EM-based Joint Estimation and Detection for Multiple Antenna Cognitive Radios

Ayman Assra and Benoit Champagne
Electrical and Computer Engineering Department, McGill University, Canada
E-mail: ayman.assra@mail.mcgill.ca and benoit.champagne@mcgill.ca

***Abstract*— In this paper, we present an iterative spectrum sensing scheme for multiantenna assisted cognitive radio (CR) using the expectation-maximization (EM) algorithm. Considering a wideband frequency spectrum, the secondary user (SU) performs an EM-based joint estimation and detection (JED), where the channel coefficients and noise variance are estimated jointly with the primary user (PU) signal variance. We also provide a semi-analytical evaluation of the proposed scheme using the Neyman-Pearson criterion. Compared with the conventional Generalized Likelihood Ratio detector (GLRD), the EM-based JED scheme enhances the detection process of the multiple antenna CR with few iterations and modest complexity.**

I. Introduction

In recent years, the cognitive radio (CR) has emerged as a key technology paradigm to alleviate the frequency spectrum scarcity [1]. The basic idea behind CR is that unlicensed or secondary users (SUs) share the frequency spectrum opportunistically with licensed or primary users (PUs) without causing harmful interference. This can be achieved by enabling the SUs to monitor the presence of PUs over a particular spectrum. In the literature, several spectrum sensing techniques have been proposed [2] . Among these, energy detection (ED) is the most common choice for spectrum sensing in CR, especially when there is no *a priori* information about PU signals [3]. However, ED suffers from a poor performance in low signal-to-noise ratio (SNR) environments, resulting from wireless multipath fading channels or shadowing [4]. Multiple antenna techniques have been employed in ED for combatting the fading effects by exploiting the spatial diversity of the observations at the SU [5]. They also require a shorter sensing time compared to single antenna ED. In addition, a main drawback of ED is inherent to its susceptibility to uncertainties of noise variance on the SU side.

Most work conducted on the spectrum sensing problem assume a perfect knowledge of channel conditions and noise variance by the SU, and few researchers have investigated the effect of estimation errors on the detection process and possible estimation techniques [6]. In [5], the authors proposes a Generalized Likelihood Ratio detector (GLRD) for multiple antenna spectrum sensing assuming unknown channel and noise variance. In [7], the authors studies the multi-antenna assisted spectrum sensing assuming a rank-P PU signals with unknown spatial covariance matrix. Recently, there has been a growing interest in iterative joint estimation and detection (JED) techniques because of its ability to achieve accurate estimation without wasting the system resources [8]. In particular, the expectation-maximization (EM) algorithm has been proposed in iterative receivers for its attractive features such as iteratively attaining the maximum likelihood (ML) solution with reduced complexity [9]. In this paper, we introduce an iterative JED scheme based on the EM algorithm for multiple antenna CR over a wideband frequency spectrum. We also present a semi-analytical evaluation for the EM-based JED scheme based on Neyman-Pearson criterion. This is done by first defining the optimum probability of detection of the proposed scheme assuming a perfect knowledge of channel state information (CSI) and noise variance at the SU side. Then, by using the EM estimates of the channel coefficients and noise variance, the optimum probability of detection is averaged over multiple realizations of the system parameters. The rest of the paper is organized as follows. The system model is described in Section II. In Section III, we present the JED scheme based on the EM algorithm. The performance evaluation and the initialization of the proposed scheme are studied in Section IV and Section V respectively. In Section VI, simulation results and discussions are presented . Finally, conclusions are drawn in Section VII.

II. System Model

In our work, we assume that the SU is equipped with P receiving antennas. We also assume a wideband frequency spectrum, which is divided into K frequency subbands. The received samples from different subbands are provided by a filter bank based on a K-point discrete Fourier Transform (DFT) operation. Under these assumptions, the m-th sample of the observed signal by the SU in the k-th subband at the p-th receive antenna is given by

$$R_{k,p}(m) = H_{k,p} S_k(m) + V_{k,p}(m), \tag{1}$$

where $k \in \{0, 1, \ldots, K-1\}$, $m \in \{0, 1, \ldots, M-1\}$, and M is the total number of samples available for detection. Also, $S_k(m)$ represents the m-th sample of the signal transmitted by the PU in the k-th subband, $H_{k,p}$ is the complex channel coefficient between the PU and SU in the k-th subband at the p-th receive antenna, and $V_{k,p}(m)$ is the m-th noise sample in the k-th subband at the p-th receive. In our work, we assume the following statistical model for $\{S_k(m)\}$, $\{H_{k,p}\}$,

This work was supported by InterDigital Canada Ltée., the Natural Sciences and Engineering Research Council of Canada, and the Government of Québec under the PROMPT program.

