A Computing Budget Allocation Approach to Multiband Spectrum Sensing

Joseph M. Bruno*, Brian L. Mark*, Yariv Ephraim*, Chun-Hung Chen†
*Dept. of Electrical and Computer Engineering
†Dept. of Systems Engineering and Operations Research
George Mason University
email: { jbruno2, bmark, yephraim, cchen9 }@gmu.edu

***Abstract*—In dynamic or opportunistic spectrum access, the primary user (PU) alternates between an idle and an active state, and a secondary user (SU) may access the channel during the idle periods. In multiband spectrum sensing, an SU tracks the PU state on a given set of channels to determine spectrum access opportunities. In this context, we address the following problem: Given M channels, determine the best subset of $N \ll M$ channels with respect to spectrum access opportunities and, at the same time, estimate the parameter of the PU state process for each channel within the selected subset. Specifically, we model the PU state on the given set of channels by M independent, two-state continuous-time Markov chains. Over a given interval of time, our goal is to determine, with high probability, the N channels with the largest mean idle periods and, at the same time, to accurately estimate the parameter of each channel in the selected subset. We adapt the optimal computing budget allocation (OCBA) methodology from the field of simulation optimization to allocate the total time budget for sensing the N channels in order to perform the channel subset selection and parameter estimation. Simulation results are presented to demonstrate the performance of the proposed algorithm.[1]**

***Index Terms*—Cognitive radio, spectrum sensing, parameter estimation, optimal computing budget allocation**

I. Introduction

Due to rapidly increasing demand for wireless bandwidth, opportunistic spectrum access systems have been proposed to increase utilization of licensed bands. In such systems, the primary user (PU) operates in a licensed band, while unlicensed secondary users (SUs) are permitted to leverage spectrum vacancies left unused by the PU. On a given channel, the PU can be in either an idle state, in which it does not transmit, or an active state, in which it transmits on the channel. To increase spectrum utilization while reducing potentially harmful interference to the PU, SUs can employ a realistic statistical model to estimate and predict the state of the PU, see, e.g., [1], [2]. The parameter of the PU model could be estimated offline [2] or online [3].

In multiband spectrum sensing, the SU tracks states of PUs operating on a given set of channels to determine spectrum access opportunities. We extend the work in [4], where the active/idle state process of each PU is modeled as a two-state homogeneous continuous-time Markov chain, and the Markov chains corresponding to different PUs are assumed statistically independent. We assume M channels, each having the same bandwidth. The parameter of each Markov chain is not known in advance and hence is estimated from observations of the state processes. We assume, as in [4], that the PU state processes are observed directly, and thus we ignore channel impairments. Our analysis is suitable for channels with very high signal-to-noise ratio. Subsequent work should address the adverse affects of the channels.

In the proposed approach, given a sensing interval of length T seconds, the SU senses each channel i for T_i seconds such that $\sum_{i=1}^{M} T_i = T$. We address the problem of allocating the sensing times T_i among the M channels such that a subset of $N \leq M$ channels with the largest mean idle times can be selected. As an additional objective, the total parameter estimation error for channels in the selected set should be minimized. In practice, $N \ll M$, i.e., the number of channels for spectrum sensing is much smaller than the total number of channels in a given spectrum band. When $N = 1$, the problem reduces to allocating the sensing time budget to determine, the channel with the largest mean idle time, and minimizing the estimation error for the associated parameter. When $N = M$, our approach defaults to the framework used in [4], where minimum mean squared error (MMSE) parameter estimation is performed over all M channels. In practice, a small value of N would be chosen to facilitate efficient multiband spectrum sensing.

To address the multichannel estimation problem described above, we adapt the optimal computing budget allocation (OCBA) methodology [5] from the field of simulation optimization. The OCBA framework was developed as a means of testing multiple designs through simulation by allocating simulation time to the designs with the objective of maximizing the probability that the best design is selected according to a given cost function [6]. The technique was subsequently extended to determine the best $N > 1$ designs among a given set of M designs [7]. In the context of multichannel parameter estimation, instead of allocating simulation time we allocate sensing times, and instead of simulating multiple designs, we perform parameter estimation of multiple channels.

