

Spectrum Sensing Using a Hidden Bivariate Markov Model

Thao Nguyen, *Student Member, IEEE*, Brian L. Mark, *Senior Member, IEEE*, and Yariv Ephraim, *Fellow, IEEE*

Abstract—A new statistical model, in the form of a hidden bivariate Markov chain observed through a Gaussian channel, is developed and applied to spectrum sensing for cognitive radio. We focus on temporal spectrum sensing in a single narrowband channel in which a primary transmitter is either in an idle or an active state. The main advantage of the proposed model, compared to a standard hidden Markov model (HMM), is that it allows a phase-type dwell time distribution for the process in each state. This distribution significantly generalizes the geometric dwell time distribution of a standard HMM. Measurements taken from real data confirm that the geometric dwell time distribution characteristic of the HMM is not adequate for this application. The Baum algorithm is used to estimate the parameter of the proposed model and a forward recursion is applied to online estimation and prediction of the state of the cognitive radio channel. The performance of the proposed model and spectrum sensing approach are demonstrated using numerical results derived from real spectrum measurement data.

Index Terms—Cognitive radio, spectrum sensing, hidden Markov model, bivariate Markov chain, Baum algorithm.

I. INTRODUCTION

SPECTRUM scarcity in wireless communications can be addressed, to some extent, through cognitive radio, without increasing the allocated bandwidth. In this regard, a critical aspect of cognitive radio is the ability to sense the wireless environment for unused portions of the spectrum. Over the past decade, several paradigms for spectrum sensing have emerged. In the temporal spectrum sensing paradigm, the cognitive radio senses a given channel for idle time intervals of the primary or licensed transmitter, to allow a secondary user to temporarily access the channel during such intervals without causing harmful interference to the primary system [1]. An alternative to temporal sensing is spatial spectrum sensing, which capitalizes on the geographical distance and propagation loss between the primary and secondary systems [2]. In wideband spectrum sensing, the focus is on the detection of idle spectrum across a band of frequencies [3]. Many of the spectrum sensing schemes proposed in the literature incorporate some form of collaboration among the cognitive radio nodes to improve sensing performance [1], [4]. Reviews of recent progress and issues in the area of spectrum sensing can be found in [5], [6].

Manuscript received November 26, 2012; revised April 12 and June 25, 2013; accepted June 28, 2013. The associate editor coordinating the review of this paper and approving it for publication was A. Giorgetti.

This work was supported in part by the U.S. National Science Foundation under Grant No. 0916568.

The authors are with the Dept. of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030 (e-mail: {tnguyed, bmark, yephrim}@gmu.edu).

Digital Object Identifier 10.1109/TWC.2013.072513.121864

In this paper we focus on temporal sensing of a single narrowband channel licensed to a primary transmitter that alternates between active and idle states with randomly distributed state dwell times. As an example, Fig. 3 in Section V-A shows a two-hour snapshot of a paging signal, obtained from real spectrum measurements [7], where the active and idle states can be clearly distinguished. This paper makes several contributions to temporal spectrum sensing. A new statistical model, in the form of a discrete-time hidden bivariate Markov chain, is proposed to characterize the transmission pattern of a primary user. We derive properties of the model and show that the dwell time in a given state has a discrete phase-type distribution, which is more suitable for modeling a cognitive radio channel than the geometric dwell time distribution of a standard hidden Markov model (HMM). Empirical evidence presented in [8] and in this paper (see Section V-A) suggest that the active and idle periods of the bursty transmissions of a wireless local area network are not geometrically distributed. To incorporate the effects of propagation loss and lognormal shadowing into the model, we assume that the bivariate Markov chain is observed through a memoryless Gaussian channel. The resulting hidden bivariate Markov model (HBMM) may be viewed as an extension of the standard HMM.

We apply the Baum algorithm [9] to estimate the parameter of the HBMM given spectrum measurement data. An HBMM parameter estimate can be used in two ways: (i) to generate simulated data for performance evaluation of cognitive radio networks; (ii) as input to a forward recursion for estimating and predicting the state of a primary transmitter in real-time. The estimated and predicted states from the forward recursion are used by the cognitive radio to make spectrum access decisions. The proposed HBMM model and HBMM-based spectrum sensing scheme are validated using real spectrum measurement data from [7]. The performance of the HBMM-based spectrum detector is compared to that of an energy detector, which is a simple spectrum sensing scheme based on comparing the received signal strength to a threshold (cf. [10]).

Most of the existing work on temporal spectrum sensing has applied the more standard *univariate* Markov chain to model the state process (cf. [11]–[16]). In [8], a continuous-time semi-Markov process was proposed as a model for the state process, but no estimation algorithms were developed. Our work may be seen as a generalization of some prior cognitive radio models in the form of hidden Markov processes. In [11], [12], a standard HMM in the form of a univariate Markov chain observed through a memoryless Gaussian channel was applied to cognitive radio. In [14], it was validated empirically

that in the 928 – 948 MHz paging band, spectrum occupancy of the primary user follows a Markovian pattern. In [15], a standard HMM with a finite number of states and a finite output alphabet was used. The adequacy of such a model was questioned in [17]. In a recent paper, optimality of a log-likelihood detector based on a standard HMM was studied [13]. A review of HMMs can be found in [18].

The hidden bivariate Markov model proposed here may be viewed as an instance of a hidden semi-Markov model. In the latter model, the dwell time of the process in each state is not necessarily geometrically distributed. Ferguson [19] proposed to supplement the standard HMM with an explicit durational distribution whose parameter is estimated along with the parameter of the HMM. An overview of hidden semi-Markov models, including HMMs with explicit durational models, can be found in [20]. An advantage of the model proposed here is that estimation of its parameter is done through the Baum algorithm, which is significantly simpler than an algorithm for estimating an explicit durational model. The explicit durational model allows more control over the exact form of the durational distribution, but this advantage is often hard to exploit since the desired form of this distribution is not typically known a priori in applications.

The remainder of the paper is organized as follows. An overview of the proposed system model for spectrum sensing is given in Section II. In Section III, we describe the HBMM and develop some of its key properties. In Section IV, we present the Baum algorithm for estimating the parameter of the HBMM and develop forward recursions for state estimation and prediction. In Section V, we give some numerical examples to evaluate the proposed model in spectrum sensing applications using data derived from real spectrum measurements. In Section VI, we provide concluding remarks.

II. SYSTEM MODEL

In this section, we discuss the application of the proposed HBMM to temporal spectrum sensing.

A. Hidden bivariate Markov model

We denote a discrete-time bivariate Markov chain by $\{Z_t = (X_t, S_t), t = 0, 1, \dots\}$. The process $\{X_t\}$ is referred to as the state process, while $\{S_t\}$ is the underlying process. The value of $\{X_t\}$ at time t represents either the idle ($X_t = 1$) or active ($X_t = 2$) state of the primary transmitter. The role of the underlying process $\{S_t\}$ is to induce a phase-type distribution on the dwell times of $\{X_t\}$ in the idle or active states. The number of states, r , that $\{S_t\}$ may take on corresponds to the number of phases in the phase-type distribution. We refer to r as the order of the bivariate Markov chain. The bivariate Markov chain is discussed more formally in Section III-A.

