Pilot-Assisted and Blind Joint Data Detection and Channel Estimation in Partial-Time Jamming

Jang-Wook Moon, *Student Member, IEEE*, Tan F. Wong, *Senior Member, IEEE*, and John M. Shea, *Member, IEEE*

***Abstract*—We consider a communication scenario in which a message is received in the presence of partial-time Gaussian jamming and additive white Gaussian noise. We consider a quasi-static channel, in which the amplitude and phase are constant over each packet transmission. The receiver does not know the amplitude and phase of the incoming signal, which symbols are jammed, or even the statistics of the jammer, such as the jamming power and jamming probability. In this scenario, the receiver must accurately estimate the parameters of the channel and the jamming to achieve good performance. We apply the expectation-maximization (EM) algorithm to iteratively approximate the maximum-likelihood (ML) estimator for all of the parameters. We find that the overall performance of the EM algorithm is very sensitive to the initial estimates, so we propose a new initial estimator that offers good performance. The EM algorithm approach requires pilot symbols to resolve a phase ambiguity. Thus, we also present a blind estimation algorithm to avoid the reduction in overall code rate from the use of pilot symbols.**

***Index Terms*—Iterative channel estimation, iterative decoding, iterative processing, partial-time jamming.**

I. Introduction

TIME-VARYING interference or jamming can severely disrupt communications. To accurately detect a message, the receiver must typically be able to characterize the jamming and discriminate between jammed and unjammed symbols. However, the presence of the interfering signal can make it difficult for the receiver to estimate many of the unknown parameters of the message signal. Typically, error-control coding is used to both provide coding gain against the interference and to aid in the discrimination between jammed and unjammed symbols.

Much previous work on hostile jamming focuses on frequency-hopping spread spectrum (FHSS) systems (cf. [1] and [2]). In an FHSS system, the most effective jamming strategy for many channels [3] is for the jammer to concentrate its power in some portions of the system's total spectrum. This approach is known as partial-band interference. After dehopping, the partial-band interference appears as partial-time interference at the input to the decoder. Since the partial-band interference is expected to be constant over each dwell interval, the receiver knows the transition times for the interference.

The use of error-control coding with jamming detection has been considered for FHSS systems in [1]–[10]. In most of these works, the receiver is assumed either to use noncoherent detection and hard-decision decoding [1], [2], [8], [10] or to have perfect knowledge of the amplitude and phase of the arriving signal [4], [6], [7]. The latter assumption requires that the receiver have some sufficiently accurate method to acquire at least the phase of the received signal if coherent detection and soft-decision decoding are used. In [5], the authors propose the use of pilot symbols to aid in this phase acquisition in the presence of jamming. In [9], they further consider a fading channel scenario in an FHSS system, in which the fading coefficient must be estimated in each dwell interval. In both papers, the detection and estimation problems are simplified by the assumption that the jamming is constant over each dwell interval.

In systems with partial-time jamming, the jammer need not turn on and off at predictable times, so it is common to use a hidden Markov model (HMM) for the interference [11]–[13]. If the interferer turns on and off according to a two-state HMM, then the channel is the classic Gilbert–Elliot channel (GEC) [11]. The use of error-control coding with jamming detection in GECs has been studied in [11]–[13]. For GECs, accurate detection of which symbols are jammed also requires an accurate estimate of the amplitude of the message signal. In the absence of interference, iterative channel estimation and decoding has also been considered [14]–[16].

In this paper, we assume that the receiver knows the thermal noise variance, but has neither knowledge of the amplitude of the message signal nor any information about the jamming signal or which symbols are jammed. The primary distinction between the scenarios considered in [5], [9], and in this paper is that we consider a partial-time jamming scenario modeled by a GEC. Unlike [5] and [9], the receiver cannot assume that the jamming starts and stops at particular times or is constant over any particular interval. Thus, the detection and estimation problems considered here are much more difficult. To deal with this difficult detection and estimation problem, we employ the expectation-maximization (EM) algorithm [17]–[20] to approximately obtain the joint maximum-likelihood (ML) estimates for the message and jamming parameters. The EM algorithm has previously been applied to channel estimation and data detection in [21]–[24].

We show that under the EM approach, the problem of detecting the message in the presence of unknown

Paper approved by G. M. Vitetta, the Editor for Equalization and Fading Channels of the IEEE Communications Society. Manuscript received September 17, 2005; revised January 4, 2006. This work was supported in part by the Office of Naval Research under Grant N00014-02-1-0554, in part by the National Science Foundation under Grant ANI-0220287, and in part by the DoD Multidisciplinary University Research Initiative administered by the Office of Naval Research under Grant N00014-00-1-0565. This paper was presented in part at the IEEE International Conference on Communications, Seoul, Korea, May 2005, and in part at the IEEE Military Communications Conference, Monterey, CA, October/November 2004.

J.-W. Moon was with the University of Florida, Gainesville, FL 32611-6130 USA. He is now with Via Telecom, Inc., San Diego, CA 92126 USA (e-mail: jmoon@via-telecom.com).

T. F. Wong and J. M. Shea are with the University of Florida, Gainesville, FL 32611-6130 USA (e-mail: twong@ece.ufl.edu; jshea@ece.ufl.edu).

Digital Object Identifier 10.1109/TCOMM.2006.881400

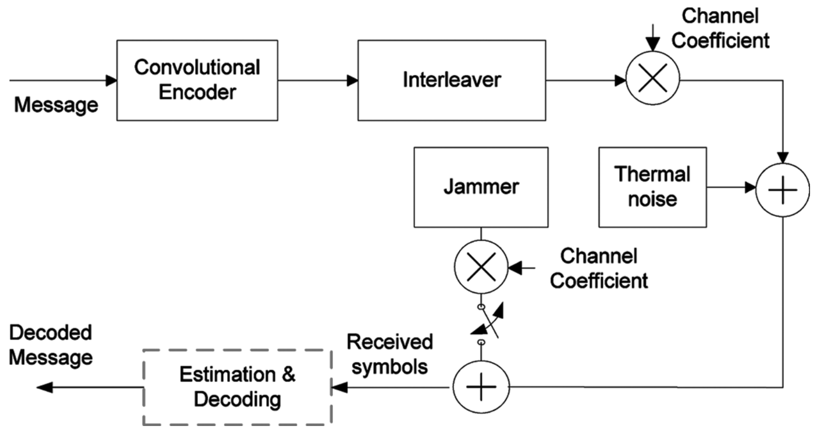


Fig. 1. System model for estimation and decoding in partial-time Gaussian jamming.

channel and jamming parameters results in an iterative detection and estimation procedure, in which two separate Bahl–Cocke–Jelinek–Raviv (BCJR) [25] or Baum–Welch [20] algorithms are used for the message and jamming states. Thus, the overall detector structure is similar to that proposed in [11]. This is in contrast to the approach used in [12] and [13], in which no channel interleaver is used, and the decoding trellis is expanded to incorporate the jamming state. The EM update process requires an initial estimate of the parameters in order to avoid converging to a local minima rather than the ML estimate. In our derivation of the EM estimators, we propose a simple initial estimate. However, the performance of this naive initial estimate is not good enough for the scenarios considered in this paper. So we propose a new estimator that provides a better result. We also present a blind estimation algorithm that does not require the use of pilot symbols.

II. System Model

We assume that the message and jamming signals are received over independent unknown channels. The jamming signal is modeled as additive white Gaussian noise (AWGN). Binary phase-shift keying (BPSK) is used for modulation. Convolutional codes are used for channel coding. To reduce the effects of jamming in the decoder, we use a rectangular channel interleaver. The overall system model is illustrated in Fig. 1. The block labeled "Estimation & Decoding" is the focus of this paper, and is developed in Sections III and IV.

The jammer is modeled using a two-state Markov model [1], [2]. When the jammer is in state 0, it does not transmit the jamming signal; when it is in state 1, the jammer does transmit the jamming signal. The characteristics of the Markov source can be described using two parameters, ρ and $E\{T_J\}$. Here, ρ is the probability that a coded bit is jammed, and $E\{T_J\}$ is the expected value of the time (in terms of number of coded bits) spent in the jamming state before returning to the unjammed state. The four transition probabilities of the two-state Markov model can be determined from ρ and $E\{T_J\}$.

At time instance k, the received symbol after demodulation can be described as

$$y_k = a\sqrt{E_s}u_k + n_k + z_k b J_k \tag{1}$$

where a and b are complex channel coefficients for the message and jamming signals, respectively. It is assumed that they are constant over each packet duration. Here, E_s is the symbol energy, and u_k is the message bit, which takes values ± 1. The parameter z_k, which is 0 or 1, is the indicator value that represents the presence of the jamming signal. The contributions from thermal noise and jamming, given by n_k and J_k, respectively, are zero-mean, circular-symmetric Gaussian random variables. The variances of n_k and J_k are given by N_0 and N_J/ρ, respectively. Without loss of generality, we can assume $b = 1$. Therefore, in the view of the receiver, the variances of the received symbols in state 0 and state 1 are N_0 and $N_0 + N_J/\rho$, respectively.