and $\{V_{k,p}(m)\}$, which is widely adopted in the literature [10]. Since we have no prior knowledge about the PU signal, the entries of $\mathbf{S}_k = [S_k(0), \ldots, S_k(M-1)]^T$, with realizations $\mathbf{s}_k = [s_k(0), \ldots, s_k(M-1)]^T$, where T denotes the transpose operation, is assumed to be independent identically distributed (i.i.d.) complex Gaussian variables with zero mean and covariance $B_k \mathcal{I}_M$, $\mathcal{CN}(\mathbf{0}_M, B_k \mathcal{I}_M)$, where $\mathbf{0}_M$ is a $M \times 1$ zero vector, and $\mathcal{I}_M$ is an identity matrix of a dimension M. We model B_k as a binary random variable, which indicates the status of the PU activity in the k^{th} subband. That is B_k has a zero value when the k-th subband is vacant, while $B_k = 1$ when the PU signal is present in the k-th subband. We also assume independent subband occupancy, where the joint probability mass function of $\{B_k\}$ is given by $P(B_0, \ldots, B_{K-1}) = \prod_{k=0}^{K-1} P(B_k)$. The samples of $\{\mathbf{S}_k\}$ are also assumed to be mutually independent across different subbands. The channel coefficients $\{H_{k,p}\}$, with realizations $\{h_{k,p}\}$, are assumed constant parameters during the sensing interval. Finally, $\mathbf{V}_{k,p} = [V_{k,p}(0), \ldots, V_{k,p}(M-1)]$, are modeled as $\mathcal{CN}(\mathbf{0}_M, \sigma_{k,v}^2 \mathcal{I}_M)$. The process $\{\mathbf{V}_{k,p}\}$ also has independent distribution across the antenna and frequency indices. Based on this model, the conditional mean and variance of $\mathbf{R}_k(m) = [R_{k,0}(m), \ldots, R_{k,P-1}(m)]^T$, with realizations $\mathbf{r}_k(m) = [r_{k,0}(m), \ldots, r_{k,P-1}(m)]^T$, given $B_k = b_k$, where $b_k \in \{0, 1\}$, and $\mathbf{H}_k = [H_{k,0}, \ldots, H_{k,P-1}]^T$, with realizations $\mathbf{h}_k = [h_{k,0}, \ldots, h_{k,P-1}]^T$, are

$$E[\mathbf{R}_k(m)|B_k = b_k, \mathbf{H}_k = \mathbf{h}_k] = \mathbf{0}_P, \tag{2}$$

$$\Sigma_k = Cov[\mathbf{R}_k(m)|B_k = b_k, \mathbf{H}_k = \mathbf{h}_k] = b_k \mathbf{h}_k \mathbf{h}_k^H + \sigma_{k,v}^2 \mathcal{I}_P, \tag{3}$$

where $E[\cdot]$ denotes the expected value, and $Cov[\cdot]$ is the covariance operator.

III. EM-based joint estimation and detection

In this case, the unknown parameter vector $\mathbf{U}$ includes $\mathbf{B}$, $\mathbf{H}$, and $\boldsymbol{\sigma}_v^2$, where $\mathbf{B} = [B_0, \ldots, B_{K-1}]^T$, with realizations $\mathbf{b} = [b_0, \ldots, b_{K-1}]^T$, $\mathbf{H} = [\mathbf{H}_0, \ldots, \mathbf{H}_{K-1}]^T$, with realizations $\mathbf{h} = [\mathbf{h}_0, \ldots, \mathbf{h}_{K-1}]^T$, and $\boldsymbol{\sigma}_v^2 = [\sigma_{0,v}^2, \ldots, \sigma_{K-1,v}^2]^T$. The vector $\mathbf{U}$ is defined as $[\mathbf{U}_0^T, \ldots, \mathbf{U}_{K-1}^T]^T$, where $\mathbf{U}_k = [B_k, \sigma_{k,v}^2, \mathbf{H}_k^T]^T$, with realizations $\mathbf{u} = [\mathbf{u}_0^T, \ldots, \mathbf{u}_{K-1}^T]^T$, where $\mathbf{u}_k = [b_k, \sigma_{k,v}^2, \mathbf{h}_k^T]^T$. This results in a multidimensional optimization problem, where the ML solution requires an exhaustive search over a multidimensional space. In the literature, the EM algorithm is proposed to achieve the ML solution iteratively with low computational complexity. Therefore, in this section, we introduce an iterative joint estimation of the unknown parameters, $\mathbf{U}$, using the EM algorithm. According to the EM terminology, the incomplete data $\mathbf{R} = [\mathbf{R}_0^T, \ldots, \mathbf{R}_{K-1}^T]^T$ are the observations form all receive antennas of the SU over the K subbands, where $\mathbf{R}_k = [\mathbf{R}_k(0)^T, \ldots, \mathbf{R}_k(M-1)^T]^T$, with realizations $\mathbf{r} = [\mathbf{r}_0^T, \ldots, \mathbf{r}_{K-1}^T]^T$, where $\mathbf{r}_k = [\mathbf{r}_k(0)^T, \ldots, \mathbf{r}_k(M-1)^T]^T$. Let $\mathbf{S} = [\mathbf{S}_0^T, \ldots, \mathbf{S}_{K-1}^T]^T$, with realizations $\mathbf{s} = [\mathbf{s}_0^T, \ldots, \mathbf{s}_{K-1}^T]^T$. We define the so-called complete data as $\mathbf{Y} = [(\mathbf{R}_0, \mathbf{S}_0)^T, \ldots, (\mathbf{R}_{K-1}, \mathbf{S}_{K-1})^T]^T$, with realizations $\mathbf{y} = [\mathbf{y}_0^T, \ldots, \mathbf{y}_{K-1}^T]^T$, where $\mathbf{y}_k = [\mathbf{r}_k^T, \mathbf{s}_k^T]^T$. For a given $\mathbf{U} = \mathbf{u}$, the K components of $\mathbf{Y}$ are statistically independent, and consequently the complete data log-likelihood function is defined by