We assume that the wireless propagation environment is characterized by a standard path loss with lognormal shadowing model. For a receiver at a distance δ from the primary transmitter, the overall log-distance path loss with shadowing, measured in dB, is given by [21, pp. 40-41],

$$L_p(\delta) = \left[\bar{L}_p(\delta_0) + 10\kappa \log_{10} \left(\frac{\delta}{\delta_0} \right) \right] + \epsilon_{(dB)}, \quad \delta \geq \delta_0, \quad (1)$$

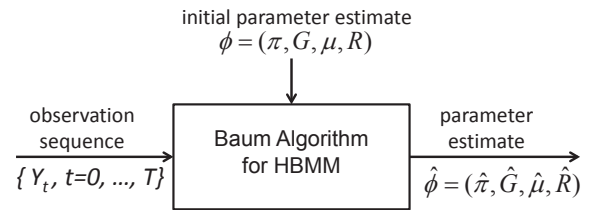


Fig. 1. Block diagram of parameter estimation procedure.

where δ_0 denotes the close-in reference distance, $\bar{L}_p(\delta_0)$ denotes the average log-distance path loss at the reference distance δ_0 , κ denotes the path loss exponent, and $\epsilon_{(dB)}$ represents random shadowing effects. The average log-distance path loss $\bar{L}_p(\delta_0)$ is typically calculated using a free-space path loss formula or through field measurements. The path loss exponent κ varies for different propagation environments. The random variable $\epsilon_{(dB)}$ is assumed to be normal with zero mean and variance σ_e^2 . We have ignored fast fading, since it can be reduced effectively by an averaging filter (cf. [22]). Based on the propagation model (1), the received signal strength Y_t , in units of dBm, is conditionally Gaussian given the state X_t . Hence, the observable process $\{Y_t\}$ is obtained by observing the state process $\{X_t\}$ through a Gaussian memoryless channel. This results in the proposed HBMM.

B. Parameter Estimation

The HBMM parameter, denoted by $\phi = (\pi, G, \mu, R)$, consists of an initial probability distribution π , a transition matrix G , a vector of mean observed signal strengths μ , and a vector of observed signal strength variances R . An appropriate model order can be determined by choosing a value of r that yields the best fit of the data in terms of the state dwell time distributions. Higher values of r generally allow more accurate modeling of the dwell-time distributions, but also require more observation data to avoid overfitting of the model. The impact of the choice of r is studied numerically in Section V.

In Section IV-B, we apply the Baum algorithm to estimate the HBMM parameter from real data in the maximum likelihood sense. Given an initial parameter estimate ϕ and a sequence of signal strength observations $\{Y_t, t = 0, \dots, T\}$, a parameter estimate $\hat{\phi}$ with higher likelihood is computed. The procedure is depicted in Fig. 1. The parameter estimate $\hat{\phi}$ could then subsequently be used to generate simulated data for performance evaluation of cognitive radio systems. Of the components in the parameter estimate $\hat{\phi}$, the most critical is the estimate \hat{G} of the transition matrix, which determines the dynamics of the primary transmitter state process. Different propagation models and shadowing variance can be represented by adjusting $\hat{\mu}$ and \hat{R} as appropriate, while retaining the same \hat{G} . In Section V, we present some numerical results using this approach.

C. Spectrum Sensing

In the spectrum sensing application, time is divided into slots. Each time slot consists of a sensing interval followed by a (longer) transmission interval. During the sensing interval, spectrum measurements are collected and processed by the

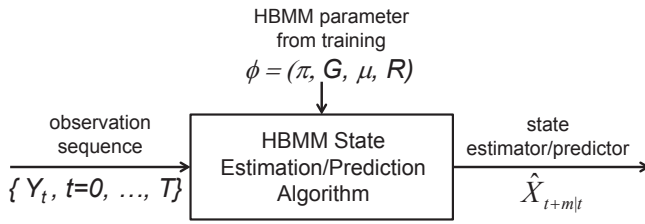


Fig. 2. Block diagram of state estimation and prediction procedure.

cognitive radio, and then a decision is made regarding the state of the primary transmitter on the channel. If the channel is detected as being idle, the cognitive radio may transmit during the transmission window. Otherwise, the cognitive radio remains silent during the transmission window. In Section IV-C, we develop a forward recursion to compute an estimate $\hat{X}_{t+m|t} \in \{1, 2\}$, where $m \geq 0$, of the primary transmitter state at time $t + m$ given the observations up to and including time t . Our approach to spectrum sensing of a cognitive radio channel consists of estimating the parameter of an HBMM from training data using the procedure in Fig. 1, and then applying this parameter into the recursion for state estimation and prediction from any given test sequence. Since spectrum sensing is performed online, it is important that the state estimation and prediction procedure be computationally efficient. For this reason, smaller values of the HBMM model order r are preferred. The procedure is illustrated in Fig. 2.

In practical hardware implementations, there exists a latency between the time the channel is sensed and the time when the cognitive radio can begin or stop its data transmission on the channel, according as the channel is sensed as idle or busy at the current time [23], [24]. With the proposed m -step predictive detector, a cognitive radio can begin vacating the channel several time slots in advance of the actual arrival of the primary user and thus avoid collision with the primary user. The predictive capability of the proposed detector also allows the cognitive radio to prepare to access the channel in advance of the departure of the primary user from the channel, resulting in higher utilization of the temporal spectrum holes. In a system with multiple channels, the predictive information provided by the HBMM-based detector could be used in a channel selection scheme to minimize the overhead of switching among different channels.

III. HIDDEN BIVARIATE MARKOV MODEL

In this section, we describe the HBMM and develop some of its properties. We use capital letters to denote random variables and lower case letters to denote their realizations. We use the notation v_l^k to denote a sequence $\{v_l, v_{l+1}, \dots, v_k\}$. We also use the generic notation of $P(\cdot)$ for probability measure and $p(\cdot)$ for a density or probability mass function (pmf) as appropriate.

A. Bivariate Markov Chain

Let $\{Z_t = (X_t, S_t), t = 0, 1, \dots\}$ denote a discrete-time, finite-state, homogeneous bivariate Markov chain. We assume without loss of generality that the state process $\{X_t\}$ takes values in $\mathbb{X} = \{1, \dots, d\}$, and that the underlying process

$\{S_t\}$ takes values in $\mathbb{S} = \{1, \dots, r\}$. In the spectrum sensing application considered in this paper, we always have $d = 2$, i.e., there are two states representing the active and idle states of the primary transmitter. The bivariate Markov chain $\{Z_t\}$ takes values in $\mathbb{Z} = \mathbb{X} \times \mathbb{S}$. The state pairs $(a, i) \in \mathbb{Z}$ are ordered lexicographically. Let $G = \{g_{ab}(ij)\}$ denote the transition matrix of the bivariate Markov chain, where

$$g_{ab}(ij) = P(Z_{t+1} = (b, j) | Z_t = (a, i)). \quad (2)$$

The transition matrix G is conveniently written as a block matrix $G = \{G_{ab}, a, b \in \mathbb{X}\}$, where $G_{ab} = \{g_{ab}(ij), i, j \in \mathbb{S}\}$ is an $r \times r$ matrix. We assume that G and $\{G_{aa}, a \in \mathbb{X}\}$ are irreducible. The underlying process $\{S_t\}$ is Markov with transition matrix Q if and only if $\sum_{b \in \mathbb{X}} G_{ab} = Q$ independently of a . A similar condition can be given for the state process $\{X_t\}$ to be Markov. Let I denote the $r \times r$ identity matrix. From the Perron-Frobenius theorem and Exercise 7 in [25, p. 536-537], the spectral radius of G_{aa} , denoted by λ_{aa} , satisfies $0 < \lambda_{aa} < 1$. From [25, p. 531, Theorem 2], for each $a \in \mathbb{X}$, the matrix $I - G_{aa}$ is an M -matrix, i.e., it is monotonic (nonsingular with nonnegative inverse), and has nonpositive off-diagonal elements. Since G is irreducible, it has a unique stationary distribution π satisfying $\pi = \pi G$ [26, Theorem 6.9.21]. The process $\{Z_t\}$ is stationary if and only if $P(Z_0 = (a, i)) = \pi_{a,i}$.

