}
$$

for $i = 1, \ldots, k$ and $\overline{\lambda}_i^*(t) = 0$ otherwise. The number of nonzero components k is determined by the interval $(b(k-1), b(k)]$ in which the budget parameter $\Lambda(t)$ falls, where $b(k)$ is a non-decreasing function defined by

$$
b(k) = \frac{r_{k+1}(t)}{\sqrt{w_{k+1}(t)}} \sum_{i=1}^{k} \sqrt{w_i(t)} - \sum_{i=1}^{k} r_i(t), \quad k = 0, \ldots, N-1,
$$

and $b(N) = \infty$. The derivation of the solution in (13) follows [11].

The solution in (13) specifies the distribution of effort over channels but, as can be seen from the form of (12), the division of $\overline{\lambda}_i^*(t)$ into $\lambda_i(t), \ldots, \lambda_i(T-1)$, i.e., the distribution over stages, is not specified. To fully determine the effort allocation, we choose $\boldsymbol{\lambda}(t)$ to be of the form $\boldsymbol{\lambda}(t) = \rho^{(T)}(t)\overline{\boldsymbol{\lambda}}^*(t)$ and perform an additional optimization over $\rho^{(T)}(t) \in [0,1]$, the fraction of the remaining effort budget used in stage t. For $T = 1, 2$, the optimal policies discussed in Section 3.1 belong to the OLFC class with $\rho^{(1)}(0) = \rho^{(2)}(1) = 1$ and $\rho^{(2)}(0)$ chosen according to (11). For $T > 2$, we follow the strategy in [11] and recursively set $\rho^{(T)}(t) = \rho^{(T-1)}(t-1)$ for $t = 1, \ldots, T-1$. Defining $J_0^{(T)}(\mathbf{x}(0))$ to be the cost of a T-stage

OLFC policy given initial state $\mathbf{x}(0)$, the first-stage fraction $\rho^{(T)}(0)$ is chosen as follows:

$$
\rho^{(T)}(0) = \arg \min_{0 \leq \rho \leq 1} \ \mathbf{E}_{\mathbf{y}(1)} \left\{ J_0^{(T-1)}(\mathbf{x}(1)) \mid \mathbf{x}(0), \rho\overline{\boldsymbol{\lambda}}^*(0) \right\}. \tag{14}
$$

Since the optimization in (14) depends only on the initial state $\mathbf{x}(0)$ and the previously determined $T-1$-stage policy, it can be performed offline. Specifically, we perform a line search with respect to ρ and for each ρ we generate samples of $\mathbf{y}(1)$ and simulate the $T-1$-stage policy, starting from the resulting states $\mathbf{x}(1)$, to evaluate the expectation in (14). As shown in [11], this optimization ensures that the OLFC policies improve with the number of stages, i.e.,

$$
J_0^{(T+1)}(\mathbf{x}(0)) \leq J_0^{(T)}(\mathbf{x}(0)), \quad T = 1, 2, \ldots.
$$

In particular, OLFC policies with more than two stages improve upon the optimal two-stage policy in Section 3.1.

4. NUMERICAL SIMULATIONS

We simulate a spectrum sensing scenario to evaluate the proposed OLFC policy and compare it to non-adaptive sensing and the adaptive policy in [4]. For concreteness, we adopt the terminology of cognitive radio although the simulation is equally applicable to cognitive radar. The number of channels N is set to 1000 and classes 0 and 1 correspond to the presence or absence of primary users. The goal therefore is to estimate the noise power in unoccupied channels assuming that they are sparse as also assumed in [4]. Spectra and observations are generated from the model in Section 2 with $\alpha_0(0) = \alpha_1(0) = 6$ and $\beta_1(0) = 5$ so that the mean noise power is normalized to 1. The parameter $\beta_0(0)$ is determined by the primary user SNR, equal to $10 \log_{10}((\beta_0(0) - \beta_1(0))/\beta_1(0))$ dB. Following [4], the sensing budget is chosen as $\Lambda(0) = 5N$ and the allocations $\boldsymbol{\lambda}(t)$ are restricted to be integer-valued. For OLFC, this integer constraint is met by rounding. We make two modifications to the policy in [4]. First, to accommodate the distribution of spectral components assumed in this work, we change the observation threshold in [4] to the median of the beta prime distribution $f(y_i(t) \mid I_i = 0)$ in (7) with $\lambda_i(t-1) = 1$. Second, estimates and decisions are made on the basis of all observations using the Bayesian update equations (5) and (6), and not just the final stage.