Let there be N BPSK symbols in a packet. Some of the u_k's may be pilot symbols and are, hence, known. Let S and D be the sets of the indices of the pilot and data symbols, respectively. Without loss of generality, we assume $P(u_k = 1) = 1$ for all of the pilot symbols. Let $\sigma_0^2 = N_0/2$, $\sigma_J^2 = N_J/(2\rho)$, and $\xi = \sigma_J^2/\sigma_0^2$. Note that the receiver has no *a priori* knowledge of the parameters of the message signal or jamming signal other than knowing the power spectral density of the thermal noise, $N_0/2$, and hence, σ_0^2.

III. Estimation Using EM Algorithm

In this section, we use the EM algorithm to estimate the unknown jamming and channel parameters. Because of channel interleaving, we treat the u_k's as independent random variables, and let $P(u_k = -1) = P_k$ and $P(u_k = +1) = 1 - P_k$. Let z_k denote the jammer state at time k. The four transition probabilities for the jammer state are given by $P(z_k = 1|z_{k-1} = 0) = \pi_0$, $P(z_k = 1|z_{k-1} = 1) = \pi_1$, $P(z_k = 0|z_{k-1} = 0) = 1-\pi_0$, and $P(z_k = 0|z_{k-1} = 1) = 1 - \pi_1$. We also define the probabilities for the initial jamming state as $P(z_0 = 1) = \pi^*$ and $P(z_0 = 0) = 1 - \pi^*$. In the following analysis, we use bold letters to denote vectors of parameters or symbols, such as the vector of message probabilities $\mathbf{P} = [P_1, P_2, \ldots, P_N]$, and similarly, the vectors of jamming states $\mathbf{z}$, received symbols $\mathbf{y}$, and transmitted symbols $\mathbf{u}$.

Consider estimating $\boldsymbol{\theta} = (a, \xi, \pi_0, \pi_1, \pi^*, \mathbf{P})$ from $\mathbf{y}$, the vector of the received symbols. The ML estimator can be written as $\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg\max_{\boldsymbol{\theta}} p(\mathbf{y}|\boldsymbol{\theta})$. We use the EM algorithm, with the transmitted information $\mathbf{u}$ and jamming states $\mathbf{z}$ treated as missing data, to iteratively obtain $\hat{\boldsymbol{\theta}}_{\rm ML}$. Note that $p(\mathbf{y},\mathbf{u},\mathbf{z}|\boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\theta},\mathbf{u},\mathbf{z})p(\mathbf{u}|\boldsymbol{\theta})p(\mathbf{z}|\boldsymbol{\theta})$, where

$$\begin{aligned}
p(\mathbf{y}|\boldsymbol{\theta},\mathbf{u},\mathbf{z}) &= \prod_{k=1}^{N} \frac{1}{2\pi\sigma_0^2(1+\xi z_k)} \exp\left[-\frac{|y_k - a\sqrt{E_s}u_k|^2}{2\sigma_0^2(1+\xi z_k)}\right] \\
p(\mathbf{u}|\boldsymbol{\theta}) &= \prod_{k=1}^{N} \left[\frac{1+u_k}{2}(1-P_k) + \frac{1-u_k}{2}P_k\right] \\
p(\mathbf{z}|\boldsymbol{\theta}) &= [z_0\pi^* + (1-z_0)(1-\pi^*)] \\
&\quad \times \prod_{k=1}^{N} [z_k(1-z_{k-1})\pi_0 \\
&\quad + (1-z_k)(1-z_{k-1})(1-\pi_0) + z_k z_{k-1}\pi_1 \\
&\quad + (1-z_k)z_{k-1}(1-\pi_1)]\,.
\end{aligned} \tag{2}$$

Let $\boldsymbol{\theta}^{(n)}$ denote the parameter estimates at the nth iteration of the EM algorithm. Then the EM algorithm amounts to updating $\boldsymbol{\theta}^{(n)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta},\boldsymbol{\theta}^{(n-1)})$, starting from an initial estimate $\boldsymbol{\theta}^{(0)}$. Here, $Q(\boldsymbol{\theta},\boldsymbol{\theta}')$ is Baum's auxiliary function, which is given by

$$\begin{aligned}
Q(\boldsymbol{\theta},\boldsymbol{\theta}') &= E\left[\log p(\mathbf{y},\mathbf{u},\mathbf{z}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] \\
&= E\left[\log p(\mathbf{y}|\mathbf{u},\mathbf{z},\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] + E\left[\log p(\mathbf{u}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] \\
&\quad + E\left[\log p(\mathbf{z}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right].
\end{aligned} \tag{3}$$

For convenience, consider the nth iteration, and let $\boldsymbol{\theta}' = \boldsymbol{\theta}^{(n-1)}$. Then the first term in (3) can be written as

$$\begin{aligned}
&E\left[\log p(\mathbf{y}|\mathbf{u},\mathbf{z},\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] \\
&= -\sum_{k=1}^{N} E\left[\log 2\pi\sigma_0^2(1+\xi z_k)|\mathbf{y},\boldsymbol{\theta}'\right] \\
&\quad - \sum_{k=1}^{N} E\left[\left.\frac{|y_k - a\sqrt{E_s}u_k|^2}{2\sigma_0^2(1+\xi z_k)}\right|\mathbf{y},\boldsymbol{\theta}'\right] \\
&= -\sum_{k=1}^{N}\left[p(z_k=1|\mathbf{y},\boldsymbol{\theta}')\log 2\pi\sigma_0^2(1+\xi)\right. \\
&\qquad \left. + p(z_k=0|\mathbf{y},\boldsymbol{\theta}')\log 2\pi\sigma_0^2\right] \\
&\quad - \sum_{k=1}^{N}\left[p(u_k=1|\mathbf{y},\boldsymbol{\theta}')p(z_k=1|\mathbf{y},\boldsymbol{\theta}')\frac{|y_k - a\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right. \\
&\qquad + p(u_k=-1|\mathbf{y},\boldsymbol{\theta}')p(z_k=1|\mathbf{y},\boldsymbol{\theta}') \\
&\qquad \times \frac{|y_k + a\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)} + p(u_k=1|\mathbf{y},\boldsymbol{\theta}') \\
&\qquad \times p(z_k=0|\mathbf{y},\boldsymbol{\theta}')\frac{|y_k - a\sqrt{E_s}|^2}{2\sigma_0^2} \\
&\qquad + p(u_k=-1|\mathbf{y},\boldsymbol{\theta}')p(z_k=0|\mathbf{y},\boldsymbol{\theta}') \\
&\qquad \left. \times \frac{|y_k + a\sqrt{E_s}|^2}{2\sigma_0^2}\right]
\end{aligned} \tag{4}$$

where in obtaining the second equality, we have approximated $p(u_k, z_k|\mathbf{y},\boldsymbol{\theta}^{(n-1)}) = p(u_k|\mathbf{y},\boldsymbol{\theta})p(z_k|\mathbf{y},\boldsymbol{\theta})$. This approximation is justified by the channel interleaving illustrated in Fig. 1. The other two terms of (3) can be written as

$$\begin{aligned}
&E\left[\log p(\mathbf{u}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] \\
&= \sum_{k=1}^{N} E\left[\log\left(\frac{1+u_k}{2}(1-P_k) + \frac{1-u_k}{2}P_k\right)\middle|\,\mathbf{y},\boldsymbol{\theta}'\right] \\
&= \sum_{k=1}^{N}\left[p(u_k=1|\mathbf{y},\boldsymbol{\theta}')\log(1-P_k)\right. \\
&\qquad \left. + p(u_k=-1|\mathbf{y},\boldsymbol{\theta}')\log P_k\right]
\end{aligned} \tag{5}$$

$$\begin{aligned}
&E\left[\log p(\mathbf{z}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}'\right] \\
&= \sum_{k=1}^{N}\left[p(z_k=1, z_{k-1}=0|\mathbf{y},\boldsymbol{\theta}')\log\pi_0\right. \\
&\qquad + p(z_k=0, z_{k-1}=0|\mathbf{y},\boldsymbol{\theta}')\log(1-\pi_0) \\
&\qquad + p(z_k=1, z_{k-1}=1|\mathbf{y},\boldsymbol{\theta}')\log\pi_1 \\
&\qquad \left. + p(z_k=0, z_{k-1}=1|\mathbf{y},\boldsymbol{\theta}')\log(1-\pi_1)\right] \\
&\quad + p(z_0=1|\mathbf{y},\boldsymbol{\theta}')\log\pi^* \\
&\quad + p(z_0=0|\mathbf{y},\boldsymbol{\theta}')\log(1-\pi^*).
\end{aligned} \tag{6}$$