The probability distribution of the state process of the bivariate Markov chain, $\{X_0, \dots, X_T\}$, can be obtained as follows. Suppose that the state process jumps at times $T_0 < T_1 < T_2 < \dots < T_N$ where $T_0 = 0$. For $n = 0, \dots, N$, define the sampled state process $\{\tilde{X}_n = X_{T_n}\}$, the sampled underlying chain $\{\tilde{S}_n = S_{T_n}\}$, and the sampled bivariate Markov chain $\{\tilde{Z}_n = Z_{T_n}\}$. For $n = 1, \dots, N$, let $\Delta T_n = T_n - T_{n-1}$ denote the dwell time of the state process $\{X_t\}$ in state \tilde{X}_{n-1} . Note that $\{X_t\}$ can be obtained from $\{\tilde{X}_n, \Delta T_n\}$ and vice versa. Hence, when $T_N = T$, the probability distribution of $\{X_t, t = 0, \dots, T\}$ can be obtained from that of $\{X_0, T_1, \tilde{X}_1, \dots, T_N, \tilde{X}_N\}$. By the Markov property of $\{\tilde{Z}_n\}$, we have

$$\begin{aligned} P(\tilde{Z}_{n+1} = z_{n+1}, \Delta T_{n+1} = \tau_{n+1} | \\ \tilde{Z}_n = z_n, \dots, \tilde{Z}_0 = z_0; T_n = t_n, \dots, T_0 = t_0) \\ = P(\tilde{Z}_{n+1} = z_{n+1}, \Delta T_{n+1} = \tau_{n+1} | \tilde{Z}_n = z_n) \end{aligned} \quad (3)$$

for any positive integer τ_{n+1} . Hence, $\{(\tilde{Z}_n, T_n)\}$ is a discrete-time Markov renewal process. The pmf of $\{(\Delta T_n, \tilde{Z}_n), n \geq 0\}$, and hence of $\{(\Delta T_n, \tilde{X}_n), n \geq 0\}$, can thus be obtained from products of the transition probabilities $\{P(\tilde{Z}_{n+1} = z_{n+1}, \Delta T_{n+1} = \tau_{n+1} | \tilde{Z}_n = z_n)\}$ for $n = 0, \dots, N-1$, and the initial distribution of $\tilde{Z}_0 = Z_0$. Let $f^{ab}(\tau) = \{f_{ij}^{ab}(\tau), i, j = 1, \dots, r\}$ denote an $r \times r$ matrix with its (i, j) element given by

$$f_{ij}^{ab}(\tau) = P(\tilde{Z}_1 = (b, j), \Delta T_1 = \tau | \tilde{Z}_0 = (a, i))$$

$$\begin{aligned}
 &= \sum_{k=1}^r P(Z_\tau = (b, j), X_2^{\tau-1} = \{a\}^{\tau-2}, \\
 &\quad Z_1 = (a, k) \mid Z_0 = (a, i)) \\
 &= \sum_{k=1}^r P(Z_\tau = (b, j), \Delta T_2 = \tau - 1 \mid Z_1 = (a, k)) \\
 &\quad \cdot P(Z_1 = (a, k) \mid Z_0 = (a, i)). \quad (4)
 \end{aligned}$$

It follows, by repeated applications of the argument in (4), that

$$f^{ab}(\tau) = G_{aa}^{\tau-1} G_{ab}. \quad (5)$$

Let $V_n = (\Delta T_n, \tilde{X}_n)$ denote the n th observation of the sampled bivariate Markov chain, and let $\nu_{x_0} = \{P(Z_0 = (x_0, i), i = 1, \dots, r)\}$. Then, the pmf of the state process of the bivariate Markov chain is given by

$$p(x_0, v_1, \dots, v_N) = \nu_{x_0} \prod_{n=1}^N f^{\tilde{x}_{n-1} \tilde{x}_n}(\tau_n) \mathbf{1} \quad (6)$$

where $\mathbf{1}$ denotes a column vector of all ones, and we have used the fact that $\tilde{Z}_0 = Z_0$.

Let $\tilde{G} = \{\tilde{G}_{ab}, a, b \in \mathbb{X}\}$ denote the transition matrix of the sampled bivariate Markov chain $\{\tilde{Z}_n\}$. By summing $f^{ab}(\tau)$ in (5) over all positive integers τ we obtain

$$\tilde{G}_{ab} = (I - G_{aa})^{-1} G_{ab}, \quad a \neq b. \quad (7)$$

By definition, \tilde{G}_{aa} is a zero matrix. Suppose that $g_{aa}(i, i) > 0$ for all $(a, i) \in \mathbb{Z}$. Then, we can write $G_{aa} = D + G_{aa}^*$ where D is a diagonal matrix with positive diagonal elements, and G_{aa}^* is irreducible. Let $\mathbf{0}$ denote the $r \times r$ zero matrix. From [25, p. 533, Exercise 1], $G_{aa}^{r-1} > \mathbf{0}$, and hence

$$(I - G_{aa})^{-1} = \sum_{n=0}^{\infty} G_{aa}^n > \mathbf{0}. \quad (8)$$

Arguing as in [27, Lemma 3], it can be shown that \tilde{G} has a single closed set of recurrent states which may be periodic. Thus, \tilde{G} has a unique stationary distribution ν with zero entries for the transient states. This distribution satisfies $\nu = \nu \tilde{G}$. The distribution ν may be written more explicitly as follows. For each $(a, i) \in \mathbb{Z}$, let $\nu_{a,i} = P(Z_0 = (a, i))$, and let $\nu_a = (\nu_{a,1}, \nu_{a,2}, \dots, \nu_{a,r})$. Then, $\nu = (\nu_1, \dots, \nu_d)$. When ν_{x_0} in (6) is consistent with the stationary distribution ν , then the sampled bivariate Markov chain is stationary with pmf given by (6). Note that we have introduced two stationary distributions, π for $\{Z_t\}$ and ν for $\{\tilde{Z}_n\}$. The two stationary distributions are related by $\nu \propto \pi \cdot \text{diag}\{(I - G_{11}), \dots, (I - G_{dd})\}$ (cf. [28]).

Next, we obtain an expression for the dwell time distribution of the state process $\{X_t\}$ in a given state. For state $a \in \mathbb{X}$, let $u_a(n) = \{P(\tilde{S}_n = i \mid \tilde{X}_n = a), i = 1, \dots, r\}$ and $w_a = \sum_{b:b \neq a} G_{ab} \mathbf{1}$. Using (5), we have

$$\begin{aligned}
 &P(\Delta T_n = m \mid \tilde{X}_{n-1} = a) = \\
 &\quad \sum_{b \neq a, i, j} P(\Delta T_n = m, \tilde{Z}_n = (b, j) \mid \tilde{Z}_{n-1} = (a, i)) \\
 &\quad \cdot P(\tilde{S}_{n-1} = i \mid \tilde{X}_{n-1} = a) = u_a(n-1) G_{aa}^{m-1} w_a, \quad (9)
 \end{aligned}$$

for $m \geq 1; n = 1, 2, \dots$. This is a discrete phase-type distribution with r phases and parameter $(u_a(n-1), G_{aa})$.

The phase-type distribution can, in principle, approximate a given nonnegative distribution to within any desired level of accuracy (cf. [29, pp. 47-50]).

B. Conditional Gaussian Observation Model

The bivariate Markov chain $\{Z_t\}$ is observed through a Gaussian memoryless channel. Let $\{Y_t\}$ denote the output process of the channel. The random variables $\{Y_t, t = 0, 1, \dots\}$ are conditionally independent given $\{X_t, t = 0, 1, \dots\}$, i.e., for any non-negative integer T ,

$$p(y_0^T \mid x_0^T) = \prod_{t=0}^T p(y_t \mid x_t). \quad (10)$$

Furthermore, for any t , $p(y_t \mid x_t)$ is normally distributed with mean μ_{x_t} and covariance matrix R_{x_t} . When Y_t is a scalar random variable, then R_{x_t} is referred to as $\sigma_{x_t}^2$ which is simply the conditional variance of Y_t given x_t . Note that we conditioned Y_t on X_t rather than on Z_t , since the latter conditioning would have resulted in a standard HMM. Let $\mu = \{\mu_a, a = 1, \dots, d\}$ and $R = \{R_a, a = 1, \dots, d\}$. The usual parameter of the proposed model is $\phi = (\pi, G, \mu, R)$. In this parameter, π denotes any distribution of Z_0 , and not necessarily the stationary distribution, which is hard to estimate. Nevertheless, it is well known that the transitional effects due to the use of this distribution die out with time.