$$L_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u}) = \sum_{k=0}^{K-1} L_{\mathbf{Y}_k|\mathbf{U}_k}(\mathbf{y}_k|\mathbf{u}_k), \tag{4}$$

where

$$\begin{aligned} L_{\mathbf{Y}_k|\mathbf{U}_k}(\mathbf{y}_k|\mathbf{u}_k) &= L_{\mathbf{R}_k|\mathbf{S}_k,\mathbf{U}_k}(\mathbf{r}_k|\mathbf{s}_k, \mathbf{u}_k) + L_{\mathbf{S}_k|\mathbf{U}_k}(\mathbf{s}_k|\mathbf{u}_k) \\ &= -MP\ln(\pi\sigma_{k,v}^2) - M\ln(\pi b_k) - \frac{\sum_{m=0}^{M-1}|s_k(m)|^2}{b_k} \\ &- \frac{\sum_{p=0}^{P-1}\sum_{m=0}^{M-1}|r_{k,p}(m) - h_{k,p}s_k(m)|^2}{\sigma_{k,v}^2}. \end{aligned} \tag{5}$$

In the expectation-step (E-step) of the EM algorithm, we estimate the conditional expectation of (4) given $\mathbf{R} = \mathbf{r}$ and the estimation of $\mathbf{U}$ at the i-th iteration, i.e., $\mathbf{U} = \hat{\mathbf{u}}^{(i)}$, where $\hat{\mathbf{u}}^{(i)} = [\hat{\mathbf{b}}^{(i)^T}, \hat{\boldsymbol{\sigma}}_v^{2^{(i)^T}}, \hat{\mathbf{h}}^{(i)^T}]^T$, and $\hat{\mathbf{b}}^{(i)}$, $\hat{\boldsymbol{\sigma}}_v^{2^{(i)}}$ and $\hat{\mathbf{h}}^{(i)}$ are the EM estimates of $\mathbf{b}$, $\boldsymbol{\sigma}_v^2$ and $\mathbf{h}$ at the i-th iteration respectively. This gives

$$\Delta(\mathbf{u}|\hat{\mathbf{u}}^{(i)}) = \sum_{k=0}^{K-1} \Delta(\mathbf{u}_k|\hat{\mathbf{u}}_k^{(i)}) \tag{6}$$

where

$$\begin{aligned} \Delta(\mathbf{u}_k|\hat{\mathbf{u}}_k^{(i)}) &= -MP\ln(\pi\sigma_{k,v}^2) - M\ln(\pi b_k) \\ &- \frac{\sum_{p=0}^{P-1}\sum_{m=0}^{M-1} E[|r_{k,p}(m) - h_{k,p}s_k(m)|^2|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}]}{\sigma_{k,v}^2} \\ &- \frac{\sum_{m=0}^{M-1} E[|s_k(m)|^2|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}]}{b_k}. \end{aligned} \tag{7}$$

Since $\{S_k(m)\}$ and $\{V_{k,p}(m)\}$ are modeled as mutually independent random processes, and the samples of each process are statistically independent across the frequency and frame indices, the unknown parameters of each subband, $\mathbf{u}_k$, are estimated individually by maximizing the corresponding likelihood function in (6), i.e., $\Delta(\mathbf{u}_k|\hat{\mathbf{u}}_k^{(i)})$. Also, the maximization process of $\Delta(\mathbf{u}_k|\hat{\mathbf{u}}_k^{(i)})$ is performed independently with respect to each of the unknown parameters in $\mathbf{u}_k$ as follows. By taking the derivative of (7) with respect to b_k and solving the resultant equation, we get

$$\hat{b}_k^{(i+1)} = \frac{\sum_{m=0}^{M-1} E[|s_k(m)|^2|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}]}{M}, \tag{8}$$

where $E[|s_k(m)|^2|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}] = |E[s_k(m)|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}]|^2 + Var[s_k(m)|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}]$, and $Var[.]$ is the variance operator. Given that $\mathbf{R}$ and $\mathbf{S}$ are jointly Gaussian, the conditional mean and variance of $S_k(m) = s_k(m)$ given $\mathbf{R} = \mathbf{r}$ and $\mathbf{U} = \hat{\mathbf{u}}^{(i)}$ are given by [11]

$$E[s_k(m)|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}] = \hat{b}_k^{(i)} \hat{\mathbf{h}}_k^{(i)^H} \hat{\Sigma}_k^{(i)^{-1}} \mathbf{r}_k(m), \tag{9}$$