Because the three terms of $E[\log p(\mathbf{y},\mathbf{u},\mathbf{z}|\boldsymbol{\theta})|\mathbf{y},\boldsymbol{\theta}']$, i.e., (4), (5), and (6), are additive and depend on different components of $\boldsymbol{\theta}$, we can maximize them separately with respect to the corresponding components in $\boldsymbol{\theta}$. First, consider (5). It is easy to see that the choice of

$$\hat{P}_k^{(n)} = p\left(u_k=-1|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right) \tag{7}$$

maximizes (5). Similarly, to maximize (6), we should choose (8)–(10), shown at the bottom of the page. By differentiating

$$\hat{\pi}_0^{(n)} = \frac{\sum_{k=1}^{N} p\left(z_k=1, z_{k-1}=0|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right)}{\sum_{k=1}^{N} p\left(z_k=0, z_{k-1}=0|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right) + p\left(z_k=1, z_{k-1}=0|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right)} \tag{8}$$

$$\hat{\pi}_1^{(n)} = \frac{\sum_{k=1}^{N} p\left(z_k=1, z_{k-1}=1|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right)}{\sum_{k=1}^{N} p\left(z_k=0, z_{k-1}=1|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right) + p\left(z_k=1, z_{k-1}=1|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right)} \tag{9}$$

$$\hat{\pi}^{*(n)} = p\left(z_0=1|\mathbf{y},\boldsymbol{\theta}^{(n-1)}\right) \tag{10}$$

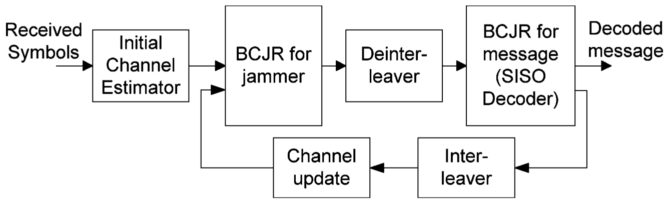


Fig. 2. Iterative estimation and decoding.

(4) with respect to a and ξ and setting the derivatives to zero, we also obtain

$$a = \frac{\sum_{k=1}^{N}\left[p_{u_k}(1) - p_{u_k}(-1)\right]\left[p_{z_k}(1)/(1+\xi) + p_{z_k}(0)\right] y_k}{\sqrt{E_s}\sum_{k=1}^{N}\left[p_{z_k}(1)/(1+\xi) + p_{z_k}(0)\right]} \tag{11}$$

and (12), shown at the bottom of the page, where $p_{z_k}(i) = p(z_k = i|\mathbf{y}, \boldsymbol{\theta})$ and $p_{u_k}(i) = p(u_k = i|\mathbf{y}, \boldsymbol{\theta})$. Thus, we need to solve for a and ξ simultaneously from (11) and (12). Such simultaneous maximization requires a complicated numerical search. So instead, we employ the approximate maximization shown in (13) and (14) at the bottom of the page, where $p_{z_k}^{(n-1)}(i) = p(z_k = i|\mathbf{y}, \boldsymbol{\theta}^{(n-1)})$ and $p_{u_k}^{(n-1)}(i) = p(u_k = i|\mathbf{y}, \boldsymbol{\theta}^{(n-1)})$. To update these parameters, we need to calculate the three kinds of probabilities, which are

$$p\left(u_k = \pm 1|\mathbf{y}, \boldsymbol{\theta}^{(n-1)}\right), \quad p\left(z_k = i|\mathbf{y}, \boldsymbol{\theta}^{(n-1)}\right),$$
$$p\left(z_k = i, z_{k-1} = j|\mathbf{y}, \boldsymbol{\theta}^{(n-1)}\right) \tag{15}$$

where $i, j = 0$ or 1. Since we use convolutional codes and the jamming signal is modeled using a two-state Markov chain, the codeword and jamming state can be directly estimated using two separate BCJR algorithms given the previous iteration parameter estimates. The values of (15) are generated in the two BCJR algorithms. The consequent overall EM algorithm is illustrated in Fig. 2. Note that the iterative decoding process is a byproduct of the EM algorithm.

An initial estimate $\boldsymbol{\theta}^{(0)}$ is needed to run the BCJR algorithms for the jammer state and message. Let $a^{(0)}$ be the average of the pilot symbols

$$a^{(0)} = \frac{1}{\sqrt{E_s}|S|}\sum_{k \in S} y_k \tag{16}$$

where $|S|$ is the cardinality of S, and then use $a^{(0)}$ in (12) with $p(u_k = 1) = p(u_k = -1) = 0.5$ for $k \in D$ and $p(z_k = 1) = p(z_k = -1) = 0.5$, $k = 1, 2, \ldots, N$, which yields

$$\xi^{(0)} = \frac{1}{2\sigma_0^2 N}\left[\sum_{k \in S}\left|y_k - a^{(0)}\sqrt{E_s}\right|^2 + \sum_{k \in D}\left(|y_k|^2 + \left(\left|a^{(0)}\right|^2 E_s\right)\right)\right] - 1. \tag{17}$$

For later use, we call this initial estimate $\boldsymbol{\theta}^{(0)}$ the *simple estimate*.

IV. Soft-Decision Decoding and Jammer State Estimation

In this section, the overall iterative decoding procedure is explained. The receiver employs two BCJR algorithms to provide updates for the probabilities in (15) given the current $\boldsymbol{\theta}^{(n)}$, as described in the previous section. One BCJR algorithm provides the *a posteriori* probabilities (APPs) for the message and coded bits, and the other estimates the APP that each symbol is jammed. They are connected in serial, as illustrated in Fig. 2, and use each other's information to refine the estimates for the message and jammer state.

A. Maximum a Posteriori (MAP) Decoder for Message

We consider first the BCJR decoder for the message [25], [26], which gives as output the APP for a bit in terms of three types of probabilities, $\alpha_k(s')$, $\beta_k(s)$, and $\gamma_k(s', s)$. Each of these probabilities is computed for different states of the convolutional code or transitions between these states. Consider the branch connecting state s' to state s corresponding to the kth bit. Then, $\alpha_k(s')$ is the forward-looking state probability, $\beta_k(s)$ is the backward-looking state probability, and $\gamma_k(s', s)$

$$1 + \xi = \frac{\sum_{k=1}^{N}\left\{|y_k|^2 + |a|^2 E_s + 2\mathrm{Re}\left(y_k^* a\sqrt{E_s}\right)\left[p_{u_k}(-1) - p_{u_k}(1)\right]\right\} p_{z_k}(1)}{2\sigma_0^2\sum_{k=1}^{N} p_{z_k}(1)} \tag{12}$$

$$\hat{a}^{(n)} = \frac{\sum_{k=1}^{N}\left[p_{u_k}^{(n-1)}(1) - p_{u_k}^{(n-1)}(-1)\right]\left[p_{z_k}^{(n-1)}(1)/\left(1+\xi^{(n-1)}\right) + p_{z_k}^{(n-1)}(0)\right] y_k}{\sqrt{E_s}\sum_{k=1}^{N}\left[p_{z_k}^{(n-1)}(1)/\left(1+\xi^{(n-1)}\right) + p_{z_k}^{(n-1)}(0)\right]} \tag{13}$$

$$\hat{\xi}^{(n)} = \frac{\sum_{k=1}^{N}\left\{|y_k|^2 + \left|\hat{a}^{(n)}\right|^2 E_s + 2\mathrm{Re}\left(y_k^*\hat{a}^{(n)}\sqrt{E_s}\right)\left[p_{u_k}^{(n-1)}(-1) - p_{u_k}^{(n-1)}(1)\right]\right\} p_{z_k}^{(n-1)}(1)}{2\sigma_0^2\sum_{k=1}^{N} p_{z_k}^{(n-1)}(1)} - 1 \tag{14}$$

is the branch probability. It can be shown that $\alpha_k(s')$ and $\beta_k(s)$ can be determined in a recursive manner using $\gamma_k(s',s)$ [25]. Assuming the use of a rate-1/2 nonsystematic convolutional code, the branch metric for the nth iteration conditioned on the current parameter estimate $\boldsymbol{\theta}^{(n)}$ is given by

$$\gamma_k\left(s',s|\boldsymbol{\theta}^{(n)}\right)=P(s|s')\cdot p\left(y_k^1|s',s,\boldsymbol{\theta}^{(n)}\right)\cdot p\left(y_k^2|s',s,\boldsymbol{\theta}^{(n)}\right) \tag{18}$$

where y_k^1 and y_k^2 are the received symbol values for the two parity bits corresponding to the kth message bit. Note that $p(y_k^1|s',s,\boldsymbol{\theta}^{(n)})$ and $p(y_k^2|s',s,\boldsymbol{\theta}^{(n)})$ are Gaussian densities if we further condition on whether the bit is jammed. Let z_k^1 and z_k^2 be the states of the jammer for the two parity symbols of the kth message bit. Then for $i=1,2$

$$\begin{aligned}p\left(y_k^i|s',s,\boldsymbol{\theta}^{(n)}\right)=&p\left(y_k^i|s',s,\boldsymbol{\theta}^{(n)},z_k^i=0\right)P\left(z_k^i=0|\boldsymbol{\theta}^{(n)}\right)\\&+p\left(y_k^i|s',s,\boldsymbol{\theta}^{(n)},z_k^i=1\right)P\left(z_k^i=1|\boldsymbol{\theta}^{(n)}\right).\end{aligned} \tag{19}$$

The probabilities $P(z_k^i=0|\boldsymbol{\theta}^{(n)})$ and $P(z_k^i=1|\boldsymbol{\theta}^{(n)})$ are replaced by the APPs generated by the MAP algorithm for the jammer state. As illustrated in Fig. 2, in each iteration, the estimator for the jammer state runs before the BCJR algorithm for the message, so $P(z_k^i=0|\boldsymbol{\theta}^{(n)})$ and $P(z_k^i=1|\boldsymbol{\theta}^{(n)})$ are replaced by $P(z_k^i=0|\mathbf{y},\boldsymbol{\theta}^{(n)})$ and $P(z_k^i=1|\mathbf{y},\boldsymbol{\theta}^{(n)})$, respectively. The MAP algorithm for the jammer state is explained in the next subsection. The decoder performance depends on the accuracy of $\gamma_k(s',s|\boldsymbol{\theta})$. Thus it is important to have accurate knowledge of the jammer state and the channel coefficient.