IV. PARAMETER AND STATE ESTIMATION ALGORITHMS

In this section, we detail the forward-backward recursions and the Baum algorithm for estimating the parameter of the HBMM. We also develop forward recursions for estimating and predicting the state process.

A. Forward-Backward Recursions

For $t \in \{0, \dots, T\}$, let $\alpha(z_t, y_0^t)$ denote the density of (Z_t, Y_0^t) . Define the $1 \times rd$ vector of such densities as $\alpha_t = \{\alpha((a, 1), y_0^t), \dots, \alpha((a, r), y_0^t), a = 1, \dots, d\}$. Define an $rd \times rd$ block diagonal matrix $B(y_t)$, with its diagonal blocks given by $\{p(y_t \mid X_t = a)I, a \in \mathbb{X}\}$. Then the forward recursion is given by

$$\alpha_0 = \pi B(y_0); \quad \alpha_t = \alpha_{t-1} G B(y_t), \quad t = 1, \dots, T. \quad (11)$$

Similarly, let $\beta(y_{t+1}^T \mid z_t)$ denote the density of y_{t+1}^T given $Z_t = z_t$. Define the $1 \times rd$ vector of such densities as $\beta_t = \{\beta(y_{t+1}^T \mid (a, 1)), \dots, \beta(y_{t+1}^T \mid (a, r)), a = 1, \dots, d\}$. Then the backward recursion is given by

$$\beta_T = \mathbf{1}'; \quad \beta_t = \beta_{t+1} B(y_{t+1}) G', \quad t = T-1, \dots, 0. \quad (12)$$

where $'$ denotes matrix transpose. The likelihood of the observed signal is given by $\alpha_T \mathbf{1}$ or more specifically,

$$p(y_0^T) = \pi B(y_0) \prod_{t=1}^T (G B(y_t)) \mathbf{1}. \quad (13)$$

To ensure the numerical stability of the forward-backward recursions, an embedded scaling procedure is implemented,

see, e.g., [18, Section V.A]. The scaled forward recursion is given by

$$\bar{\alpha}_0 = \frac{\pi B(y_0)}{c_0}, \quad \bar{\alpha}_t = \frac{\bar{\alpha}_{t-1} G B(y_t)}{c_t}, \quad t = 1, \dots, T, \quad (14)$$

where $c_0 = \pi B(y_0)\mathbf{1}$, and $c_t = \bar{\alpha}_{t-1} G B(y_t)\mathbf{1}$ for $t = 1, \dots, T$. Clearly, the scaled forward recursion satisfies $\bar{\alpha}(z_t, y_0^t) = p(z_t | y_0^t)$. The scaled backward recursion is given by

$$\bar{\beta}_T = \mathbf{1}'; \quad \bar{\beta}_t = \frac{\bar{\beta}_{t+1} B(y_{t+1}) G'}{c_t}, \quad t = T-1, \dots, 0. \quad (15)$$

The computational complexity of the forward and backward recursions is $O(d^2 r^2)$, or $O(r^2)$ when $d = 2$, for each step. The scaled and unscaled forward vectors are related by $\bar{\alpha}_t = \alpha_t / \prod_{k=0}^t c_k$. The likelihood in (13) can be expressed in terms of the scaling coefficients as follows:

$$p(y_0^T) = \alpha_T \mathbf{1} = \left(\prod_{t=0}^T c_t \right) \bar{\alpha}_T \mathbf{1} = \prod_{t=0}^T c_t. \quad (16)$$

B. Parameter Estimation

Estimation of the parameter of the HBMM can be done by extending the Baum algorithm [9], see, e.g., [18, Section VI.C]. In this section we make explicit the dependence of all probability functions on the parameter of the HBMM. For a given parameter value, say ϕ , a new parameter value, say $\hat{\phi}$, is obtained from re-estimation formulas. These formulas rely on the conditional probabilities $p(z_{t-1}, z_t | y_0^T; \phi)$, $t = 1, \dots, T$, which can be efficiently calculated as follows:

$$p(z_{t-1}, z_t | y_0^T; \phi) = \frac{\bar{\alpha}(z_{t-1}, y_0^{t-1}) \bar{\beta}(y_{t+1}^T | z_t) [G]_{z_{t-1}, z_t} p(y_t | x_t)}{\sum_{z_{t-1}, z_t} \bar{\alpha}(z_{t-1}, y_0^{t-1}) \bar{\beta}(y_{t+1}^T | z_t) [G]_{z_{t-1}, z_t} p(y_t | x_t)}, \quad (17)$$

where $[G]_{z_{t-1}, z_t}$ denotes the (z_{t-1}, z_t) entry of the transition matrix G . In terms of the conditional probabilities in (17), the re-estimation formulas are given by

$$\hat{\pi}_{a,i} = p(z_0 = (a, i) | y_0^T; \phi), \quad (18)$$

$$\hat{g}_{ab}(ij) = \frac{\sum_{t=1}^T p(z_{t-1} = (a, i), z_t = (b, j) | y_0^T; \phi)}{\sum_{(b,j) \in \mathbb{Z}} \sum_{t=1}^T p(z_{t-1} = (a, i), z_t = (b, j) | y_0^T; \phi)}. \quad (19)$$

When the bivariate Markov chain is observed through a noisy channel with conditional density $p(y_t | x_t) = \mathcal{N}(\mu_{x_t}, \sigma_{x_t}^2)$, then the estimate of $\{(\mu_a, \sigma_a^2), a \in \mathbb{X}\}$ is given by

$$\hat{\mu}_a = \frac{\sum_{t=0}^T \sum_{i=1}^r p(z_t = (a, i) | y_0^T; \phi) y_t}{\sum_{t=0}^T \sum_{i=1}^r p(z_t = (a, i) | y_0^T; \phi)}, \quad (20)$$

$$\hat{\sigma}_a^2 = \frac{\sum_{t=0}^T \sum_{i=1}^r p(z_t = (a, i) | y_0^T; \phi) (y_t - \hat{\mu}_a)^2}{\sum_{t=0}^T \sum_{i=1}^r p(z_t = (a, i) | y_0^T; \phi)}. \quad (21)$$

C. State Estimation and Prediction

Suppose that the parameter of the HBMM is ϕ and that $\{X_t\}$ has two states, i.e., $d = 2$, such that $X_t = 1$ represents the idle state of the primary user while $X_t = 2$ represents its active state. The parameter ϕ is either given or is a parameter estimate. It is easy to see that

$$p(z_{t+m} | y_0^t; \phi) = \sum_{z_t} p(z_t | y_0^t; \phi) p(z_{t+m} | z_t; \phi) = \sum_{z_t} \bar{\alpha}(z_t, y_0^t) [G^m]_{z_t, z_{t+m}}, \quad (22)$$

for $m \geq 0$ and $t \geq 0$. As mentioned above, the complexity of the forward recursion for computing $\bar{\alpha}(z_t, y_0^t)$ is $O(d^2 r^2)$ per step. Since G^m can be pre-computed, the computational complexity of the forward recursion (22) is also $O(d^2 r^2)$, or $O(r^2)$ when $d = 2$. A detection scheme for the state process of the bivariate Markov chain at time $t+m$, given y_0^t , is obtained from

$$\hat{X}_{t+m|t} = \begin{cases} 1, & \text{if } p(x_{t+m} = 1 | y_0^t; \phi) \geq \gamma, \\ 2, & \text{otherwise,} \end{cases} \quad (23)$$

where using (22) we can calculate

$$p(x_{t+m} = 1 | y_0^t; \phi) = \sum_{s_{t+m}} p(z_{t+m} = (1, s_{t+m}) | y_0^t; \phi), \quad (24)$$

and $0 < \gamma < 1$ is a decision threshold. The computational complexity of the detector given by (23) consists of $O(r^2)$ multiplications. When $\gamma = .5$, this detector implements the maximum a-posteriori (MAP) decision rule for testing whether $X_{t+m} = 1$ or $X_{t+m} = 2$ given y_0^t and the parameter ϕ . By varying the value of γ we obtain a receiver operating characteristic (ROC) curve as shown in the next section. The detection scheme (23) can be couched in terms of the log-likelihood ratio (LLR), as was done in [13] for a detector based on the standard HMM, by setting $\gamma = e^{\theta_{\text{LLR}}} / (1 + e^{\theta_{\text{LLR}}})$, where θ_{LLR} is the threshold against which the LLR is compared.