The OCBA approach is generally applied iteratively to a sequence of simulation time intervals. Likewise, our proposed

[1]This work was supported in part by the U.S. National Science Foundation under Grants No. 1421869 and 1205453.

algorithm for multichannel parameter estimation iterates over a sequence of sensing intervals. In a departure from the traditional OCBA, we have as an additional objective, minimizing the estimation error for the parameters of the channels in the selected subset. During each iteration of the algorithm, we employ the Bhattarcharyya distance metric [8] to eliminate from consideration channels that are unlikely to belong to the selected subset. This approach allows the sensing resources to be concentrated, in subsequent iterations, on estimation of the channels that are more likely to belong to the selected subset.

The work in [4] allocates the sensing times $\{T_i\}$ by minimizing the Cramér-Rao bound on the minimum mean squared error in estimating the parameters of *all* M channels. Our proposed approach focuses the estimation effort on a much smaller subset of the N most promising channels with respect to mean idle time. In [4], an asymptotic expression for large observation time of the inverse Fisher information matrix (FIM) is used. We refine this result to apply to any finite time interval. By using the asymptotic expression, closed-form formulas for the MMSE sensing time allocations were obtained in [4], as shown in (8). In this paper, we use an exact expression for the FIM, but resort to a numerical optimization approach to solve for the sensing time allocations.

A number of articles on multiband spectrum sensing have approached the problem as a type of multi-armed bandit problem [9]–[12] or the related partially observable Markov decision process (POMDP) [13]. Several assume knowledge of the parameters of the underlying Markov chains, but do not address the important issue of parameter estimation [11]–[13]. Our proposed multichannel parameter estimation algorithm obtains estimates of this parameter, and thus could, in principle, be used in conjunction with these approaches. Moreover, knowledge of the parameter can be used to improve spectrum detection performance and allows the prediction of future PU state, which provides clear advantages for spectrum sensing [2], [12]. In [2], for example, a likelihood ratio detector for PU state on a given channel is proposed based on an estimate of the associated parameter.

The rest of the paper is organized as follows. In Section II, we present the system model assumed in the paper. In Section III, we summarize the multichannel estimation algorithm in [4], which is based on minimizing the mean squared error over all channels. In Section IV, we develop the proposed algorithm for multichannel estimation based on OCBA and the Bhattaacharyya distance measure. In Section V, we present simulation results that demonstrate the performance of the algorithm. The paper is concluded in Section VI with additional comments.

II. System Model

Consider a multiband spectrum sensing scenario consisting of M channels, which an SU may leverage for opportunistic spectrum access. In each band, an independent PU is operating. Each PU is modelled by a two-state continuous-time Markov chain as depicted in Fig. 1, where state 0 represents an idle PU and state 1 represents a busy PU. An SU may only use the band when the PU is in the idle state. For a given PU, let $\{X_t, t \geq 0\}$ denote the Markov chain associated with the state process. The transition rate from state 0 to 1 is denoted λ_0, and the transition rate from state 1 to 0 is denoted λ_1. The parameter of the Markov chain is given by $\theta = (\lambda_0, \lambda_1)$. For simplicity, we assume, as in [4], that the SU directly observes the PU state process $\{X_t, t \geq 0\}$. The model could be extended to incorporate channel impairments, as was done in [2], for example.

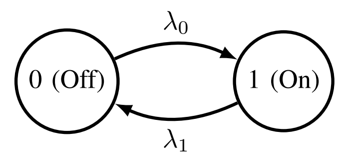

Fig. 1. Continuous-time Markov chain model for PU state of a single channel.

Let $N_t(j,k)$ denote the number of jumps of the PU state from state j to state k over the time interval $[0,t)$, and denote its expected value by $\overline{N}_t(j,k)$, where $j,k \in \{0,1\}$. We assume that the Markov chain $\{X_t\}$ has initial state probabilities $\{\pi_0, \pi_1\}$, where $\pi_j = P(X_0 = j)$, $j = 0,1$. The Fisher information matrix (FIM) for $\{X_t\}$ can be derived as follows:

$$I(t) = \begin{bmatrix} \dfrac{\overline{N}_t(0,1)}{\lambda_0^2} & 0 \\ 0 & \dfrac{\overline{N}_t(1,0)}{\lambda_1^2} \end{bmatrix}. \tag{1}$$