B. MAP Algorithm for Jammer State Estimation

The BCJR algorithm directly applies for jammer state estimation because the partial-time jammer is modeled using a Markov chain. The APP that a symbol is jammed can be estimated in a similar way as in the previous subsection. Then in the current iteration, the estimated parameters from the previous iteration will be used in the MAP estimate of the current jamming state probabilities. As for the BCJR for the message, all of the probabilities in the MAP algorithm for the jammer state can be determined from $\Gamma_k(z',z)$, which is the branch metric for the transition from jammer state z' to state z at time k. Then

$$\begin{aligned}\Gamma_k&\left(z',z|\boldsymbol{\theta}^{(n)}\right)\\&=P(z|z')\cdot p\left(y_k|z',z,\boldsymbol{\theta}^{(n)}\right)\\&=P(z|z')\cdot\left[p\left(y_k|z',z,\boldsymbol{\theta}^{(n)},u_k=1\right)P\left(u_k=1|\boldsymbol{\theta}^{(n)}\right)\right.\\&\quad\left.+p\left(y_k|z',z,\boldsymbol{\theta}^{(n)},u_k=-1\right)P\left(u_k=-1|\boldsymbol{\theta}^{(n)}\right)\right]\end{aligned} \tag{20}$$

where y_k and u_k represent the received symbol and code bit corresponding to the time instance k, respectively. For the parameters of the form $P(z|z')$, the current estimates from the EM algorithm given by (8) and (9) are used. In place of the probabilities $P(u_k=1|\boldsymbol{\theta}^{(n)})$ and $P(u_k=-1|\boldsymbol{\theta}^{(n)})$, we use the APPs from the BCJR algorithm for the message from the previous iteration, which are $P(u_k=1|\mathbf{y},\boldsymbol{\theta}^{(n-1)})$ and $P(u_k=-1|\mathbf{y},\boldsymbol{\theta}^{(n-1)})$, respectively. Thus, we see that the EM algorithm is an iterative algorithm for estimating $\boldsymbol{\theta}$ and detecting the jamming state and message bits.

V. Improved Initial Estimation

For some cases, the simple initial estimate $\boldsymbol{\theta}^{(0)}$ in (16) and (17) may not be good enough because some of the y_k's can have a very large variance due to the jamming signal. As a result, the EM algorithm may be stuck at a local minimum, giving poor results [27]. In Section VIII, we will present simulation results that illustrate the performance problems of the simple estimate. To increase the accuracy of the initial guess, we propose a new estimator.

A. Derivation of the New Initial Estimator

Let us revisit $p(\mathbf{y}|\boldsymbol{\theta})$, which is given by

$$\begin{aligned}p(\mathbf{y}|\boldsymbol{\theta})&=\sum_{\mathbf{u},\mathbf{z}}p(\mathbf{y}|\mathbf{u},\mathbf{z},\boldsymbol{\theta})p(\mathbf{u}|\boldsymbol{\theta})p(\mathbf{z}|\boldsymbol{\theta})\\&=\sum_{\mathbf{u},\mathbf{z}}\{[z_0\pi^*+(1-z_0)(1-\pi^*)]\\&\quad\times\prod_k\left[\frac{1}{2\pi\sigma_0^2(1+\xi z_k)}\exp\left(-\frac{|y_k-a\sqrt{E_s}u_k|^2}{2\sigma_0^2(1+z_k\xi)}\right)\right.\\&\quad\times\left(\frac{1+u_k}{2}(1-P_k)+\frac{1-u_k}{2}P_k\right)\\&\quad\times(z_k(1-z_{k-1})\pi_0+(1-z_k)(1-z_{k-1})\\&\quad\left.\times(1-\pi_0)+z_kz_{k-1}\pi_1+(1-z_k)z_{k-1}(1-\pi_1))\right]\}.\end{aligned}$$

Direct ML estimation of $\boldsymbol{\theta}$ is difficult. For the initial estimate $\boldsymbol{\theta}^{(0)}$, the decoder has no knowledge of π^*, π_0, π_1, or $\mathbf{P}$ (except for the pilot symbols), and we have observed from simulation that the initial estimates for these values do not contribute significantly to the performance of the EM algorithm. Thus, we set the initial values of these probabilities to 0.5. Then under these assumptions, $p(\mathbf{y}|\boldsymbol{\theta})$ reduces to

$$\begin{aligned}p(\mathbf{y}|a,\xi)=\sum_{\mathbf{u},\mathbf{z}}\prod_{k\in D}&\left[\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}u_k|^2}{2\sigma_0^2(1+z_k\xi)}\right\}}{2\pi\sigma_0^2(1+\xi z_k)}\cdot\frac{1}{4}\right]\\&\times\prod_{k\in S}\left[\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}|^2}{2\sigma_0^2(1+z_k\xi)}\right\}}{2\pi\sigma_0^2(1+\xi z_k)}\cdot\frac{1}{2}\right].\end{aligned}$$

Hence, we have

$$\begin{aligned}&\log p(\mathbf{y}|a,\xi)\\&=\sum_{k\in D}\log\left[\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}|^2}{2\sigma_0^2}\right\}}{8\pi\sigma_0^2}+\frac{\exp\left\{-\frac{|y_k+a\sqrt{E_s}|^2}{2\sigma_0^2}\right\}}{8\pi\sigma_0^2}\right.\\&\qquad\left.+\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right\}}{8\pi\sigma_0^2(1+\xi)}+\frac{\exp\left\{-\frac{|y_k+a\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right\}}{8\pi\sigma_0^2(1+\xi)}\right]\\&\quad+\sum_{k\in S}\log\left[\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}|^2}{2\sigma_0^2}\right\}}{4\pi\sigma_0^2}+\frac{\exp\left\{-\frac{|y_k-a\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right\}}{4\pi\sigma_0^2(1+\xi)}\right]\\&=\sum_{k\in D}\log\left[\exp\left\{-\frac{|y_k|^2+|a|^2E_s}{2\sigma_0^2}\right\}\cosh\frac{\mathrm{Re}\left(y_k^*a\sqrt{E_s}\right)}{\sigma_0^2}\right]\\&\quad+\sum_{k\in D}\log\left[1+\frac{1}{1+\xi}\exp\left\{\frac{|y_k|^2+|a|^2E_s}{2\sigma_0^2(1+1/\xi)}\right\}\right.\end{aligned}$$

$$\left.\times\frac{\cosh\frac{\mathrm{Re}\left(y_k^* a\sqrt{E_s}\right)}{\sigma_0^2(1+\xi)}}{\cosh\frac{\mathrm{Re}\left(y_k^* a\sqrt{E_s}\right)}{\sigma_0^2}}\right] + \sum_{k\in S} -\frac{|y_k - a\sqrt{E_s}|^2}{2\sigma_0^2} + \log\left[1+\frac{1}{1+\xi}\exp\left\{\frac{|y_k - a\sqrt{E_s}|^2}{2\sigma_0^2(1+1/\xi)}\right\}\right] - N\log 4\pi\sigma_0^2. \tag{21}$$

For high signal-to-noise ratios (SNRs), where $\sqrt{E_s}/\sigma_0^2 \gg 1$, $\cosh[\mathrm{Re}(y_k^* a\sqrt{E_s})\sqrt{E_s})/\sigma_0^2] \approx 0.5\exp\{|\mathrm{Re}(y_k^* a\sqrt{E_s})|/\sigma_0^2\}$, and $\cosh[\mathrm{Re}(y_k^* a\sqrt{E_s})/\{\sigma_0^2(1+\xi)\}] \approx 0.5 \exp\{|\mathrm{Re}(y_k^* a\sqrt{E_s})|/\{\sigma_0^2(1+\xi)\}]$. Therefore

$$\log p(\mathbf{y}|a,\xi) \approx \sum_{k\in D} A_k + \sum_{k\in S} B_k - |D|\log 2 - N\log 4\pi\sigma_0^2$$

where $|D|$ is the cardinality of the set of data symbols

$$A_k = -\frac{\bar{d}(y_k,a)}{2\sigma_0^2} + \log\left[1+\frac{1}{1+\xi}\exp\left\{\frac{\bar{d}(y_k,a)}{2\sigma_0^2(1+1/\xi)}\right\}\right]$$
$$B_k = -\frac{d(y_k,a)}{2\sigma_0^2} + \log\left[1+\frac{1}{1+\xi}\exp\left\{\frac{d(y_k,a)}{2\sigma_0^2(1+1/\xi)}\right\}\right].$$