$$Var[s_k(m)|\mathbf{R} = \mathbf{r}, \mathbf{U} = \hat{\mathbf{u}}^{(i)}] = \hat{b}_k^{(i)} - \hat{b}_k^{(i)} \hat{\mathbf{h}}_k^{(i)^H} \hat{\Sigma}_k^{(i)^{-1}} \hat{\mathbf{h}}_k^{(i)} \hat{b}_k^{(i)}, \tag{10}$$

where $\hat{\Sigma}_k^{(i)} = \hat{b}_k^{(i)}\hat{\mathbf{h}}_k^{(i)}\hat{\mathbf{h}}_k^{(i)^H} + \sigma_{k,v}^{2^{(i)}}\mathcal{I}_P$, and H denotes the conjugate transpose operation. Similarly, we get the EM estimates of $\mathbf{h}_k$ and $\sigma_{k,v}^2$ at the $(i+1)$-th iteration as follows. First, $\hat{h}_{k,p}^{(i+1)}$ is obtained by maximizing the corresponding summand in the right-hand expression of (7)

$$\hat{h}_{k,p}^{(i+1)} = \arg\min_{h_{k,p}} \sum_{m=0}^{M-1} E[|r_{k,p}(m) - h_{k,p}s_k(m)|^2|\mathbf{R}=\mathbf{r}, \mathbf{U}=\hat{\mathbf{u}}^{(i)}], \tag{11}$$

which results in

$$\hat{h}_{k,p}^{(i+1)} = \frac{\sum_{m=0}^{M-1} r_{k,p}(m)E[s_k(m)^*|\mathbf{R}=\mathbf{r},\mathbf{U}=\hat{\mathbf{u}}^{(i)}]}{\sum_{m=0}^{M-1} E[|s_k(m)|^2|\mathbf{R}=\mathbf{r},\mathbf{U}=\hat{\mathbf{u}}^{(i)}]}, \tag{12}$$

where x^* denotes the complex conjugate of x. Then, we substitute $\mathbf{h}_k = \hat{\mathbf{h}}_k^{(i+1)}$ in (7) and perform the maximization process with respect to $\sigma_{k,v}^2$. This yields to

$$\hat{\sigma}_{k,v}^{2^{(i+1)}} = \frac{\sum_{p=0}^{P-1}\sum_{m=0}^{M-1} X_{k,p}(m)}{MP} \tag{13}$$

where $X_{k,p}(m) = |r_{k,p}(m)|^2 - r_{k,p}(m)^*\hat{h}_{k,p}^{(i+1)}E[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{U}=\hat{\mathbf{u}}^{(i)}] - r_{k,p}(m)\hat{h}_{k,p}^{(i+1)^*}E[s_k(m)^*|\mathbf{R}=\mathbf{r},\mathbf{U}=\hat{\mathbf{u}}^{(i)}] + |\hat{h}_{k,p}^{(i+1)}|^2 E[|s_k(m)|^2|\mathbf{R}=\mathbf{r},\mathbf{U}=\hat{\mathbf{u}}^{(i)}]$. When achieving a certain performance criteria, i.e., $|\hat{b}_k^{(i+1)} - \hat{b}_k^{(i)}| = \epsilon$, where $\epsilon \to 0$, a hard limiting with threshold γ_k is applied on $\hat{b}_k^{\mathrm{EM}} = \hat{b}_k^{(i+1)}$, which gives $\hat{b}_k^{\mathrm{EM}} \underset{\tilde{b}_k=0}{\overset{\tilde{b}_k=1}{\gtrless}} \gamma_k$. We notice that by using the EM-based JED scheme, the K-subband optimization problem over the wideband spectrum is decomposed into K idependent one subband optimization problems, leading to a computationally feasible scheme. We notice that the estimation of (9) and (10) dominates the computational complexity of the proposed EM-based JED scheme. That is it requires $\mathcal{O}\{KMP^4\}$ mathematical operations per sensing interval.

IV. Performance Evaluation of EM-based JED

The performance of spectrum sensing schemes is evaluated using the Neyman-Pearson criterion, where for a given probability of false alarm in the k-th subband, the optimum threshold in the decision-making process, γ_k^{opt}, and subsequently the optimum probability of detection, $P_{d,k}^{\mathrm{opt}}(\gamma_k^{\mathrm{opt}})$, are provided [11]. This analytical evaluation is not applicable in iterative JED schemes since the derivation of closed-form expressions of the final decision variables, i.e., $\hat{b}_k^{(i)}$ as $i \to \infty$, is not always feasible. Therefore, we propose a semi-analytical evaluation of the proposed EM-based JED scheme, where an analytical expression of γ_k^{opt} and $P_{d,k}^{\mathrm{opt}}(\gamma_k^{\mathrm{opt}})$ are introduced assuming a perfect knowledge of $\boldsymbol{\sigma}_v^2$ and $\mathbf{H}$ by the SU. Then, by substituting $\boldsymbol{\sigma}_v^2 = \hat{\boldsymbol{\sigma}}_v^{2^{(i)}}$ and $\mathbf{H} = \hat{\mathbf{h}}^{(i)}$, the receiver operating characteristic (ROC) of the proposed JED scheme is obtained by averaging $\hat{P}_{d,k}^{\mathrm{opt}^{(i)}}(\hat{\gamma}_k^{\mathrm{opt}^{(i)}})$ over different realizations of the system parameters. In the following we present the EM estimation of $\mathbf{B}$ assuming that $\boldsymbol{\sigma}_v^2$ and $\mathbf{H}$ are perfectly known by the SU. In this case, we get an explicit expression of $\hat{b}_k^{(i)}$ as $i \to \infty$ through deriving the ML solution of the same problem, $\hat{b}_k^{\mathrm{ML}}$. Finally, using $\hat{b}_k^{\mathrm{ML}}$, a closed-form expressions of γ_k^{opt} and $P_{d,k}^{\mathrm{opt}}(\gamma_k^{\mathrm{opt}})$ are derived.