V. NUMERICAL RESULTS

In this section, we present some numerical examples based on real spectrum measurement data to demonstrate the application of the HBMM to spectrum sensing.

A. Spectrum Measurement Data

In selecting the real data for this work, we have examined spectrum occupancy measurements collected by Shared Spectrum Company during the first week of September 2009 [7]. The data was collected using an antenna located on the rooftop of a building in Vienna, Virginia. The elevation of the antenna was 28.96 meters, with latitude of 38.9260 degrees, and longitude of -77.2456 degrees. The measurements were collected over a spectrum band from 30 MHz to 3 GHz once every 137.83 seconds for a duration of 86.835 hours. This fixes the time slot in our proposed model at 137.83 seconds. The number of data samples for each frequency bin was approximately 2268. Out of 32 frequency bands collected, we have considered measurements in one spectrum band with bandwidth ranging from 928 MHz to 1 GHz. There were 501

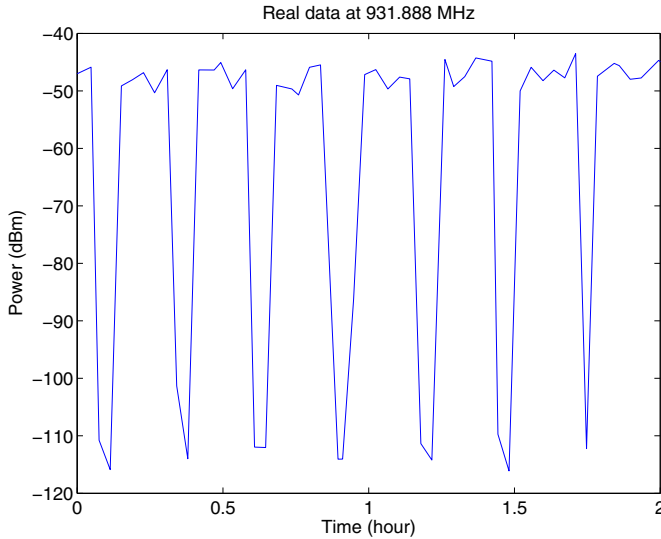


Fig. 3. 2-hour snapshot of spectrum measurement data from paging band.

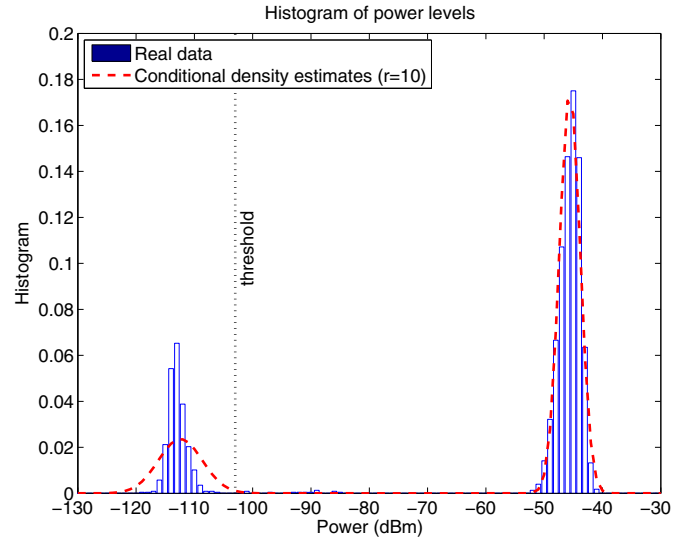


Fig. 4. Histogram of power levels from paging band data.

frequency bins collected in this band, which resulted in a frequency resolution of 144 kHz. The integration time for the data samples was 160 ms. Further details on the equipment setup and measurements can be found in [7].

Within the bandwidth of 928 MHz to 1 GHz, we have selected the spectrum measurements in a paging band with a center frequency of 931.888 MHz. Since the frequency resolution is 144 kHz, the selected spectrum measurements contain all signals that have frequencies in the range from 931.816 MHz to 931.960 MHz. The total time for collecting and processing the data samples in the paging band of interest was about 5 s. In addition, a relatively small amount of additional time would be needed to compute the sensing decision $\hat{X}_{t+m|t}$ for the time slot via (23). In the spectrum sensing application, the remainder of the 137.83 s time slot would be allocated for the transmission interval. From the paging database [30], we found only one registered paging tower in the vicinity with an assigned frequency between 931.816 MHz and 931.960 MHz. This paging tower has a call sign KNKI478 and is located in McLean, Virginia. The elevation of the tower's antenna is 57.6 meters, with coordinates 38.9223 degrees latitude and -77.2289 degrees longitude [31]. The assigned frequency for this tower is 931.9375 MHz and the channel bandwidth is 20 kHz. The maximum effective radiated power (ERP) of the transmitter was $\Gamma = 690$ watts. The radius of the associated macro-cell ranges from 1 to 30 km.

In Fig. 3, a two-hour snapshot of the spectrum measurement data in the time domain is shown. Although the signal may appear to be approximately periodic in this figure, it is not periodic or even quasi-periodic, since the dwell times in the active and idle states are randomly distributed, as evidenced by the empirical dwell time distributions presented in Fig. 5. Fig. 4 shows a histogram of the power levels from the paging band data. Since these signals were recorded using highly elevated antennas for the transmitter and receiver, over a line-of-sight path of 1.5044 km, identifying the idle and active periods is trivial, and can be done using a simple energy detector. The threshold for this detector was set based on

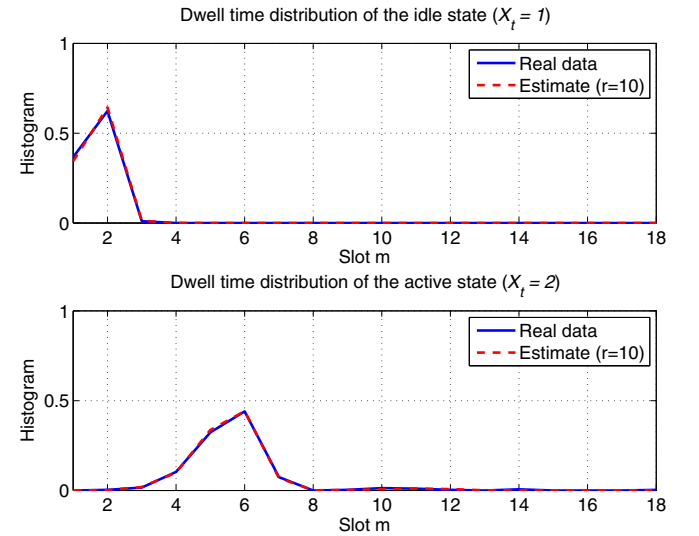


Fig. 5. Dwell time distributions from paging band data.

the noise level of the spectrum band and noise figure of the receiver. This threshold is shown in Fig. 4. Applying an energy detector to the data, we have identified the idle periods and calculated the empirical distributions of the idle and active dwell time periods. These distributions are shown in Fig. 5.