The derivation follows directly from the definition of the FIM and an expression for the likelihood function of a continuous-time Markov chain given in [14, Sec. 4]. Let $\lambda = \lambda_0 + \lambda_1$. The expected number of jumps from state j to state k can be expressed as follows (see Appendix A for a sketch of the derivation):

$$\overline{N}_t(j,k) = \frac{\lambda_j \lambda_k}{\lambda} t + \left(\pi_j - \frac{\lambda_k}{\lambda}\right) \frac{\lambda_j}{\lambda} \left(1 - e^{-\lambda t}\right). \tag{2}$$

An expression for the inverse FIM in the nonstationary case, where $\{\pi_0, \pi_1\}$ are assumed to be constants, follows from (1) and (2). The stationary case does not lead to a simple closed-form expression. An asymptotic expression for the FIM of a stationary two-state Markov chain, valid in the regime of large t, was derived in [4, Theorem 1]. The asymptotic FIM leads to a closed-form solution for the sensing time allocations, see [4, Eq. (17)], but incurs non-negligible approximation error for smaller values of t.

We use a subscript i to denote the ith channel, e.g., we denote the PU Markov process for the ith channel as $\{X_{i,t}, t \geq 0\}$ and its associated parameter by $\theta_i = (\lambda_{i,0}, \lambda_{i,1})$. Let $\sigma_i^2(t)$ denote the sum of the variances in estimating the two components of θ_i by an unbiased estimator over a sensing interval of length t seconds. Let $I_i(t)$ denote the FIM for the ith channel over the same sensing interval. The Cramér-Rao bound (CRB) for a single channel i over time t is given by [4]

$$\sigma_i^2(t) \geq \mathrm{trace}[I^{-1}(t)]. \tag{3}$$

The asymptotic form of the CRB used in [4] follows from the asymptotic approximation for the inverse FIM [4, Eq. (6)] for the two-state Markov chain. The asymptotic CRB was derived earlier by Albert [14] for a general finite-state Markov chain. Applying (1), the CRB is given by

$$\sigma_i^2(t) \geq \frac{\lambda_{0,i}^2}{\overline{N}_t(0,1)} + \frac{\lambda_{1,i}^2}{\overline{N}_t(1,0)}. \tag{4}$$

The sensing interval of length t is partitioned into subintervals of length t_i, $i = 1, \ldots, M$, where t_i is the time spent estimating channel i, with $t = \sum_{i=1}^{M} t_i$. The SU observes $\{X_{i,t}\}$ for an interval of length t_i in the sequence $i = 1, \ldots, M$. The CRB for estimation of all M channels over a sensing interval of length t is given by

$$\sigma^2(t) \geq \sum_{i=1}^{M} \left(\frac{\lambda_{0,i}^2}{\overline{N}_{t_i}(0,1)} + \frac{\lambda_{1,i}^2}{\overline{N}_{t_i}(1,0)} \right). \tag{5}$$

III. MMSE Multichannel Estimation [4]

In this section, we summarize the multichannel parameter estimation algorithm in [4], which iteratively determines the sensing intervals and parameter estimation using (5). We assume that time is divided into a sequence of sensing intervals, $\{T_n\}_{n=0}^{\infty}$. Each sensing interval is in turn subdivided into M sensing subintervals, one for each of the M channels. Let T_n denote the duration of the nth sensing interval, and let $T_{i,n}$ denote the duration of the nth sensing interval that is devoted to channel i, such that $T_n = \sum_{i=1}^{M} T_{i,n}$. To perform the initial iteration of multiband sensing, all M channels are sensed for exactly the same amount of time, i.e., we set

$$T_{i,0} = \frac{T_0}{M}, \; i = 1, \ldots, M. \tag{6}$$

During each sensing interval, each channel is sensed in sequence.

For channel i, the number of sojourns in each state is counted, and the counts recorded up to and including time t are denoted $n_i^{\mathrm{off}}(t)$ and $n_i^{\mathrm{on}}(t)$ for the off and on states, respectively. The jth recorded sojourn times for channel i are denoted by $z_{i,j}^{\mathrm{off}}$ and $z_{i,j}^{\mathrm{on}}$, respectively. The estimator for the transition rates proposed in [4], referred to in that paper as the "moment estimator," is given by:

$$\hat{\lambda}_{0,i}(t) = \frac{n_i^{\mathrm{off}}(t) - 1}{\sum_{j=1}^{n_i^{\mathrm{off}}(t)} z_{i,j}^{\mathrm{off}}}, \; \hat{\lambda}_{1,i}(t) = \frac{n_i^{\mathrm{on}}(t) - 1}{\sum_{j=1}^{n_i^{\mathrm{on}}(t)} z_{i,j}^{\mathrm{on}}}, \tag{7}$$

respectively. The parameter estimate $\hat{\theta}_i$ obtained at the end of the nth sensing interval is given by (7), and is used to calculate the sensing time allocations for the next sensing interval. This estimator is the well-known maximum likelihood estimator (MLE), with consistency and asymptotic normality proved by Albert [14, Theorem 6.10]. Asymptotic efficiency of this estimator was also proved by Albert [14, Section 7]. Efficiency of the moment estimator was demonstrated in [4] through simulations.