A. *EM-Based Detector with Known $\boldsymbol{\sigma}_v^2$ and $\mathbf{H}$*

In this case, the only unknown parameter is $\mathbf{B}$. Therefore, we denote the EM solution by EM-based spectrum sensing (EM-based SS). Following the same procedure as above, $\hat{b}_k^{(i+1)}$ is reduced to

$$\hat{b}_k^{(i+1)} = \frac{\sum_{m=0}^{M-1} E[|s_k(m)|^2|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}]}{M}, \tag{14}$$

where $E[|s_k(m)|^2|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}] = |E[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}]|^2 + Var[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}]$. The conditional mean and variance of $s_k(m)$ given $\mathbf{R}=\mathbf{r}$ and $\mathbf{B}=\hat{\mathbf{b}}^{(i)}$ are given by [11]

$$E[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}] = \hat{b}_k^{(i)}\mathbf{h}_k^H\Sigma_k^{-1}\mathbf{r}_k(m), \tag{15}$$

and

$$Var[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}] = \hat{b}_k^{(i)} - \hat{b}_k^{(i)}\mathbf{h}_k^H\Sigma_k^{-1}\mathbf{h}_k\hat{b}_k^{(i)}. \tag{16}$$

To justify the convergence of $\hat{b}_k^{(i+1)}$ to $\hat{b}_k^{\mathrm{ML}}$ as $i \to \infty$, we first present the ML estimation of $\mathbf{B}$ or the ML-based SS as follows. The log-likelihood function of $\mathbf{R}$ given $\mathbf{B}=\mathbf{b}$ and $\mathbf{H}=\mathbf{h}$ is

$$L_{\mathbf{R}|\mathbf{B},\mathbf{H}}(\mathbf{r}|\mathbf{b},\mathbf{h}) = \sum_{k=0}^{K-1} L_{\mathbf{R}_k|B_k,\mathbf{H}_k}(\mathbf{r}_k|b_k,\mathbf{h}_k), \tag{17}$$

where

$$L_{\mathbf{R}_k|B_k,\mathbf{H}_k}(\mathbf{r}_k|b_k,\mathbf{h}_k) = -PM\ln(\pi) - M\ln(\det(\Sigma_k)) - \sum_{m=0}^{M-1} tr(\Sigma_k^{-1}\mathbf{r}_k(m)\mathbf{r}_k(m)^H), \tag{18}$$

In (18), $\det(\cdot)$ is the matrix determinant, and $tr(\cdot)$ is the matrix trace. Using the matrix properties, $\det(\Sigma_k)$ given $B_k = b_k$ and $\mathbf{H}_k = \mathbf{h}_k$ is defined by [5]

$$\det(\Sigma_k) = (b_k||\mathbf{h}_k||^2 + \sigma_{k,v}^2)(\sigma_{k,v}^2)^{P-1}, \tag{19}$$

where $||\cdot||$ denotes the vector norm. Also, by using the matrix inversion lemma [12], Σ_k^{-1} conditioned on $B_k = b_k$ and $\mathbf{H}_k = \mathbf{h}_k$ is given by

$$\Sigma_k^{-1} = \frac{\mathcal{I}_P}{\sigma_{k,v^2}} - \frac{\mathbf{h}_k\mathbf{h}_k^H}{\sigma_{k,v}^2\left(\frac{\sigma_{k,v}^2}{b_k} + ||\mathbf{h}_k||^2\right)}. \tag{20}$$

Substituting (19) and (20) in (18), we have

$$L_{\mathbf{R}_k|B_k,\mathbf{H}_k}(\mathbf{r}_k|b_k,\mathbf{h}_k) = -PM\ln(\pi) - M\ln(b_k||\mathbf{h}_k||^2 + \sigma_{k,v}^2) - M(P-1)\ln(\sigma_{k,v}^2) - \frac{\sum_{m=0}^{M-1} tr(\mathbf{r}_k(m)\mathbf{r}_k(m)^H)}{\sigma_{k,v}^2} + \frac{\sum_{m=0}^{M-1}|\mathbf{h}_k^H\mathbf{r}_k(m)|^2}{\sigma_{k,v}^2\left(\frac{\sigma_{k,v}^2}{b_k} + ||\mathbf{h}_k||^2\right)}. \tag{21}$$