B. Parameter Estimation From Real Data

We applied the Baum algorithm detailed in Section III to the real spectrum measurements discussed above to obtain an estimate of the parameter of an HBMM for various values of the model order r . We used all available samples of the real data, i.e., $T = 2268$ samples, for this estimation. The algorithm was initialized by a randomly chosen parameter ϕ_0 and it was terminated when the relative difference in consecutive log-likelihood values was smaller than 10^{-5} . Using the estimated parameter, we calculated the dwell time distribution (9) for the active and idle states of $\{X_t\}$. We found that the choice of $r = 10$ resulted in a close fit to the

empirical dwell time distribution, as shown in Fig. 5. Note that both dwell time distributions deviate significantly from that of a geometric distribution. For the case $r = 10$, the estimated values of the initial distribution π , and transition matrix G , may be found in [32].

The estimated means and variances of the two conditional Gaussian output distributions in this case were given by

$$\begin{aligned}\mu &= (\mu_1, \mu_2) = (-112.4026, -45.6073), \\ R &= (\sigma_1^2, \sigma_2^2) = (14.2279, 3.1357).\end{aligned}\quad (25)$$

The corresponding conditional Gaussian densities are shown in Fig. 4. Note that histogram in Fig. 4 represents the empirical distribution of the measured power levels irrespective of the primary transmitter state. Thus, the shape of the conditional Gaussian density estimates will not in general match that of the histogram. The larger variance estimate in the idle state is due to the presence of intermittent adjacent channel interference with power in the range $[-100, -85]$ dBm. Since this interference power is closer to the mean power in the idle state than that in the active state, it effectively contributes to the estimate of the conditional density in the idle state, resulting in a larger variance estimate, but has negligible impact on the density estimate in the active state.

The parameter $\phi = (\pi, G, \mu, R)$ was then used to generate simulated data to evaluate the performance of spectrum sensing using the state estimation and prediction algorithms as will be discussed in the next subsection. Although the numerical results presented here are based on data pertaining to a single channel, we have done an extensive study of other channels captured in the spectrum measurement data. The particular channel that we selected required a relatively high model order, i.e., $r = 10$, to achieve a close match of the empirical dwell time distributions. The majority of the other channels required much smaller orders of HBMM for accurate modeling of the dwell time distributions. In this sense, the particular paging channel that was chosen was the most difficult to model accurately and hence may be considered a worst-case representative of all the channels appearing in the spectrum measurement data.

It should be noted that the dwell time distribution depends also on the duration of the time slot, which in this case was 137.83 s, due to the 3 GHz bandwidth over which the original spectrum data was collected. A smaller time slot would generally result in higher correlation between consecutive time slots and consequently dwell time distributions that are very different from the geometric distribution. We also note that when the time slot is smaller, the accuracy of the state predictor $\hat{X}_{t+m|t}$ for values of $m > 1$ becomes more important in practical spectrum sensing scenarios.

C. Spectrum Sensing Performance

Next, we present numerical results to demonstrate the performance of our proposed spectrum sensing approach. To assess the performance of the proposed detector, the true state sequence of the channel must be known. Since this sequence is usually not available for real data, we simulate the observation sequence using the high order ($r = 10$) HBMM parameter $\phi = (\pi, G, \mu, R)$ obtained in the previous

subsection. As noted previously with respect to Fig. 5, the state dwell time distributions corresponding to this parameter, which are determined by the bivariate Markov chain parameter (π, G) , match closely with the empirical distributions. Therefore, the simulated observation sequences generated using this parameter are expected to provide an accurate representation, in a statistical sense, of the actual primary transmitter state process.

We distinguish between two representative scenarios. In the first case, there exists a line-of-sight between the primary user transmitter and the cognitive radio receiver. We refer to this case as spectrum sensing *under no shadowing*, although in reality there is some random noise in the measurements. This situation is essentially the same as the one in which the real spectrum data was obtained. Here, state estimation can be performed accurately with a simple energy detector. On the other hand, the energy detector does not provide predictive information, so we will focus on the m -step prediction performance. In the second scenario, we assume that the cognitive radio receiver is located farther away from the primary transmitter. This results in higher path loss and shadowing effects in the reception of the primary signal compared to the real data measurements under no shadowing case. We refer to this case as spectrum sensing *under shadowing effects*. For this situation, we modified the parameter (μ, R) , of the conditional Gaussian densities to generate new training and testing data. The new training data was used to re-train the low order model, which was then tested on the new testing data. In this case, the performance of the energy detector can be compared meaningfully against the model-based detector for state estimation.

1) *Under No Shadowing*: We used the HBMM parameter ϕ of order $r = 10$ derived from the real spectrum data to generate a simulated observation sequence with $T = 2268$ samples. This observation sequence constituted the training data from which the parameter of a lower order HBMM was estimated using the procedure of Fig. 1. We used three different orders given by $r = 1, 2, 5$. In all cases, the estimated parameter values of the conditional Gaussian densities were very close to the corresponding values obtained for the high-order model in (25) (see [32]).

Next, the parameter ϕ was used to generate a new simulated observation sequence with $T = 2268$ samples. This observation sequence constituted the test data that was used to evaluate the performance of spectrum sensing using the lower order models based on the state detector/predictor of Fig. 2. Fig. 6 shows the probability of error, as a function of the threshold γ in (23), of the one-step prediction using parameter estimates obtained for $d = 2$ and $r = 1, 2, 5$. All values of the decision threshold γ satisfying the inequality $\gamma \leq p(x_{t+m} = 1 | y_0^t; \phi)$ result in the estimate $\hat{X}_{t+m|t} = 1$ and therefore lead to the same prediction error P_e . For a given realization of the processes $\{Y_t\}$ and $\{Z_t\}$, $p(x_{t+m} = 1 | y_0^t; \phi)$ take on a finite set of values, which results in the staircase curves in Fig. 6, i.e., the prediction error performance is constant over certain ranges of the threshold γ . For comparison purposes, the figure also shows the one-step prediction performance obtained when the true parameter ($r = 10$) used to generate the test data was applied in detecting the state sequence from this data.

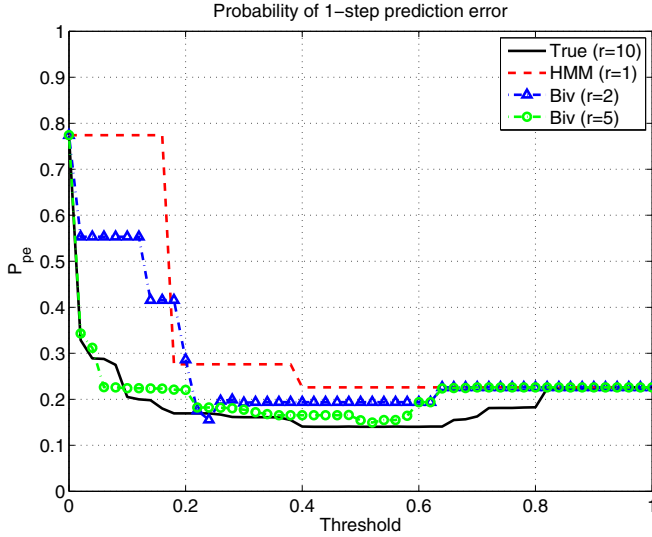


Fig. 6. Under no shadowing: One-step state prediction performance obtained using parameter estimates with $r = 1, 2, 5$ and using the true parameter ($r = 10$).

Practically, this represents the highest achievable performance. We note that there is a significant performance improvement when using the HBMM with $r = 2$ compared to using the standard HMM, which corresponds to $r = 1$. Improvement is also seen when the order r is increased from 2 to 5. When $r = 5$, the performance is nearly as good as in the case when the true parameter is used to detect the state sequence of the primary user.

In Fig. 7, the performance of m -step prediction applying the HBMM parameter estimate with $r = 5$ is shown. The curves show the probability of prediction error vs. the detection threshold γ for $m = 1, 2, 5, 10$. Consider the special case where $\gamma = 0.5$. We observe that the prediction performances for $m = 1$ and $m = 2$ are similar. However, a major performance degradation is observed when $m = 5$, and especially in the case $m = 10$, as would be expected. Although not shown here, the m -step prediction performance when $r = 2$ is significantly worse than the case of $r = 5$ shown in Fig. 7.