Multichannel parameter estimation should be designed in such a way that the sensing intervals are used most effectively. In [4], sensing time allocations were derived such that the right-hand side of the multichannel CRB in Eq. (5) is minimized. In this derivation, an asymptotic approximation for inverse FIM was used, which led to closed-form expressions for the proportion α_i of the sensing interval T_n that should be allocated to channel i, given as follows [4, Eq. (17)]:

$$\alpha_i = \frac{\sqrt{\left(\hat{\lambda}_{0,i}^2 + \hat{\lambda}_{1,i}^2\right)\left(\frac{1}{\hat{\lambda}_{0,i}} + \frac{1}{\hat{\lambda}_{1,i}}\right)}}{\sum_{j=1}^{M} \sqrt{\left(\hat{\lambda}_{0,j}^2 + \hat{\lambda}_{1,j}^2\right)\left(\frac{1}{\hat{\lambda}_{0,j}} + \frac{1}{\hat{\lambda}_{1,j}}\right)}}, \tag{8}$$

where the moment estimates, $\hat{\lambda}_{0,i}$ and $\hat{\lambda}_{1,i}$, are computed by applying (7) and $T_{i,n} = \alpha_i T_n$, $i = 1, \ldots, M$. The allocation strategy based on (8) was shown in [4] to approach, as $T_n \to \infty$, the CRB for the joint estimation of all M independent channels.

IV. OCBA Multichannel Parameter Estimation

In this section, we focus on an approach for selecting a smaller subset of the most promising channels for opportunistic spectrum access, while estimating their associated parameters. We adapt OCBA to determine the appropriate sensing time allocations to achieve this objective.

A. OCBA Sensing Allocations

Our goal is to determine the N channels with largest mean dwell time in the idle state, where ideally $N \ll M$. Equivalently, we seek the N channels with minimum cost, where the cost function for channel i is defined by

$$J_i = \lambda_{0,i}. \tag{9}$$

We assume that the initial sensing interval T_0 is allocated according to (6). At the end of the nth sensing interval of length T_n, the PU parameter is re-estimated using Eq. (7) and we apply OCBA [5] to determine the channels with the lowest cost functions. In the context of simulation optimization, given a fixed total computing budget and M alternative designs, OCBA determines the computing budget allocation for simulating the M designs that maximizes the probability of selecting the subset of N designs out of M with minimum cost. The OCBA methodology requires knowledge of the standard deviation of the cost function, which we denote by s_i for channel i. A lower bound on the standard deviation of the cost function (9) follows from the CRB for estimating λ_0, which can be derived from the FIM in Eq. (1):

$$s_i(t) \geq \sqrt{\frac{\lambda_{0,i}^2}{\overline{N}_t(0,1)}}. \tag{10}$$

To find the subset of the N best channels, we first sort the estimated values of the cost function J in (9), denoted by $\hat{J}_i = \hat{\lambda}_{0,i}(t)$, such that $\hat{J}_1 \leq \hat{J}_2 \leq \ldots \leq \hat{J}_{M-1} \leq \hat{J}_M$.

We then compute a reference constant c, which in [7] is the midpoint between the highest cost value of the selected subset and the next highest cost value among the M channels, i.e.,

$$c = \frac{\hat{J}_N + \hat{J}_{N+1}}{2}. \tag{11}$$

We denote the total sensing time on channel i up to and including iteration n as $\Sigma_{T,i,n}$. Applying OCBA, we must next find sensing intervals such that [7]

$$\frac{\Sigma_{T,1,n}}{\left(s_1/(\hat{J}_1 - c)\right)^2} = \ldots = \frac{\Sigma_{T,M,n}}{\left(s_M/(\hat{J}_M - c)\right)^2} \tag{12}$$

and

$$\sum_{i=1}^{M} \left(\Sigma_{T,i,n} - \Sigma_{T,i,n-1}\right) = T_n. \tag{13}$$

The resulting sensing time allocations are then given by [7]

$$T_{i,n} = \Sigma_{T,i,n} - \Sigma_{T,i,n-1}. \tag{14}$$

The standard deviations s_i in (12) can be approximated by computing the right-hand side of (10) using the current estimates $\hat{\lambda}_{0,i}(t)$.