Here $d(y_k,a) \triangleq |y_k - a\sqrt{E_s}|^2$ and $\bar{d}(y_k,a) \triangleq \min\{|y_k - a\sqrt{E_s}|^2, |y_k + a\sqrt{E_s}|^2\}$. For fixed (a,ξ) and k, such that $\bar{d}(y_k,a) \gg \Psi \triangleq 2\sigma_0^2(1+1/\xi)\log(1+\xi)$, we can approximate A_k by

$$A_k \approx \frac{1}{2\sigma_0^2}\left[-2\sigma_0^2\log(1+\xi) - \frac{\bar{d}(y_k,a)}{1+\xi}\right]. \tag{22}$$

On the other hand, for k such that $\bar{d}(y_k,a) \ll \Psi$

$$A_k \approx -\bar{d}(y_k,a)/\left(2\sigma_0^2\right). \tag{23}$$

Similarly, B_k can be approximated by using (22) and (23), with $\bar{d}$ replaced by d. Here, we can further approximate $\log p(\mathbf{y}|a,\xi)$ by partitioning D into subsets $\{k \in D : \bar{d}(y_k,a) < \Psi\}$ and $\{k \in D : \bar{d}(y_k,a) \geq \Psi\}$. Similar approximations are also applied for the set S. Eventually, we have

$$\begin{aligned}\log p(\mathbf{y}|a,\xi) &\approx -\frac{1}{2\sigma_0^2}\sum_{k\in A_{00}}\bar{d}(y_k,a) - \frac{1}{2\sigma_0^2}\sum_{k\in A_{10}} d(y_k,a)\\ &\quad - \frac{1}{2\sigma_0^2(1+\xi)}\sum_{k\in A_{01}}\bar{d}(y_k,a) - \frac{1}{2\sigma_0^2(1+\xi)}\\ &\quad \times\sum_{k\in A_{11}} d(y_k,a) - (|A_{01}|+|A_{11}|)\log(1+\xi)\\ &\quad - |D|\log 2 - N\log 4\pi\sigma_0^2\end{aligned} \tag{24}$$

where $A_{00} = \{k \in D : \bar{d}(y_k,a) < \Psi\}$, $A_{01} = \{k \in D : \bar{d}(y_k,a) \geq \Psi\}$, $A_{10} = \{k \in S : d(y_k,a) < \Psi\}$, and $A_{11} = \{k \in S : d(y_k,a) \geq \Psi\}$. Then the approximate ML estimator for (a,ξ) can be obtained by minimizing

$$\begin{aligned}c(a,\xi) \triangleq &\sum_{k\in A_{00}}\bar{d}(y_k,a) + \sum_{k\in A_{10}} d(y_k,a) + \frac{1}{1+\xi}\sum_{k\in A_{01}}\bar{d}(y_k,a)\\ &+\frac{1}{1+\xi}\sum_{k\in A_{11}} d(y_k,a) + (|A_{01}|+|A_{11}|)\,2\sigma_0^2\log(1+\xi).\end{aligned} \tag{25}$$

The overall steps to find $\hat{\xi}$ and $\hat{a}$ from this approximate ML estimator are as follows.

1) Start with some small initial guess for $\hat{\xi}$ and find the corresponding $\hat{\Psi}$.
2) Partition the complex plane into squares of size $2\hat{\Psi}$.
3) Assume that the center of each bin is the temporary $\hat{a}\sqrt{E_s}$. Determine the four regions $A_{ij},\ i,j = 0,1$ for each $\hat{a}$. Calculate (25) for each bin.
4) Choose the bin that gives minimum value of (25) and store the corresponding $\hat{a}$, $\hat{\xi}$, and $c(\hat{a},\hat{\xi})$.
5) Repeat steps 2)–5) by increasing $\hat{\xi}$ and updating the optimal $\hat{a}$, $\hat{\xi}$, and $c(\hat{a},\hat{\xi})$.
6) After finishing the iterations, the stored $\hat{a}$ and $\hat{\xi}$ give the improved initial estimate.

We call this estimate the *improved estimate*. Note that for the improved estimate, it is important to have enough pilot symbols to ensure that the estimate is not π radians out of phase. To better understand this, we revisit (25). Note that the first and third summations generate the same value for the bins at a π radians offset. In the absence of pilot symbols, two possible candidate bins will have the same value of (25). The second and fourth summations, which depend on the pilot symbols, decide between the two bins. Because the pilot symbols are also susceptible to jamming, we need enough pilots to ensure an accurate initial estimate. The performance with various numbers of pilot symbols is presented in Section VIII.

B. Cramer–Rao Bounds

In this section, we derive the Cramer–Rao bound (CRB) for the variance of the proposed initial estimators for a and ξ. Assume that we estimate the vector parameter $[a_r\ a_i\ \xi]$, where a_r and a_i are the real and imaginary parts of a, respectively. The CRB for each component is the diagonal element of the inverse of the Fisher information matrix (FIM). We show how to calculate the element of the FIM that is in the first row and first column. From (21), let M_D and M_S be the argument of the log for the data and pilot symbols, respectively, and M_D', M_D'', M_S', and M_S'' be the first and second partial derivatives of M_D and M_S with respect to a_r, respectively. For both M_D and M_S, it can be shown that

$$\begin{aligned}M' = &v_0\frac{\sqrt{E_s}(y_r - a_r\sqrt{E_s})}{8\pi\sigma_0^4}\exp\left[\frac{-|y-a_r\sqrt{E_s}|^2}{(2\sigma_0^2)}\right]\\ &- v_1\frac{\sqrt{E_s}(y_r + a_r\sqrt{E_s})}{8\pi\sigma_0^4}\exp\left[\frac{-|y+a_r\sqrt{E_s}|^2}{(2\sigma_0^2)}\right]\\ &+ v_0\frac{\sqrt{E_s}(y_r - a_r\sqrt{E_s})}{8\pi\sigma_0^4(1+\xi)^2}\exp\left[\frac{-|y-a_r\sqrt{E_s}|^2}{(2\sigma_0^2(1+\xi))}\right]\\ &- v_1\frac{\sqrt{E_s}(y_r + a_r\sqrt{E_s})}{8\pi\sigma_0^4(1+\xi)^2}\exp\left[\frac{-|y+a_r\sqrt{E_s}|^2}{(2\sigma_0^2(1+\xi))}\right]\end{aligned}$$

and

$$
\begin{aligned}
M'' = & v_0\left[\exp\left(-\frac{|y-a_r\sqrt{E_s}|^2}{2\sigma_0^2}\right)\right. \\
& \left.\times\left(\frac{E_s(y_r-a_r\sqrt{E_s})^2}{\sigma_0^2}-1\right)\right]/\left(8\pi\sigma_0^4\right) \\
& - v_1\left[\exp\left(-\frac{|y+a_r\sqrt{E_s}|^2}{2\sigma_0^2}\right)\right. \\
& \left.\times\left(\frac{E_s(y_r+a_r\sqrt{E_s})^2}{\sigma_0^2}-1\right)\right]/\left(8\pi\sigma_0^4\right) \\
& + v_0\left[\exp\left(-\frac{|y-a_r\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right)\right. \\
& \left.\times\left(\frac{E_s(y_r-a_r\sqrt{E_s})^2}{\sigma_0^2(1+\xi)}-1\right)\right]/\left(8\pi\sigma_0^4(1+\xi)^2\right) \\
& - v_1\left[\exp\left(-\frac{|y+a_r\sqrt{E_s}|^2}{2\sigma_0^2(1+\xi)}\right)\right. \\
& \left.\times\left(\frac{E_s(y_r+a_r\sqrt{E_s})^2}{\sigma_0^2(1+\xi)}-1\right)\right]/\left(8\pi\sigma_0^4(1+\xi)^2\right)
\end{aligned}
$$

where $v_0 = 1$ for M_D and 2 for M_S, and $v_1 = 1$ for M_D and 0 for M_S. Then it can be shown that

$$
\begin{aligned}
\frac{\partial^2 \log p(\mathbf{y}|a,\xi)}{\partial a_r^2} = & |D|\left[M_D'' M_D^{-1} - M_D'^2 M_D^{-2}\right] \\
& + |S|\left[M_S'' M_S^{-1} - M_S'^2 M_S^{-2}\right]. \qquad (26)
\end{aligned}
$$

Therefore

$$
\begin{aligned}
& E\left[\frac{\partial^2 \log p(\mathbf{y}|a,\xi)}{\partial a_r^2}\right] \\
& = |D|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[M_D'' M_D^{-1} - M_D'^2 M_D^{-2}\right] M_D\, dy_r\, dy_i \\
& \quad + |S|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[M_S'' M_S^{-1} - M_S'^2 M_S^{-2}\right] M_S\, dy_r\, dy_i. \qquad (27)
\end{aligned}
$$

Note that the other required elements of the FIM can be obtained in a similar manner. Then, the CRB for a is obtained by adding the first two diagonal elements of the inverse of the FIM, and the CRB for ξ is the last diagonal element.