In Fig. 1, we show the reduction in MSE relative to non-adaptive sensing for the adaptive policies as a function of SNR. The averages are computed from 40000 simulations. For the same number of stages, the OLFC policy yields lower MSE than the policy in [4] at all SNR, with the differences becoming larger at higher SNR. The $T = 2$ OLFC policy is optimal for policies with two stages. We also compare the integer-rounded OLFC policy ('OLFC') to an unrounded version ('OLFC-NR'). The loss due to rounding is minor and occurs when the MSE reduction is already quite large. Larger gains are seen for $p(0) = 0.01$ since resources can be concentrated to a greater extent. We note that increasing the sensing budget $\Lambda(0)$ has a similar effect as increasing the SNR since both make the two classes more distinguishable.

Figure 2 plots the MSE reduction as a function of the number of stages T for $p(0) = 0.01$ and three SNR levels. The monotonic improvement property of the OLFC policy is verified, except for small finite-sample deviations visible at SNR = 2 dB because of the small gain overall. Incremental gains diminish as T increases but less quickly at lower SNR. For SNR = 10, 20 dB, a 3-stage OLFC policy gives lower MSE than a 10-stage policy from [4], whereas for SNR = 2 dB, [4] does not yield any MSE improvement.
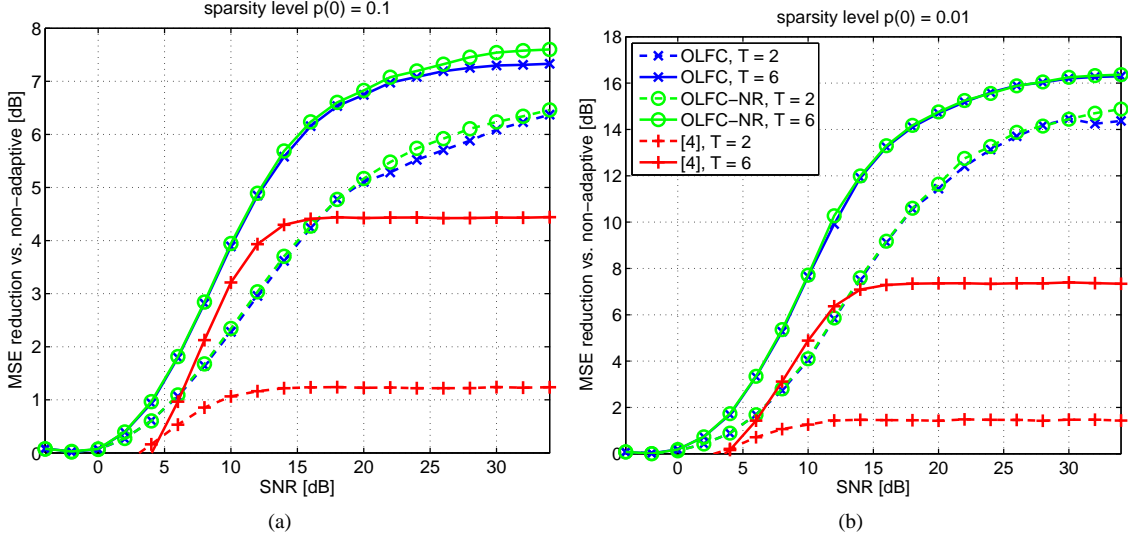
Fig. 1. MSE reduction relative to non-adaptive estimation as a function of SNR for (a) $p(0) = 0.1$ and (b) $p(0) = 0.01$. For the same number of stages T, the proposed OLFC policy yields lower MSE than the policy in [4] at all SNR and especially at higher SNR. At high SNR the benefit of the OLFC policy increases by more than a factor of 6 as the sparsity increases from $p(0) = 0.1$ in (a) to $p(0) = 0.01$ in (b).
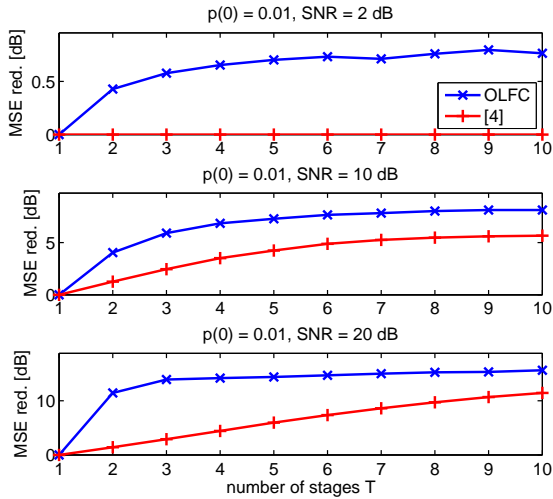

Fig. 2. MSE reduction as a function of the number of stages T. Gains diminish as T increases but less quickly at lower SNR. In all cases shown, a 3-stage OLFC policy results in lower MSE than a 10-stage policy from [4].