B. Channel Elimination

Once a desired level of certainty that a channel is not a member of the selected subset has been reached, our algorithm ceases allocating sensing time to that channel for the current and future iterations. This results in a smaller parameter estimation error for the channels in the eventual selected subset compared to the standard OCBA. To determine whether a channel i is unlikely to be part of the eventually selected subset, we compare it to the member of the current selected subset with the highest cost, $\overline{J}_N$. We make use of the Bhattacharyya distance for this purpose. The Bhattacharyya distance between a pair of normal random variables $U \sim \mathcal{N}(\mu_u, \sigma_u^2)$ and $V \sim \mathcal{N}(\mu_v, \sigma_v^2)$ is given by [15, p. 777]

$$D_B(U,V) = \frac{(\mu_u - \mu_v)^2}{4(\sigma_u^2 + \sigma_v^2)} + \frac{1}{4}\log\left[\frac{1}{4}\left(\frac{\sigma_u^2}{\sigma_v^2} + \frac{\sigma_v^2}{\sigma_u^2} + 2\right)\right]. \tag{15}$$

Channel i is eliminated if

$$D_B(\hat{J}_N, \hat{J}_i) > \gamma, \;\; \hat{J}_i > \hat{J}_N, \tag{16}$$

where γ is a threshold chosen by the system designer. A larger value of γ will allow for increased certainty of correct decision at the expense of longer time required to obtain the final selected subset of N channels.

When we are left with the selected subset of N channels, subsequent estimation effort can be applied to these channels. In the case of $N = 1$, the optimal sensing strategy is simply to allocate all sensing time to the selected channel. More generally, MMSE allocations are applied to the N selected channels, and the other $M - N$ channels receive no sensing time allocation. Thus, parameter estimation proceeds along the lines of [4] for the selected set of N channels, except that we use the closed-form expression (2) to calculate the FIM. Consequently, we resort to sequential quadratic programming [16] to numerically solve for the sensing time allocations. Nevertheless, the closed-form sensing time allocations given by (8) may serve as good initial values for the algorithm.

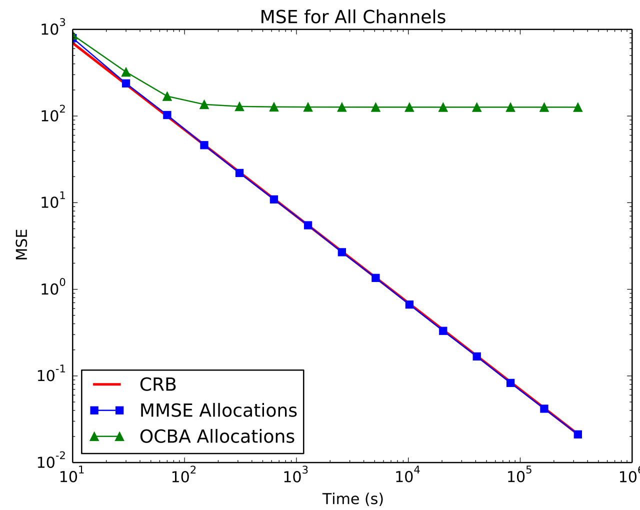

Fig. 2. MSE for all channels using MMSE and OCBA allocations

V. Numerical Examples

To test the proposed algorithm, we present numerical results of two example scenarios based on examples that were considered in [4]. Results were obtained using the Python packages SciPy, NumPy, and Matplotlib.

A. Example 1

In the first example, four channels were defined with parameter values given as follows:

$$\{\lambda_{0,i}\} = \{10, \; 60, \; 20, \; 50\},$$
$$\{\lambda_{1,i}\} = \{70, \; 10, \; 80, \; 90\}.$$

We used an initial allocation of $T_0 = 10$ seconds, increased sensing time by $T_n = 2T_{n-1}$ for each sensing iteration, and performed 15 sensing iterations. As a baseline, we performed the MMSE allocations as proposed in [4], and we compared to OCBA allocations as proposed in this paper. For the fist simulation, we searched for a selected subset of size $N = 1$, i.e., we only searched for a single best channel. We recorded the MSE for the known best channel as well as the system MSE for all 4 channels. A total of 200 simulations were performed, and a minimum Bhattacharyya distance of $\gamma = 18$ was used.