VI. Blind Estimation Algorithm

The initial estimators proposed in the previous sections require pilot symbols to resolve a π-radian ambiguity in the phase of the message signal. This ambiguity results from the use of BPSK modulation, which causes some symbols to have mean a, and others to have mean $-a$. Pilot symbols allow this ambiguity to be resolved, because the mean values of those symbols are known *a priori*. However, the use of pilot symbols reduces the overall code rate. Therefore, in this section, we propose a blind decoding algorithm to avoid the use of pilot symbols.

The approach we use to deal with this ambiguity is to decode for both of the cases, $\hat{a}$ and $-\hat{a}$. In each iteration, the two decodings will result in two possible decoded message sequences. We use the path metrics for the message sequence to select a candidate sequence for use in further iterations or as the final decoder output. The use of path metrics to choose the correct codeword is reasonable for the following reason. Note that the flipped sequence (0 to 1, 1 to 0) of a valid codeword is not a valid codeword, because the all $\mathbf{1}$ sequence is not a valid codeword for the rate-1/2 convolutional code that we use. So with high probability, the path metric obtained from $\hat{a}$ with the correct sign will be greater than the other path metric from $\hat{a}$ with the incorrect sign.

Note that the path metric is easily obtained using the Viterbi algorithm, and thus, from the max-log-MAP algorithm [28], which provides the same hard decision as the Viterbi algorithm. The ML path metric in the Viterbi algorithm is the metric of the terminating state, and this value is the summation of all branch metrics along the ML path. For the max-log-MAP algorithm, the summation of all the branch metrics on the ML path is

$$
\sum_k \gamma_k(s_{k-1}, s_k|\boldsymbol{\theta}) = \sum_k p(y_k|s_{k-1}, s_k, \boldsymbol{\theta})P(u_k) \qquad (28)
$$

where $P(u_k)$ is the *a priori* probability for the message bit u_k, which we assume is 0.5. Here the sequence of states $\{s_k\}$ corresponds to the states on the ML path. Therefore, the ML path metric from the Viterbi algorithm is a linear function of the summation of all $\gamma_k(s', s)$ along the ML path. Exploiting these facts, we perform decoding as follows.

1) Calculate improved initial estimate as described in Section V without using pilot symbols.
2) Decode for two cases, $\hat{a}$ and $-\hat{a}$.
3) Select one of the two $\pm\hat{a}$ that gives the greater path metric. Stop and output the decoded sequence if this is the last iteration.
4) Update estimates of $\hat{a}$ and the jamming parameters and probabilities. Return to step 2).

We call this decoding procedure the *blind* scheme. Performance results for this scheme are presented in Section VIII.

VII. Algorithm Complexity

In this section, we briefly study the complexity of the iterative estimation and decoding algorithms described in the previous sections. First, we consider the pilot-assisted case, as depicted in Fig. 2. In the discussion below, when the order of complexity of an algorithm is discussed, we roughly refer to the complexity order in the unit of a complex-valued multiplication and a few complex-valued additions. Moreover, we assume a rate-1/2 code is employed. Thus, the number of data bits is roughly $N/2$.

For the improved initial estimator, the procedure described in Section V can be effectively implemented by forming a two-dimensional histogram from N received symbols with bin size $2\hat{\Psi}$. The calculation of (25) in Step 3) of the procedure requires a complexity of $\mathcal{O}(N)$ per bin. Due to the use of BPSK and circular symmetry of the noise and jamming signal, we can see that the maximum number of bins needed is roughly N. Thus, the complexity order of the improved initial estimator is $\mathcal{O}(N^2)$. In the iterative loop, the BCJR decoder and jammer state estimator are standard BCJR algorithms. Their combined complexity order is $\mathcal{O}(N2^m + N2^2)$ per iteration, where m is the constraint length of the convolutional code. The "Channel Update" block evaluates (8)–(10), (13), and (14). Its order of complexity is $\mathcal{O}(N)$ per iteration.

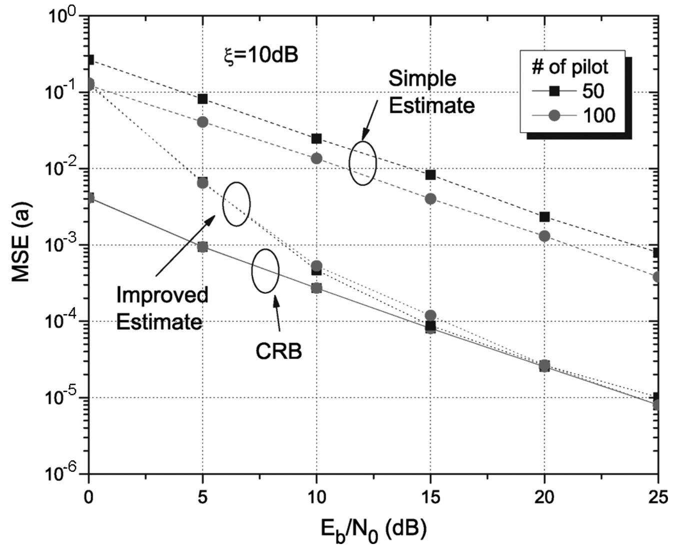


Fig. 3. Comparison of MSE for simple and improved estimates of channel gain a.

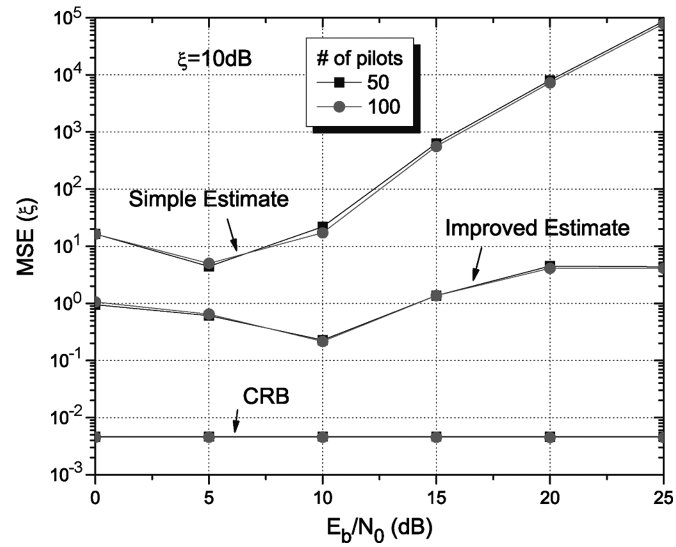


Fig. 4. Comparison of MSE for simple and improved estimates of ξ.

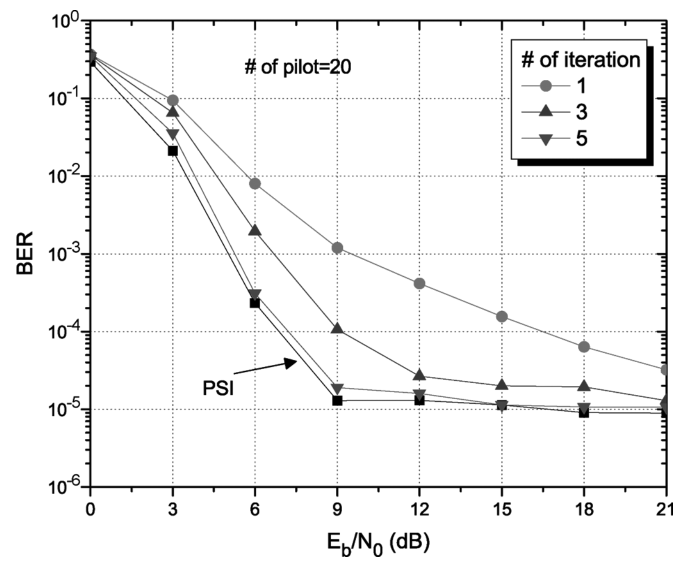


Fig. 5. Performance versus number of iterations in the EM algorithm.

From the above, if I iterations are performed, then the overall complexity of the pilot-assisted joint channel estimation and decoding algorithm is $\mathcal{O}(N^2 + IN(2^m + 5))$. For the blind algorithm described in Section VI, we need to run two BCJR decoders simultaneously, and hence, its order of complexity is $\mathcal{O}(N^2 + IN(2^{m+1} + 5))$. In summary, we can see that the initial improved estimator contributes mostly to the computational complexity of both the pilot-assisted and blind algorithms, when the number of received symbols N is large. Nevertheless, we can further bring down the complexity order of the improved initial estimator to be close to linear in N, at the expense of a slight performance loss, by only calculating the metric in (25) for bins that contain a significant number of received symbols, and ignoring bins that contain only a few symbols in the procedure described in Section V.

VIII. Simulation Results

In this section, we present performance results for the proposed iterative estimation and decoding algorithms. For the results presented in this paper, the rate-1/2 convolutional code with memory 6 and generator polynomials 133 and 171 (in octal) is used. The information block size (without tail bits) is 1000. Therefore, $N = 2012 + |S|$, where $|S|$ is the number of pilot symbols. The channel interleaver is a rectangular interleaver of size 50×43. We consider a quasi-static AWGN channel. We set $|a| = 1$ and $b = 1$. The phase of a varies from packet to packet.