In Fig. 8, the dwell time distributions of the state process of the HBMM, as calculated from (9) using the parameter estimates corresponding to models of orders $r = 1, 2, 5$ and the true parameter of order $r = 10$ are compared. Clearly, the geometric dwell time distribution of the HMM does not provide a good approximation. When $r = 2$, the dwell time distribution in the idle state lines up closely with the distribution corresponding to the high order model, but the dwell time distribution in the active state is very far from that of the high order model. When $r = 5$, a good approximation is also obtained for the active state dwell time distribution.

2) *Under Shadowing Effects*: To accommodate the effects of higher path loss and shadowing, we modified parameter of the Gaussian output density, (μ, R) , in the estimated parameter of the high order HBMM, and generated new training and testing sequences that were subsequently used to train and test the low order HBMMs. Only the means and variances of the conditional Gaussian distributions of the high order model

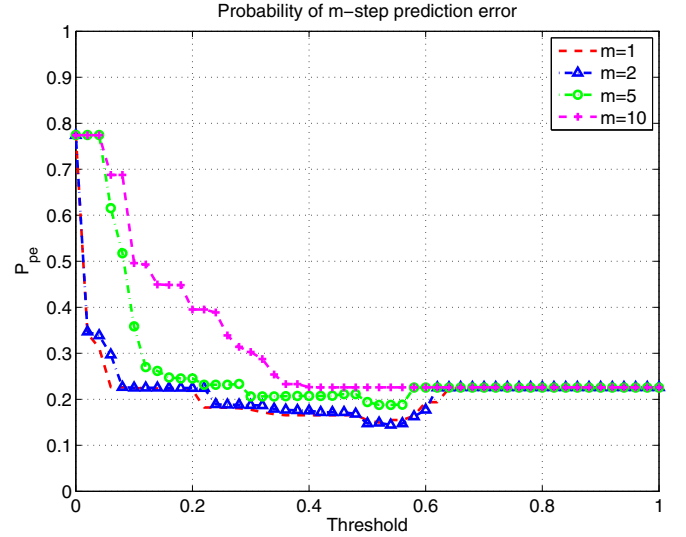


Fig. 7. Under no shadowing: m -step state prediction performance using a parameter estimate with $r = 5$ and $m = 1, 2, 5, 10$.

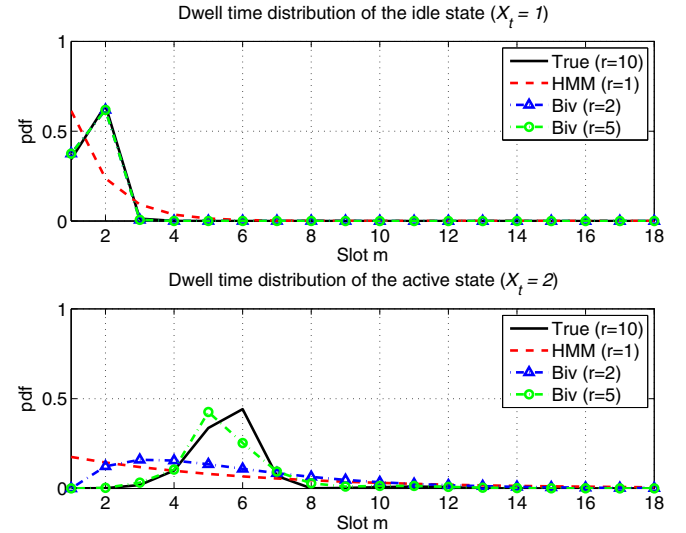


Fig. 8. Under no shadowing: Inferred dwell time distributions obtained using parameter estimates with $r = 1, 2, 5$ and using the true parameter ($r = 10$).

were modified, while the initial distribution and transition matrix, (π, G) , were kept as in Section V-C1.

We applied the propagation model given in (1) with $\delta = 15$ km; $\delta_0 = 1.5044$ km, which was the original distance between the transmitter and receiver under no shadowing case; $\kappa = 5$, which is an appropriate value for the shadowed urban area of McLean, Virginia; and $\sigma_\epsilon^2 = 64$. With these values, $10\kappa \log_{10} \left(\frac{\delta}{\delta_0} \right) = 49.9364$ dB. This loss affects the mean of the received signal Y_t at time t in the active state. Given $\mu_2 = -45.6073$ dBm at distance δ_0 in (25), this mean becomes $\mu_2 - 49.9364 = -95.5437$ dBm at distance δ . Thus, the received signal Y_t in the active state in this case is normally distributed with mean -95.5437 dBm, and variance given by $\sigma_2^2 + \sigma_\epsilon^2 = 3.1357 + 64 = 67.1357$. To summarize, the parameter of the

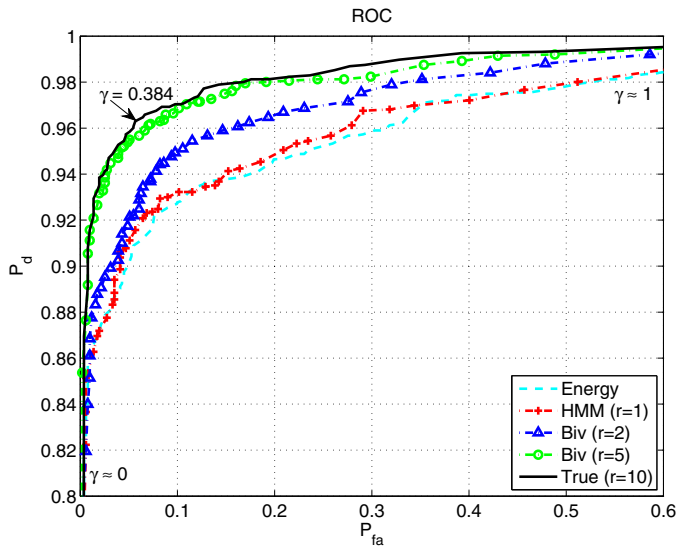


Fig. 9. Under shadowing effects: State estimation performance using parameter estimates with $r = 1, 2, 5$ and using the true parameter ($r = 10$).

conditional Gaussian densities was adjusted from (25) to

$$\mu = \{-112.4026, -95.5437\}, R = \{14.2279, 67.1357\}, \quad (26)$$

for the modified model, while the parameter of the underlying bivariate Markov chain was kept the same.

Under the model (26), Fig. 9 shows ROC curves for state estimation using the energy detector and the HBMM-based detector using parameter estimates with $r = 1, 2$, and 5 . We remark that compared to the energy detector, the HBMM-detector requires some additional overhead to compute the forward recursion for $\hat{X}_{t+m|t}$ (see Section IV-C), but for moderate values of r , this would be insignificant relative to the integration time and other processing overhead required by both detectors. We observe that the HMM-based detector performs only slightly better than the energy detector. When r is increased to 2 , the detection performance improves significantly. A noticeable improvement is also seen in going from $r = 2$ to $r = 5$. The improved performance with increasing value of r can be attributed to more accurate characterization of the dwell time distributions. We also note that the ROC curve for $r = 5$ is very close to the best possible curve obtained by applying the true parameter ($r = 10$) in the detection procedure. Thus, very good spectrum sensing performance can be achieved using a relatively low order HBMM.

Fig. 10 shows the performance of m -step prediction for the HBMM using a parameter estimate with $r = 5$. Similar to the no shadowing case, the prediction performance decreases as the number of steps m increases. Comparing Fig. 7 and Fig. 10, we notice that fading and shadowing effects have slightly degraded the performance of the proposed detection scheme.