The results of this example are plotted in Figs. 2 and 3. Fig. 2 shows the total MSE in estimating the parameters of all channels compared to the associated CRB. After the best channel is determined, i.e., channel $i = 1$ with $\lambda_{0,i} = 0.1$, the MSE for all channels in the case of OCBA allocation diverges from the CRB. This is because all of the sensing time is devoted to estimation of the parameter of the selected channel,

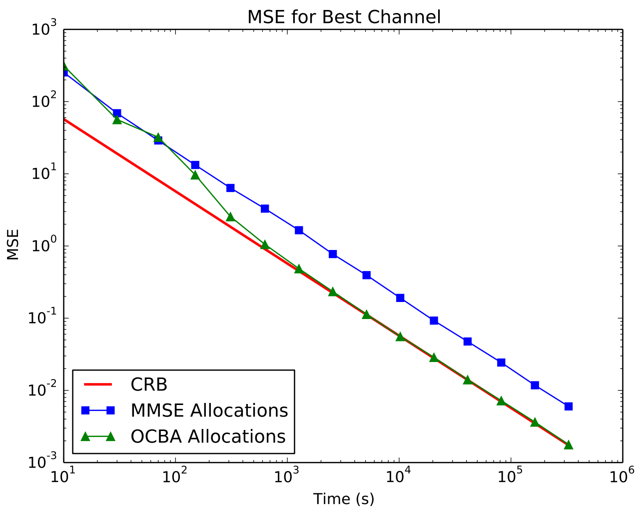

Fig. 3. MSE for best channel using MMSE and OCBA allocations

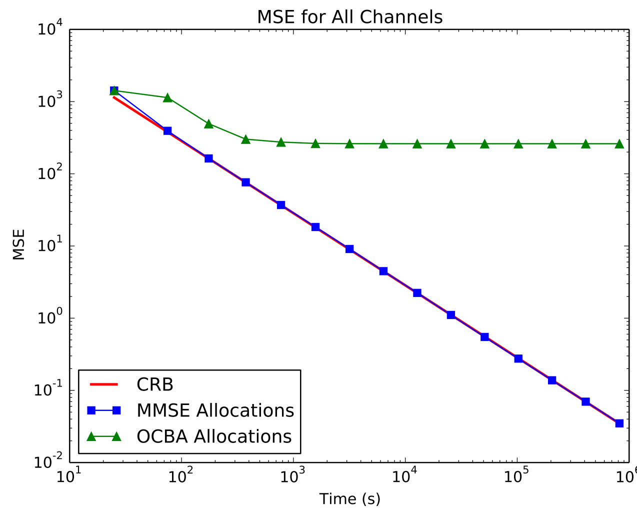

Fig. 4. MSE for all channels using MMSE and OCBA allocations

while the MSE for the other channels remains constant. Fig. 3 shows the MSE for only the best channel in terms of the longest mean dwell time in the idle state. The MSE resulting from both allocation strategies is compared to the CRB under the assumption that all sensing time is given to the best channel.

B. Example 2

Here, 10 channels were defined with the following parameter values:

$$\{\lambda_{0,i}\} = \{10,\ 20,\ 30,\ 40,\ 50,\ 60,\ 70,\ 80,\ 90,\ 100\},$$
$$\{\lambda_{1,i}\} = \{20,\ 70,\ 10,\ 10,\ 40,\ 80,\ 40,\ 90,\ 10,\ 40\}.$$

We used an initial allocation of $T_0 = 25$ seconds, increased sensing time by $T_n = 2T_{n-1}$ for each sensing iteration, and performed 15 sensing iterations. Again, the resulting MSE in estimating the parameters of the channels from MMSE allocations was compared to that obtained from the OCBA allocations. For the OCBA allocations, we set $N = 3$. We recorded the MSE for the set consisting of the best three channels, i.e., $\{1, 2, 3\}$, as well as the system MSE for all 10 channels. In total, 200 simulation runs were performed, and a minimum Bhattacharyya distance of $\gamma = 18$ was used.