Figure 3 compares the hole detection performance of the OLFC and non-adaptive policies and the policy from [4]. The probability of primary user misdetection is shown on a log scale since the limits on interference to primary users are usually stringent while the requirements for hole detection are less so, i.e., it is often unnecessary to detect all or even most holes. In terms of the receiver operating characteristics (ROC), the OLFC policy is competitive with [4] and significantly better than non-adaptive sensing even though it is not specifically designed for detection. At 0 dB SNR, the policy in [4] has a slightly higher hole detection probability at low primary user misdetection probabilities, while the OLFC policy is slightly better at higher misdetection probabilities. At 4 and 10 dB SNR, the OLFC

policy consistently outperforms [4]. Above 10 dB SNR or below 0 dB SNR, the policy in [4] is better once again but both policies perform either very well or very poorly.
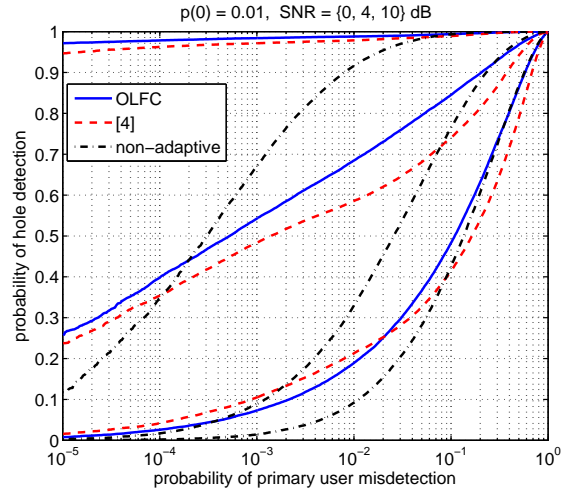

Fig. 3. ROC for the proposed OLFC policy with 6 stages, the policy in [4] with 6 stages, and non-adaptive sensing at SNR = 0, 4, 10 dB (lower to upper within each set of curves). The OLFC policy is competitive with [4] and significantly better than non-adaptive sensing.

5. CONCLUSION

We have discussed a multistage adaptive approach to spectrum sensing and estimation based on dynamic programming. The proposed effort allocation policy results in significantly lower estimation MSE compared to non-adaptive sensing and to the policy of [4]. Gains in hole detection performance are also observed, and further improvements may be expected by developing a dynamic programming policy aimed specifically toward detection.

6. REFERENCES

[1] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio," *IEEE Signal Process. Mag.*, vol. 29, pp. 101–116, May 2012.

[2] M. Wicks, "Cognitive radar: A way forward," in *IEEE Radar Conference (RADAR)*, 2011, pp. 012–017.

[3] J. Haupt, R. M. Castro, and R. Nowak, "Distilled sensing: Adaptive sampling for sparse detection and estimation," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6222–6235, Sept. 2011.

[4] A. Tajer, R. M. Castro, and X. Wang, "Adaptive sensing of congested spectrum bands," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 6110–6125, Sept. 2012.

[5] M. Malloy and R. Nowak, "Sequential analysis in high-dimensional multiple testing and sparse recovery," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aug. 2011, pp. 2661–2665.

[6] M. Malloy and R. Nowak, "On the limits of sequential testing in high dimensions," in *Conf. Rec. Asilomar Conf. Signals Syst. Comput.*, Nov. 2011, pp. 1245–1249.

[7] J. Unnikrishnan and V. V. Veeravalli, "Algorithms for dynamic spectrum access with learning for cognitive radio," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 750–760, Feb. 2010.

[8] W. Zhang, A. K. Sadek, C. Shen, and S. J. Shellhammer, "Adaptive spectrum sensing," in *Proc. Info. Theory Appl. Workshop (ITA)*, 2010, pp. 1–7.

[9] 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.

[10] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, Athena Scientific, Nashua, NH, 3rd edition, 2005.

[11] D. Wei and A. O. Hero, "Multistage adaptive estimation of sparse signals," arXiv:1210.1473, Oct. 2012.

[12] D. Wei and A. O. Hero, "Adaptive spectrum sensing and estimation," in preparation, 2012.

[13] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, Wiley, New York, 2nd edition, 1995.