Since the subband occupancies, $\{B_k\}$, are assumed to be statistically independent, the maximization process of (17) with respect to $\mathbf{b}$ is done separately for each subband. In this case, $\hat{b}_k^{\mathrm{ML}} = \arg\max_{b_k} L_{\mathbf{R}_k|B_k,\mathbf{H}_k}(\mathbf{r}_k|b_k,\mathbf{h}_k)$, which results in

$$\hat{b}_k^{\mathrm{ML}} = \frac{\sum_{m=0}^{M-1}|\mathbf{h}_k^H\mathbf{r}_k(m)|^2 - M\sigma_{k,v}^2||\mathbf{h}_k||^2}{M||\mathbf{h}_k||^4}. \tag{22}$$

The convergence of $\hat{b}_k^{(i+1)}$ to the ML solution is proven as follows. Using the matrix inversion lemma [12], (15) and (16) are expressed as

$$E[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}] = \frac{\mathbf{h}_k^H\mathbf{r}_k(m)}{\frac{\sigma_{k,v}^2}{b_k}+||\mathbf{h}_k||^2}, \tag{23}$$

$$Var[s_k(m)|\mathbf{R}=\mathbf{r},\mathbf{B}=\hat{\mathbf{b}}^{(i)}] = \frac{\sigma_{k,v}^2}{\frac{\sigma_{k,v}^2}{b_k}+||\mathbf{h}_k||^2}. \tag{24}$$

Substituting (23) and (24) in (14), we get

$$\hat{b}_k^{(i+1)} = \frac{\frac{1}{M}\sum_{m=0}^{M-1}|\mathbf{h}_k^H\mathbf{r}_k(m)|^2 + \frac{\sigma_{k,v}^4}{b_k} + \sigma_{k,v}^2||\mathbf{h}_k||^2}{\left(\frac{\sigma_{k,v}^2}{b_k}+||\mathbf{h}_k||^2\right)^2}. \tag{25}$$

Let $\hat{b}_k^* = \lim_{i\to\infty}\hat{b}_k^{(i)}$. Then, by substituting $b_k = \hat{b}_k^{(i+1)} = \hat{b}_k^*$ in (25) and solving the resultant equation, we get $\hat{b}_k^* = \hat{b}_k^{\mathrm{ML}}$.

B. Performance Evaluation of EM-based SS

The decision statistic of the ML-based SS or the EM-based SS, i.e., assuming $i\to\infty$, is defined by

$$Z_k = \frac{\sum_{m=0}^{M-1}|\mathbf{H}_k^H\mathbf{R}_k(m)|^2 - M\sigma_{k,v}^2||\mathbf{H}_k||^2}{M||\mathbf{H}_k||^4}. \tag{26}$$

Under the hypothesis $B_k = 0$, the distribution of $\mathbf{R}_k(m)$ is $\mathcal{CN}(\mathbf{0}_P, \sigma_{k,v}^2\mathcal{I}_P)$. Therefore, conditioned on $\mathbf{H}_k$, $\mathbf{H}_k^H\mathbf{R}_k(m)$ has a complex Gaussian distribution with zero mean and variance $||\mathbf{H}_k||^2\sigma_{k,v}^2$. Assume $z_{k,0} = \frac{\sum_{m=0}^{M-1}|\mathbf{H}_k^H\mathbf{R}_k(m)|^2}{||\mathbf{H}_k||^2\sigma_{k,v}^2}$, then $z_{k,0}$ has a chi-square distribution with $2M$ degrees of freedom, χ_{2M}^2. Using the closed-form expression of the complementary cumulative distribution function (CCDF) of the chi-square distribution, $P_{f,k}(\gamma_k)$ is given by [5]

$$\begin{aligned}P_{f,k}(\gamma_k) &= P\left(Z_k > \gamma_k|B_k=0\right)\\ &= P\left(z_{k,0} > \frac{M\gamma_k||\mathbf{H}_k||^4 + M\sigma_{k,v}^2||\mathbf{H}_k||^2}{||\mathbf{H}_k||^2\sigma_{k,v}^2}\right)\\ &= \Gamma\left(M, \frac{M\gamma_k||\mathbf{H}_k||^2 + M\sigma_{k,v}^2}{\sigma_{k,v}^2}\right),\end{aligned} \tag{27}$$

where $\Gamma(c,y) = \frac{1}{\Gamma(c)}\int_y^\infty x^{c-1}e^{-x}dx = w$ is the normalized upper incomplete gamma function, and $\Gamma(c) = \int_0^\infty x^{c-1}e^{-x}dx$ represents the complete gamma function. Under the constraint, $P_{f,k}(\gamma_k) = \alpha_k$, the optimum threshold is given by

$$\gamma_k^{\mathrm{opt}} = \frac{\sigma_{k,v}^2\Gamma^{\mathrm{inv}}(M,\alpha_k) - M\sigma_{k,v}^2}{M||\mathbf{H}_k||^2}, \tag{28}$$