The results for Example 2 are plotted in Figs. 4 and 5. Fig. 4 shows the total MSE for estimation of all channels, compared to the CRB for all channels. As in Fig. 2, the MSE for the OCBA allocation in Fig. 4, diverges from the CRB after the best subset is selected. Fig. 5 shows the MSE for only the selected subset. The MSE resulting from both allocation strategies is compared to the CRB under the assumption that all sensing time is given to the selected subset.

C. Discussion

The numerical examples demonstrate that when total system MSE is to be minimized for all channels, the MMSE approach proposed in [4] achieves the CRB. However, when

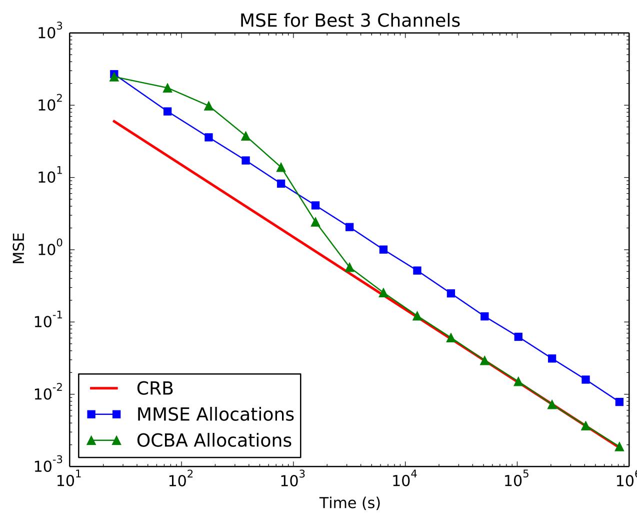

Fig. 5. MSE for optimal subset using MMSE and OCBA allocations

the performance of a smaller subset of channels is more important, the proposed algorithm based on OCBA may be used to quickly determine the selected subset and to focus subsequently on the most promising channels. This is an important feature of our approach, as the total number of channels M may be significantly larger than the eventual number of channels of interest, i.e., $M \gg N$. In such a scenario, diluting the spectrum sensing effort among M channels is heavily resource-intensive, and ultimately is likely to result in degraded opportunistic spectrum access. Thus, focusing the parameter estimation on a much smaller set of N candidate channels is more efficient both computationally and in terms of exploiting the spectrum access opportunities available among the original set of M channels. Given accurate parameter estimates for the chosen set of N channels, detection and/or prediction of the PU states on these channels can proceed using a variety of approaches based on Markovian models, see, e.g., [2], [11]–[13].

VI. Conclusion

We proposed a parameter estimation algorithm for multi-band spectrum sensing based on a Markovian model and compared it with an earlier algorithm of [4] that relies on an MMSE approach. Over a sequence of sensing intervals, the proposed algorithm iteratively allocates a partition of each interval among M channels in such a manner as to determine a smaller set of $N \ll M$ channels that are "best" with respect to a certain cost function while obtaining estimates of the corresponding channel parameters. The algorithm allocates initial sensing times to each of the M channels and then updates the allocations after each subsequent sensing interval. The OCBA methodology from [7] is applied to adjust the allocations with the goal of determining the best N channels at each iteration. To further reduce estimation error, the Bhattacharyya distance is used to eliminate certain channels from further consideration in subsequent iterations. The algorithm was verified through simulation and shown to approach the CRB on the variance of the parameter estimator for channels in the selected set.

In this paper, we have assumed that the PU state is directly observable. The Markov chain model could be extended to take into account channel impairments and to accommodate non-exponential PU state sojourn time distributions, as was done in [2]. We have also assumed a simple cost function for the channels based on the mean idle time of the PU. Alternative cost functions may be considered depending on the spectrum access needs of the SU. For example, in addition to the mean idle time, the variance of the idle time and the channel bandwidth could, in principle, be incorporated into the cost function.

Appendix

A. Expected Number of Jumps Between States

Let X_t denote the state of the PU on a given channel at time t. Consider a sample function of the process such that $X_t = 1$. Then the PU state alternates from 0 to 1, and the cycle of a sojourn in state 0 followed by a sojourn in state 1 repeats. Alternatively, if $X_0 = 1$, a cycle consists of a sojourn in state 1 followed by a sojourn in state 0. For our purposes, the sequence of states within a cycle is immaterial. Let us call the completion of each cycle a renewal event. Let $\Lambda(t)$ denote the number of renewal events in $(0, t]$. The number of transitions from state 0 to 1 in $(0, t]$ can be related to $\Lambda(t)$ as follows:

$$N_t(0,1) = \Lambda(t) + 1_{\{X_0=0, X_t=1\}}, \tag{17}$$

where 1_A is the indicator function of event A. This additional term is necessary to account for the event of an additional transition from state 0 to state 1 without the completion of an entire renewal, i.e., a transition back to state 0.