First, we compare performance of the proposed initial estimates for a and ξ with the CRB. We compare the performance with independent jamming and $\pi_0 = \pi_1 = \pi^* = 0.5$, which is the assumption under which the initial estimators and the CRB are derived. For 50 and 100 pilot symbols and $\xi = 10$ dB, the mean-squared error (MSE) of the initial estimates for a are compared with the CRB in Fig. 3. The MSE of the improved estimate converges to the CRB at high SNR. The MSE of the simple estimate is much larger than that of the improved estimate. As previously mentioned, the simple estimate may be problematic, as the EM algorithm can get stuck in a local minimum if the initial estimate is not accurate. The comparison between the MSE of the initial estimates for ξ and the CRB is shown in Fig. 4. Again, the improved estimate shows better performance than the simple estimate. It can be shown that the simple estimate of ξ is biased, and the bias term increases as SNR increases. This is the reason that the MSE of the simple estimate increases. In comparison, the MSE of the improved estimate is relatively flat.

For the results in Figs. 5–7, we consider the performance for $E_b/N_J = 0$ dB, $\rho = 0.2$, and $E\{T_J\} = 35$. The results in Fig. 5 illustrate the performance of the EM algorithm with different numbers of iterations. Here we show the results for the improved estimate only. The performance gradually converges to that of the case with perfect side information (PSI) as the number of iterations increases, regardless of E_b/N_0. As E_b/N_0 is increased, the performance is eventually dominated by the fixed jamming power, which results in the error floor evident in the figure. The results show that five iterations of the EM algorithm provides performance close to the PSI performance. To ensure that our

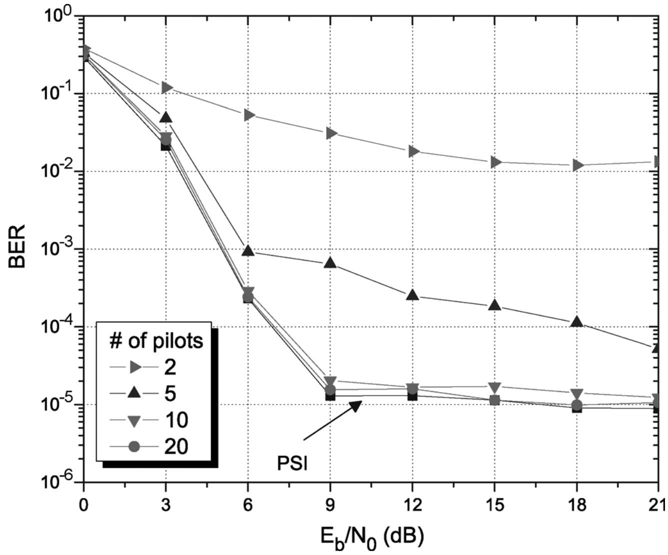


Fig. 6. Performance of the improved estimate with various numbers of pilot symbols.

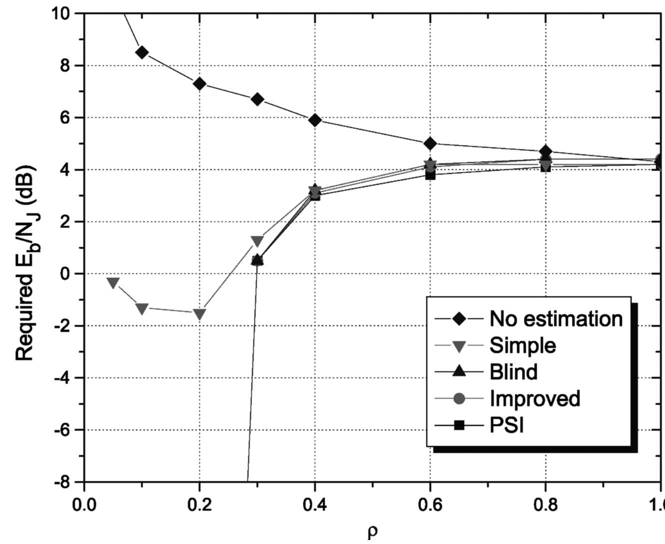


Fig. 8. Required E_b/N_J to obtain a FER of 10^{-2} ($E_b/N_0 = 15$ dB).

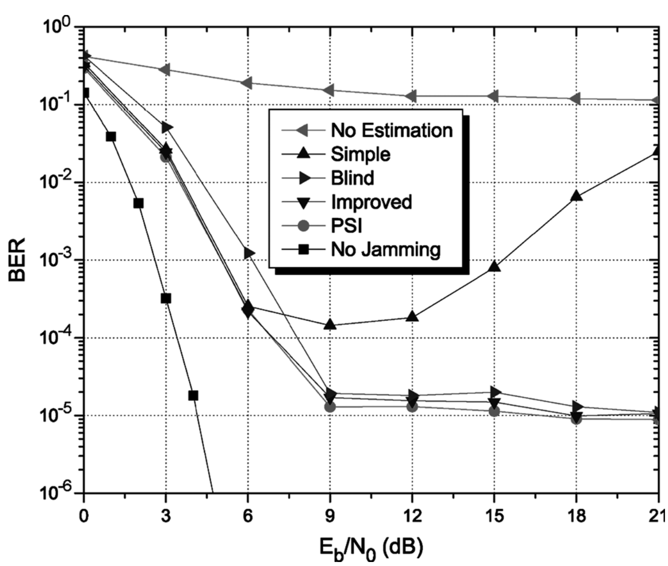


Fig. 7. Performance of overall estimation and decoding process with various initial estimates.

iterative approach can achieve performance close to that of PSI as other parameters are varied, we use 10 EM iterations for the remainder of the results presented in this paper.

We next consider the performance of the improved estimate with various numbers of pilot symbols. These results are illustrated in Fig. 6. As previously explained, the improved estimate will not work properly if we do not use enough pilot symbols. We can see that the performance with 10–20 pilot symbols is very close to the performance when PSI is available. Note that further increasing the number of pilot symbols will decrease the overall code rate. This decreases the effective E_s/N_0, which will increase the error probabilities.

The bit error rates (BERs) for the pilot-assisted and blind detection and decoding processes are illustrated in Fig. 7. The curve labeled "PSI" illustrates the results with PSI for the channel and jamming parameters. For the curves labeled "Simple" and "Improved," the receiver uses the pilot-assisted estimation algorithm with the simple estimate or the improved estimate, respectively. For the curve labeled "Blind," the receiver uses the improved estimator with blind detection. For these results, the total number of iterations is 10, and the number of pilot symbols is 20. We observe that the performance with the simple estimate is very poor at high E_b/N_0. From Figs. 3 and 4, we can see that the MSE of the simple estimate is too large, and hence, the EM algorithm often converges to local maxima that are different from the correct message. The thermal noise provides random jitter that allows the EM algorithm to escape from these local maxima. At high E_b/N_0, the thermal noise is negligible, and hence, the EM algorithm very often remains trapped in the local maxima. As a result, the BER performance gets worse at high E_b/N_0. However, using the improved estimator, we see that the decoding algorithm offers performance close to that of PSI.

The performance of the blind scheme is nearly as good as the performance with pilots. Thus, the use of the path metric is an effective way to select the candidate codeword to be retained at the end of each iteration. The disadvantage of this scheme is the decoder complexity and decoding time. The label "No estimation" represents the case that only the channel coefficient is estimated using pilot symbols, and the receiver operates as if there is no jamming signal. The poor performance of this approach illustrates the need to accurately identify jammed symbols and the jamming parameters.

As suggested in [1], the anti-jam capability of the jamming mitigation schemes can be measured by determining ρ^*, which is the value of ρ required to prevent the receiver from achieving an acceptable error probability. The higher the value of ρ^*, the more symbols that must be jammed in order to significantly degrade communications. In Fig. 8, the value of E_b/N_J that is required to achieve a frame-error rate (FER) of 10^{-2} is shown as a function of ρ, the probability that a symbol is jammed. For these results, $E_b/N_0 = 15$ dB. The curve labeled "PSI" corresponds to the performance with perfect knowledge of which bits are jammed and perfect knowledge of E_b/N_J and ρ. The value

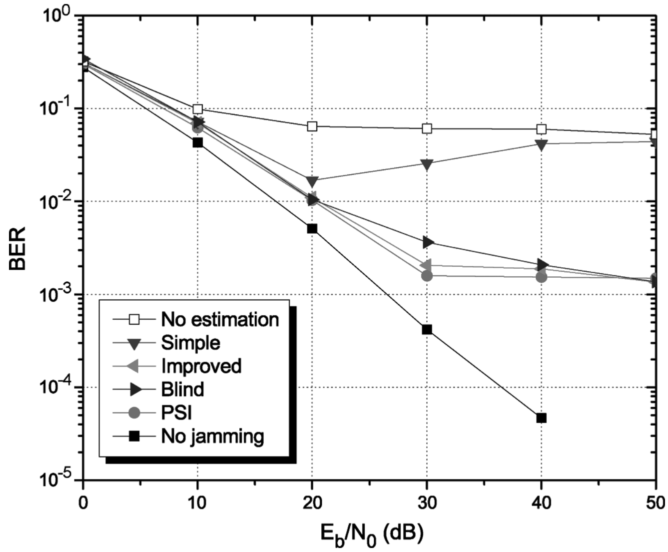


Fig. 9. Performance of jamming detection with various initial estimates in quasi-static fading channels ($E_b/N_J = 8$ dB, $\rho = 0.4$).

of ρ^* for all of the estimates except the simple estimate is approximately 0.3. The blind scheme again shows approximately the same performance as the improved estimate. The EM algorithm with the simple estimate offers poor performance under moderate jamming ($\rho < 0.3$). For the "No estimation" scheme, the required E_b/N_J is a monotonically decreasing function of ρ. This is because the decoder treats all symbols as unjammed; as ρ increases, the jamming signal is spread more evenly over all the symbols, so the performance improves.