where $\Gamma^{\mathrm{inv}}(c,w)$ is the inverse incomplete gamma function of $\Gamma(c,y)$ [13]. Similarly, under the hypothesis $B_k = 1$, $\mathbf{R}_k(m)$ is $\mathcal{CN}(\mathbf{0}_P, \mathbf{H}_k\mathbf{H}_k^H + \sigma_{k,v}^2\mathcal{I}_P)$, and consequently conditioned on $\mathbf{H}_k$, $\mathbf{H}_k^H\mathbf{R}_k(m)$ follows $\mathcal{CN}(0, ||\mathbf{H}_k||^2(||\mathbf{H}_k||^2+\sigma_{k,v}^2))$. In this case, the distribution of $z_{k,1} = \frac{\sum_{m=0}^{M-1}|\mathbf{H}_k^H\mathbf{R}_k(m)|^2}{||\mathbf{H}_k||^2(||\mathbf{H}_k||^2+\sigma_{k,v}^2)}$ is χ_{2M}^2. Therefore, the optimum probability of detection is obtained by

$$\begin{aligned}P_{d,k}^{\mathrm{opt}}(\gamma_k^{\mathrm{opt}}) &= P\left(Z_k > \gamma_k^{\mathrm{opt}}|B_k=1\right)\\ &= P\left(z_{k,1} > \frac{M\gamma_k^{\mathrm{opt}}||\mathbf{H}_k||^2 + M\sigma_{k,v}^2}{||\mathbf{H}_k||^2+\sigma_{k,v}^2}\right)\\ &= \Gamma\left(M, \frac{M(\nu_k\gamma_k^{\mathrm{opt}}+1)}{\nu_k+1}\right),\end{aligned} \tag{29}$$

where $\nu_k = \frac{||\mathbf{H}_k||^2}{\sigma_{k,v}^2}$ is the average received signal to noise ratio at the SU under $B_k = 1$.

V. Initialization

Since the EM algorithm is sensitive to the initialization of the parameters to be estimated [8], we assume that our proposed EM-based JED is initialized by reliable estimates of the unknown parameters. This guarantees that the performance of our proposed detector converges to ML solution with few iterations. In our case, we assume that the SU receives a training sequence of J_h samples from a CR in the vicinity of the PU. Then, the initial value of $h_{k,p}$, $\hat{h}_{k,p}^{(0)}$, is estimated using the least squares (LS) channel estimation technique [11]. Assuming that the SU has *a priori* knowledge about the history of the PU activities, the initialization of $\sigma_{k,v}^2$, $\hat{\sigma}_{k,v}^{2^{(0)}}$, is performed by estimating the sample variance of the observations over J_n-sample period when the PU is absent, i.e., $R_{k,p}(m) = V_{k,p}(m)$. Finally, each b_k is initialized by γ_k, i.e., $\hat{b}_k^{(0)} = \gamma_k$. In the following iterations of the proposed EM-based JED scheme, $\hat{b}_k^{(i)}$, for $i > 0$, takes values greater than γ_k if $b_k = 1$, while it becomes less than γ_k if $b_k = 0$. This process continues until the convergence is achieved, i.e., $|\hat{b}_k^{(i+1)} - \hat{b}_k^{(i)}| = \epsilon$. At this stage, $\hat{b}_k^{(i+1)}$ is compared with γ_k resulting in a hard estimate of b_k as follows: $\hat{b}_k^{(i+1)} \underset{\tilde{b}_k=0}{\overset{\tilde{b}_k=1}{\gtrless}} \gamma_k$.

VI. Simulation Results

The performance of the proposed JED scheme based on the EM algorithm is evaluated through the ROC curve. Throughout our simulations, we assume that the CR has $P = 1, 2$ receiving antennas, operating in a wideband spectrum with $K = 8$ subbands. Since the estimation of the unknown parameters is performed independently for each subband, our results are presented only for one subbande, for example, $k = 2$. The simulations are implemented for $M = 50$, $J_p = J_n = 30$, $\sigma_{2,v}^2 = 1$, SNR=-3dB, and 10^5 trials for each value of γ_2. Also, the performance of the proposed EM-based JED scheme is evaluated after 5 iterations.

Fig. 1 represents the performance of different spectrum sensing schemes for a single antenna CR. As a reference, we plot the ROC curve of the EM-based SS and the optimum ED [11] where $H_{2,1}$ and $\sigma^2_{2,v}$ are perfectly known by the SU user. The results show that the performance of the EM-based SS perfectly matches with that of the optimum ED, proving the optimality of the EM solution. Also, the proposed EM-based JED scheme achieves a performance close to the perfect estimation case, when $H_{2,1}$ and b_2 are the only unknown parameters. We also notice that the proposed scheme is sensitive to the error in the noise variance estimation. As a benchmark, the noise variance estimation error can be quantified through the comparison with the optimum ED under noise variance mismatch of 2dB [4]. The noise variance mismatch is defined by $10\log_{10}(\frac{\tilde{\sigma}^2_{2,v}}{\sigma^2_{2,v}})$, where $\tilde{\sigma}^2_{2,v}$ is the actual noise variance.