Letting $R(t) = \mathrm{E}[\Lambda(t)]$, the expected number of transitions from 0 to 1 can be written as

$$\overline{N}_t(0,1) = R(t) + \pi_1 P_{01}(t). \tag{18}$$

where $P_{01}(t) = \mathrm{P}[X_t = 1 \mid X_0 = 0]$. The transition probability $P_{01}(t)$ is given by [17, Ch. 3, Eqs. (28)-(29)]:

$$P_{01}(t) = \frac{\lambda_0}{\lambda_0 + \lambda_1}\left(1 - e^{-(\lambda_0+\lambda_1)t}\right). \tag{19}$$

An expression for $R(t)$ can be obtained using some results from renewal theory, see, e.g., [18], [19], as follows:

$$R(t) = \frac{\lambda_0\lambda_1}{\lambda_0 + \lambda_1}t - \frac{\lambda_0\lambda_1}{(\lambda_0+\lambda_1)^2}\left(1 - e^{-(\lambda_0+\lambda_1)t}\right). \tag{20}$$

Applying (20) and (19) in (18), we obtain (2) when $j = 0$ and $k = 1$. The case $j = 1$ and $k = 0$ is derived similarly.

References

[1] I. Akbar and W. Tranter, "Dynamic spectrum allocation in cognitive radio using hidden Markov models: Poisson distributed case," in *SoutheastCon, 2007. Proceedings. IEEE*, March 2007, pp. 196–201.

[2] T. Nguyen, B. L. Mark, and Y. Ephraim, "Spectrum sensing using a hidden bivariate Markov model," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4582–4591, Aug. 2013.

[3] Y. Sun, B. L. Mark, and Y. Ephraim, "Online parameter estimation for temporal spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4105–4114, Aug. 2015.

[4] P. Tehrani, L. Tong, and Q. Zhao, "Asymptotically efficient multi-channel estimation for opportunistic spectrum access," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5347–5360, Oct. 2012.

[5] C. H. Chen and L. H. Lee, *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. World Scientific Publishing Co., 2010.

[6] C.-H. Chen, J. Lin, E. Yücesan, and S. E. Chick, "Simulation budget allocation for further enhancing the efficiency of ordinal optimization," *Discrete Event Dynamic Systems*, vol. 10, no. 3, pp. 251–270, 2000.

[7] C.-H. Chen, D. He, M. Fu, and L. H. Lee, "Efficient simulation budget allocation for selecting an optimal subset," *INFORMS Journal on Computing*, vol. 20, no. 4, pp. 579–595, 2008.

[8] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol. 15, no. 1, pp. 52–60, 1967.

[9] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: Structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 5431– 5440, Dec. 2008.

[10] J. Oksanen, V. Koivunen, and H. V. Poor, "A sensing policy based on confidence bounds and a restless multi-armed bandit model," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2012, pp. 318–323.

[11] S. Bagheri and A. Scaglione, "The restless multi-armed bandit formulation of the cognitive compressive sensing problem," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1183–1198, March 2015.

[12] K. Wang, L. Chen, Q. Liu, W. Wang, and F. Li, "One step beyond myopic probing policy: A heuristic lookahead policy for multi-channel opportunistic access," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 759–769, Feb 2015.

[13] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, April 2007.

[14] A. Albert, "Estimating the infinitesimal generator of a continuous time, finite state Markov process," *Annals of Mathematical Statistics*, vol. 23, no. 2, pp. 727–753, Jun. 1962.

[15] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proceedings of IEEE*, vol. 67, no. 5, pp. 773–785, May 1969.

[16] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.

[17] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to stochastic processes*. Waveland Press, 1986.

[18] E. Çinlar, *Introduction to Stochastic Processes*. Dover Publications, 2013, reprint of 1975 edition published by Prentice-Hall.

[19] H. Kobayashi, B. L. Mark, and W. Turin, *Probability, Random Processes, and Statistical Analysis*. Cambridge University Press, 2012.