VI. CONCLUSION

We proposed an approach to temporal spectrum sensing of a narrowband channel in which the received signal is

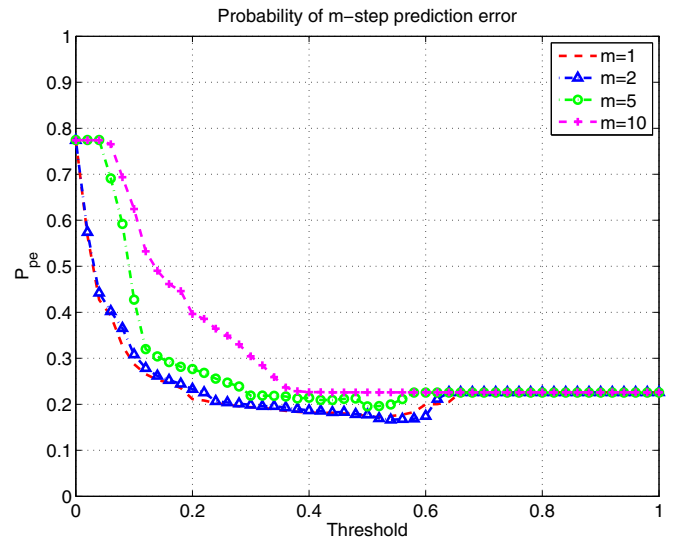


Fig. 10. Under shadowing effects: m -step prediction performance using a parameter estimate with $r = 5$ and $m = 1, 2, 5, 10$.

modeled by a discrete-time bivariate Markov chain observed through a memoryless Gaussian channel. In contrast to a simple energy detector, the detector based on the proposed HBMM can predict the state at a future time instant. Our numerical results based on real spectrum measurement data show that the HBMM-based detector leads to more accurate state estimation and prediction than a detector based on a standard hidden Markov model (HMM) or an energy detector, particularly in scenarios with high path loss and/or strong shadowing effects.

ACKNOWLEDGMENTS

We would like to thank Dr. Mark McHenry for making available the spectrum measurement data from [7]. We would also like to thank the editor and the anonymous reviewers for their comments, which were helpful in improving the presentation of the paper.

REFERENCES

- [1] J. Unnikrishnan and V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 18–27, Feb. 2008.
- [2] B. L. Mark and A. O. Nasif, "Estimation of maximum interference-free transmit power level for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2505–2513, May 2009.
- [3] 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.
- [4] T. Do and B. L. Mark, "Joint spatial-temporal spectrum sensing for cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3480–3490, Sept. 2010.
- [5] E. Axell, G. Leus, E. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: state-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.
- [6] T. Yücek 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.
- [7] Shared Spectrum Company, "General survey of radio frequency bands: 30 MHz to 3 GHz," Tech. Rep., August 2010. Available: <http://www.sharedspectrum.com>
- [8] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: modeling and exploiting white space," *IEEE Commun. Mag.*, pp. 66–72, May 2007.

- [9] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *Ann. Math. Statist.*, vol. 37, pp. 1554–1563, 1966.
- [10] A. Mariani, A. Giorgetti, and M. Chiani, "Effects of noise power estimation on energy detection for cognitive radio applications," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3410–3420, Dec. 2011.
- [11] Z. Sun, G. Bradford, and N. Laneman, "Sequence detection algorithms for PHY-layer sensing in dynamic spectrum access networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 97–109, Feb. 2011.
- [12] T. Nguyen, B. L. Mark, and Y. Ephraim, "Hidden Markov process based dynamic spectrum access for cognitive radio," in *2011 Conf. Inf. Sciences Syst.*, pp. 1–6.
- [13] K. Haghighi, E. G. Ström, and E. Agrell, "On optimum causal cognitive spectrum reutilization strategy," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1911–1921, Nov. 2012.
- [14] C. Ghosh, C. Cordeiro, D. P. Agrawal, and M. B. Rao, "Markov chain existence and hidden Markov models in spectrum sensing," *IEEE Pervasive Comput. Commun.*, pp. 1–6, Mar. 2009.
- [15] I. A. Akbar and W. H. Tranter, "Dynamic spectrum allocation in cognitive radio using hidden Markov models: Poisson distributed case," in *Proc. 2007 IEEE SoutheastCon*, pp. 196–201, Mar. 2007.
- [16] S. Geirhofer, L. Tong, and B. M. Sadler, "A measurement-based model for dynamic spectrum access in WLAN channels," in *Proc. 2006 IEEE Military Commun. Conf.*, pp. 1–7.
- [17] K. W. Sung, S.-L. Kim, and J. Zander, "Temporal spectrum sharing based on primary user activity prediction," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3848–3855, Dec. 2010.
- [18] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518–1569, June 2002.
- [19] J. D. Ferguson, "Variable duration models for speech," in *Proc. 1980 Symp. Application Hidden Markov Models Text Speech*, pp. 143–179.
- [20] S. Z. Yu, "Hidden semi-Markov models," *Artificial Intelligence*, vol. 174, pp. 215–243, 2010.
- [21] J. W. Mark and W. Zhuang, *Wireless Communications and Networking*. Prentice Hall, 2002.
- [22] B. L. Mark and A. E. Leu, "Local averaging for fast handoffs in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 866–874, Mar. 2007.
- [23] T. C. Clancy and B. D. Walker, "Predictive dynamic spectrum access," in *Proc. 2006 SDR Forum Tech. Conf.*
- [24] Z. Chen, N. Guo, Z. Hu, and R. C. Qiu, "Channel state prediction in cognitive radio, part II: single-user prediction," in *Proc. 2011 IEEE SoutheastCon*, pp. 50–54.
- [25] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed. Academic Press, 1985.
- [26] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*. Oxford Science Publications, 2001.
- [27] Y. Ephraim and B. L. Mark, "Consistency of maximum likelihood parameter estimation for bivariate Markov chains," *Stochastic Models*, vol. 29, no. 1, pp. 89–111, 2013.
- [28] B. L. Mark and Y. Ephraim, "An EM algorithm for continuous-time bivariate Markov chains," *Computational Statistics Data Analysis*, vol. 57, no. 1, pp. 504–517, Jan. 2013.
- [29] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA/SIAM Series on Statistics and Applied Probability, 1999.
- [30] City-Data, "Tysons Corner, Virginia." Available: <http://www.city-data.com/city/Tysons-Corner-Virginia.html>
- [31] FCC, "Universal licensing system." Available: <http://wireless2.fcc.gov/UlsApp/UlsSearch/searchLicense.jsp>
- [32] T. Nguyen, "Hidden Markov model based spectrum sensing for cognitive radio," Ph.D. dissertation, George Mason University, Apr. 2013.



Thao Nguyen received her B.S., M.S., and Ph.D. degrees, all in Electrical Engineering, from George Mason University, Fairfax, VA, in 2004, 2006, and 2013, respectively. In 2006, she joined Shared Spectrum Company, Vienna, VA, where she is currently a Senior Communications Systems Engineer. Her main topics of interest are wireless communications and signal processing. Her focus is on research and development of cognitive radio technology.



Brian L. Mark (S'91-M'95-SM'08) received the Ph.D. in Electrical Engineering from Princeton University, in 1995 and the B.A.Sc. degree in Computer Engineering with an option in Mathematics from the University of Waterloo in 1991. He was a Research Staff Member at the C&C Research Laboratories, NEC USA, from 1995 to 1999. In 1999, he was on part-time leave from NEC as a visiting researcher at Ecole Nationale Supérieure des Télécommunications in Paris, France. In 2000, he joined George Mason University, where he is currently Professor of Electrical and Computer Engineering. His research interests lie in the design and performance analysis of communication networks. He served as an associate editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2006-2009.



Yariv Ephraim (S'82-M'84-SM'90-F'94) received the D.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1984. During 1984-1985, he was a Rothschild Postdoctoral Fellow at the Information Systems Laboratory of Stanford University, Stanford, CA. From 1985 to 1993, he was a Member of Technical Staff at the Information Principles Research Laboratory of AT&T Bell Laboratories, Murray Hill, NJ. In 1993, he joined George Mason University, Fairfax, VA, where he currently is Professor of Electrical and Computer Engineering. His research interests are in statistical signal processing.