Fig. 2 evaluates the performance of the proposed scheme for a CR with P=2 receiving antennas. For the comparison, we include the ROC curve of the GLRD proposed in [5] and the EM-based SS. The results show that the EM-JED scheme provides more reliable estimates for the channel coefficients and noise variance, and consequently enhances the spectrum detection process compared with the GLRD. The error in the channel and noise variance estimation can be quantified through the comparison with the conventional multiple antenna ED under noise variance mismatch of 1.8dB.The semi-analytical evaluation of the proposed JED scheme after 5 iterations is also presented. By targeting specific α_2, $\{\hat{h}^{(5)}_{2,p}\}$ and $\hat{\sigma}^{2^{(5)}}_{2,v}$ are estimated for 10^5 system parameters realizations, i.e., 10^5 trials, from (12) and (13) respectively. Then, these estimates are used to get $\hat{\gamma}^{\text{opt}^{(5)}}_2$ and $\hat{P}^{\text{opt}^{(5)}}_{d,2}(\hat{\gamma}^{\text{opt}^{(5)}}_2)$ of each trial from (28) and (29) respectively. This process is repeated for each value of α_2, where at the end the average values of $\hat{P}^{\text{opt}^{(5)}}_{d,2}(\hat{\gamma}^{\text{opt}^{(5)}}_2)$ over 10^5 trials are plotted against different values of α_2. From the results, we notice the closeness of the semi-analytical solution to the simulation results.

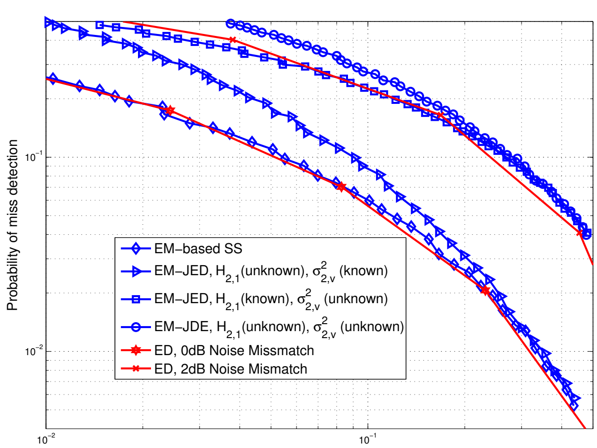

Fig. 1. ROC of EM-based JED scheme for a single antenna CR (M=50, SNR=-3dB, and 5 iterations)

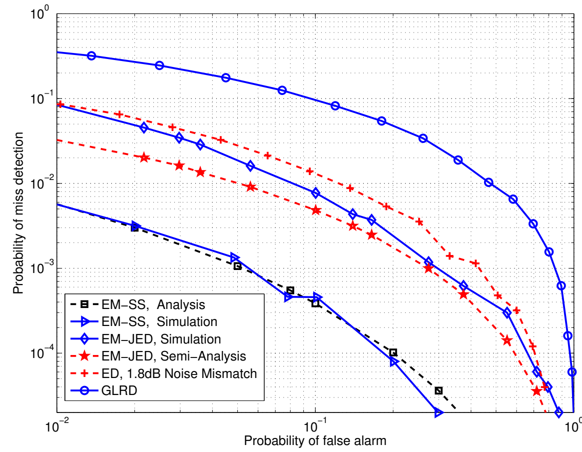

Fig. 2. ROC of EM-based JED scheme for 2-antenna CR (M=50, SNR=-3dB, and 5 iterations)

VII. Conclusion

We developed an EM-based JED scheme for multiple antenna cognitive radios over wideband frequency spectrum. We also introduced a semi-analytical evaluation of the proposed scheme based on the Neyman-Pearson criterion, which was justified through the comparison with the simulation results. Compared with the conventional GLRD, the proposed scheme provided a potential improvement in the PU detection process.

References

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

[2] Z. Zhang, Q. Yang, L. Wang, and X. Zhou, "A novel hybrid matched filter structure for IEEE 802.22 standard", in *Proc. IEEE Asia Pacific Conf. Circuits and Systems*, Dec. 2010, pp. 652–655.

[3] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks", *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1232–1241, Apr. 2011.

[4] A. Pandharipande and J.-P. Linnartz, "Performance analysis of primary user detection in a multiple antenna cognitive radio", in *IEEE Int. Conf. Communications (ICC 07)*, June 2007, pp. 6482–6486.

[5] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios", *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814–823, Feb. 2010.

[6] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio", *IEEE Trans. Vehicular Technology*, vol. 59, no. 4, pp. 1791–1800, May 2010.

[7] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Detection of rank-P signals in cogntive radio networks with uncalibrated multiple antennas", *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.

[8] Q. Guo and D. Huang, "EM-based joint channel estimation and detection for frequency selective channels using gaussian message passing", *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4030–4035, Aug. 2011.

[9] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm", *J. Royal Statist Soc.*, vol. 39, pp. 1–38, Jan. 1977.

[10] K. Hossain and B. Champagne, "Wideband spectrum sensing for cognitive radios with correlated subband occupancy", *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 35 –38, Jan. 2011.

[11] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 2nd edition, 1994.

[12] G. Strang, *Introduction to Linear Algebra*, Cambridge Publication, 3rd edition, 2003.

[13] N. M. Temme, "Asymptotic inversion of incomplete gamma functions", *Mathematics of Computation*, vol. 58, no. 198, pp. 755–764, Apr. 1992.