The proposed improved estimate and jamming detection scheme can be applied to quasi-static fading channels (where the fading coefficient is constant over each codeword) without any modification. We consider the case of independent fading for the message and jamming signal. So, a and b are independent, complex Gaussian random variables with mean 0 and variance 0.5 per each dimension. The BERs for the various iterative detection and decoding processes are illustrated in Fig. 9 for an average bit energy-to-jamming power spectral density of $E_b/N_J = 8$ dB and $\rho = 0.4$. Again, the EM algorithm with the improved estimate and either the pilot-assisted or blind estimator provides performance close to that of PSI. The EM algorithm with the simple estimate offers poor performance.

IX. Conclusions

In this paper, we consider data communication in the presence of partial-time jamming. For best performance, the receiver requires an accurate estimate of the channel amplitude and phase along with estimates of the jamming parameters and which symbols are jammed. We use the EM algorithm to iteratively approximate the ML estimates for the channel and jamming parameters. The simulation results demonstrate that using simple initial estimates for the channel coefficient and relative fading strength is not good enough for the EM algorithm to converge in many cases of interest. Thus, to achieve better performance, we develop an improved estimator that offers performance close to that of PSI. Both the simple and improved estimators require pilot symbols to resolve a π-radian ambiguity. The use of pilot symbols reduces the overall code rate, and hence, the error performance. So we also propose a blind decoding algorithm to avoid the use of pilots. We show that this blind decoding scheme also gives performance close to that of the pilot-assisted estimators, at the expense of higher complexity.

References

[1] M. B. Pursley and W. E. Stark, "Performance of Reed–Solomon coded frequency-hop spread-spectrum communications in partial-band interference," *IEEE Trans. Commun.*, vol. COM-33, no. 8, pp. 767–774, Aug. 1985.

[2] C. D. Frank and M. B. Pursley, "Concatenated coding for frequency-hop spread-spectrum with partial-band interference," *IEEE Trans. Commun.*, vol. 44, no. 3, pp. 377–387, Mar. 1996.

[3] A. A. Ali, "Worst-case partial-band noise jamming of Rician fading channels," *IEEE Trans. Commun.*, vol. 44, no. 6, pp. 660–662, Jun. 1996.

[4] J. H. Kang and W. E. Stark, "Turbo codes for coherent FH-SS with partial band interference," in *Proc. IEEE Mil. Commun. Conf.*, 1997, pp. 5–9.

[5] H. El Gamal and E. Geraniotis, "Turbo codes with channel estimation and dynamic power allocation for anti-jam FH/SSMA," in *Proc. IEEE Mil. Commun. Conf.*, Boston, MA, Oct. 1998, vol. 1, pp. 170–175.

[6] J. H. Kang and W. E. Stark, "Turbo codes for noncoherent FH-SS with partial band interference," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1451–1458, Nov. 1998.

[7] ——, "Iterative estimation and decoding for FH-SS with slow Rayleigh fading," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2014–2023, Dec. 2000.

[8] M. B. Pursley and J. S. Skinner, "Turbo product coding in frequency-hop wireless communications with partial-band interference," in *Proc. IEEE Mil. Commun. Conf.*, 2002, pp. 774–779.

[9] H. El Gamal and E. Geraniotis, "Iterative channel estimation and decoding for convolutionally coded anti-jamming FH signals," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 321–331, Feb. 2002.

[10] L.-L. Yang and L. Hanzo, "Low-complexity erasure insertion in RS-coded SFH spread-spectrum communications with partial-band interference and Nakagami-m fading," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 914–925, Jun. 2002.

[11] J. H. Kang, W. E. Stark, and A. O. Hero, "Turbo codes for fading and burst channels," in *Proc. IEEE Commun. Theory Mini-Conf. Globecom*, Nov. 1998, pp. 40–45.

[12] J. Garcia-Frias and J. D. Villasenor, "Turbo decoders for Markov channels," *IEEE Commun. Letters*, vol. 2, no. 9, pp. 257–259, Sep. 1998.

[13] ——, "Turbo decoding of Gilbert–Elliot channels," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 357–363, Mar. 2002.

[14] M. C. Valenti and B. D. Woerner, "Refined channel estimation for coherent detection of turbo codes over flat-fading channels," *Electron. Lett.*, vol. 34, no. 17, pp. 1648–1649, Aug. 1998.

[15] ——, "Iterative channel estimation and decoding of pilot assisted turbo codes over flat-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[16] Q. Li, C. N. Georghiades, and X. Wang, "An iterative receiver for turbo-coded pilot-assisted modulation in fading channels," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 145–147, Apr. 2001.

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

[18] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

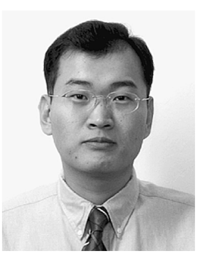
[19] J. Bilmes, "A gentle tutorial on the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models," 1997 [Online]. Available: http://citeseer.ist.psu.edu/bilmes98gentle.html

[20] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Stat.*, vol. 41, pp. 164–171, 1970.

[21] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear and nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, Jul. 1994.

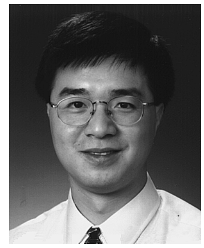
[22] C. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, no. 3, pp. 300–308, Mar. 1997.

[23] H. Zamiri-Jafarian and S. Pasupathy, "EM-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1297–1302, Sep. 1999.
[24] A. Kocian and B. Fleury, "EM-based joint data detection and channel estimation of DS-CDMA signals," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1709–1720, Oct. 2003.
[25] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rates," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
[26] W. E. Ryan, "Concatenated convolutional codes and iterative decoding," in *Wiley Encyclopedia of Telecommunications*, J. G. Proakis, Ed. New York: Wiley, 2003.
[27] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Statist.*, vol. 11, pp. 95–103, 1983.
[28] M. P. C. Fossorier, F. Burkert, S. Lin, and J. Hagenauer, "On the equivalence between SOVA and max-log-MAP decodings," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 137–139, May 1998.



Jang-Wook Moon (S'03) received the B.S. degree in electrical engineering from Sung-Kyun-Kwan University, Suwon, Korea in 1999, the M.S. degree in electrical and computer engineering from Yonsei University, Seoul, Korea in 2001, and the Ph.D. degree in electrical and computer engineering from the University of Florida, Gainesville, in 2005.

He is currently a Senior Systems Engineer with VIA Telecom, Inc., San Diego, CA. His research interest is in wireless communications with emphasis on turbo coding and iterative decoding, anti-jamming communication techniques, collaborative communications, CDMA, and ad hoc networks.



Tan F. Wong (S'96–M'98–SM'03) received the B.Sc. degree (with first-class honors) in electronic engineering from the Chinese University of Hong Kong in 1991, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1992 and 1997, respectively.

He was a Research Engineer with the Department of Electronics, Macquarie University, Sydney, Australia, where he was involved with the high-speed wireless networks project. He also served as a Postdoctoral Research Associate with the School of Electrical and Computer Engineering, Purdue University. Since August 1998, he has been with the University of Florida, Gainesville, where he is currently an Associate Professor of Electrical and Computer Engineering.

Prof. Wong serves as the Editor for Wideband and Multiple-Access Wireless Systems for the IEEE TRANSACTIONS ON COMMUNICATIONS and as an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

John M. Shea (S'92–M'99) received the B.S. degree (with highest honors) in computer engineering in 1993, and the M.S. and Ph.D. degrees in electrical engineering in 1995 and 1998, respectively, all from Clemson University, Clemson, SC.

He is currently an Associate Professor of Electrical and Computer Engineering with the University of Florida, Gainesville. Prior to that, he was an Assistant Professor with the University of Florida from July 1999 to August 2005, and a Postdoctoral Research Fellow with Clemson University from January 1999 to August 1999. He was a Research Assistant in the Wireless Communications Program at Clemson University from 1993 to 1998. He is currently engaged in research on wireless communications with emphasis on error-control coding, cross-layer protocol design, cooperative diversity techniques, and hybrid ARQ.

Dr. Shea was selected as a Finalist for the 2004 Eta Kappa Nu Outstanding Young Electrical Engineer Award. He received the Ellersick Award from the IEEE Communications Society in 1996. He was a National Science Foundation Fellow from 1994 to 1998. He is an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.