EURASIP Journal on Wireless Communications and Networking 2005:2, 100–116

Soft-In Soft-Output Detection in the Presence of Parametric Uncertainty via the Bayesian EM Algorithm

A. S. Gallo
Department of Information Engineering, University of Modena and Reggio Emilia, via Vignolese 905, 41100 Modena, Italy
Email: asgallo@unimo.it

G. M. Vitetta
Department of Information Engineering, University of Modena and Reggio Emilia, via Vignolese 905, 41100 Modena, Italy
Email: giorgio.vitetta@unimo.it

Received 30 April 2004; Revised 6 October 2004

We investigate the application of the *Bayesian expectation-maximization* (BEM) technique to the design of *soft-in soft-out* (SISO) detection algorithms for wireless communication systems operating over channels affected by parametric uncertainty. First, the BEM algorithm is described in detail and its relationship with the well-known *expectation-maximization* (EM) technique is explained. Then, some of its applications are illustrated. In particular, the problems of SISO detection of spread spectrum, single-carrier and multicarrier space-time block coded signals are analyzed. Numerical results show that BEM-based detectors perform closely to the *maximum likelihood* (ML) receivers endowed with perfect channel state information as long as channel variations are not too fast.

Keywords and phrases: expectation-maximization algorithm, soft-in soft-out detection, fading channels, space-time coding, OFDM.

1. INTRODUCTION

In recent years, many research efforts have been devoted to the study of detection algorithms for digital signals transmitted over channels affected by random parametric uncertainty, like multipath fading channels and AWGN channels with phase jitter (see, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and references therein). In this field the attention has been progressively shifting from *maximum likelihood* (ML) sequence detection [2, 3, 4] to *maximum a posteriori* (MAP) symbol detection techniques [5, 6, 7, 8, 9, 10, 11, 12, 13] producing *a posteriori probabilities* (APPs) on the possible data decisions. This has been mainly due to the need of robust receiver structures for coded modulations and, more specifically, to the advent of the *turbo processing principle* applied to efficient iterative decoding of concatenated coding structures [14, 15, 16, 17, 18, 19, 20, 21, 22]. Such a principle has been also exploited to design iterative detection/equalization/decoding algorithms for interleaved coded signals transmitted over channels with memory [10, 11, 12, 13, 23]. In all these cases good error performance is achieved by means of concatenated detection/decoding structures exchanging among each other soft information about the detected data. The basic building blocks of these structures are the so-called *soft-in soft-out* (SISO) modules [18, 22].

A wealth of technical papers on the design techniques for ML sequence detectors operating on channels with parametric uncertainty is available (see [1, 2, 3, 4] and references therein). Since in many problems the implementation of the ML strategy is prohibitively complicated, general tools, like the principle of *per-survivor processing* (PSP) [2] and the *expectation-maximization* (EM) algorithm [3, 4, 24, 25], have been proposed to devise quasioptimal receivers. The EM technique is an iterative algorithm generating the ML estimate of a set of deterministic unknown parameters, if properly initialized. It has been successfully applied to a number of problems and, in particular, to the ML detection of digital signals transmitted over fading channels [3, 4, 6, 26] and to carrier phase recovery [3, 7, 27, 28]. The EM algorithm, however, being a technique for ML estimation, is unable to incorporate any statistical information about the unknown parameters to be estimated, even if such information are available.

Recently, an extension of the EM, dubbed *Bayesian* EM (BEM), has been applied to solve MAP estimation problems and to derive SISO receivers [29, 30, 31, 32] for single-user detection over frequency-flat Rayleigh fading channels. The BEM algorithm allows to design SISO modules estimating the channel state, incorporating the symbol *a priori probabilities* (APRPs) and the statistics of the channel uncertainty, and generating the symbol APPs. Therefore, it can be easily employed in iterative equalization/decoding structures for coded transmissions [17, 23]. The favorable features of the BEM technique have suggested to further investigate its application to other communication scenarios.

This paper offers both a tutorial introduction to BEM-based estimation techniques and some recent research results about its applications. In fact, in its first part it describes the BEM technique, its relationship with the EM algorithm, and how it can be used to derive SISO algorithms for the detection of digital data transmitted over channels having memory and affected by parametric uncertainty. Then, in the second part of the paper, the application of the BEM approach to some detection problems of current interest is illustrated. In particular, we consider

(1) the multiuser detection of *direct sequence spread spectrum* (DSSS) signals in a synchronous CDMA system [33];
(2) the detection of single-carrier space-time block coded signals transmitted over frequency-flat fading channels [34];
(3) the detection of multicarrier space-time block coded signals transmitted over frequency-selective fading channels [35].

For each specific problem, in the third scenario, a BEM-based SISO algorithm is described and some numerical results are illustrated. Moreover, the use of a BEM-based SISO module in an iterative receiver is described in detail.

The paper is organized as follows. The EM and BEM techniques are described in Section 2. The use of the BEM technique to devise SISO modules for channels with parametric uncertainty and memory is illustrated in Section 3. Specific applications of the BEM tool are analyzed in Section 4. Finally, Section 5 offers some conclusions.

2. EXPECTATION-MAXIMIZATION ALGORITHMS FOR PARAMETER ESTIMATION

2.1. The EM algorithm

Let $\mathbf{\Theta} \doteq [\Theta_0, \Theta_1, \ldots, \Theta_{L-1}]^T$ denote an L-dimensional *deterministic* vector to be estimated from an N-dimensional received vector $\mathbf{R} \doteq [R_0, R_1, \ldots, R_{N-1}]^T$ of noisy data (with $N \geq L$).[1] The ML estimate of $\mathbf{\Theta}$ is the solution of the problem [36]

$$\boldsymbol{\theta}_{\mathrm{ML}} = \arg\max_{\tilde{\boldsymbol{\theta}}} L_{\mathbf{r}}(\tilde{\boldsymbol{\theta}}), \tag{1}$$

where $L_{\mathbf{r}}(\tilde{\boldsymbol{\theta}}) \doteq \log f(\mathbf{r}|\tilde{\boldsymbol{\theta}})$ is a log-likelihood function and $f(\mathbf{x}|\mathbf{y})$ denotes the *probability density function* (pdf) of the random vector $\mathbf{X}$ conditioned on the event $\{\mathbf{Y} = \mathbf{y}\}$. Solving the problem (1) in a direct fashion requires a closed form expression for $L_{\mathbf{r}}(\tilde{\boldsymbol{\theta}})$ but, even if this expression is available, the search for its maximum may entail an unacceptable computational burden. When this occurs, a feasible alternative can be offered by the EM algorithm [3, 25]. The EM approach develops from the assumption that a *complete* data vector $\mathbf{C} = [C_0, C_1, \ldots, C_{P-1}]^T$ (with $P \geq N$) is observed in place of the *incomplete* data set $\mathbf{R}$. The vector $\mathbf{C}$ is characterized by a couple of relevant properties: (1) it is not observed directly but, if available, would ease the estimation of $\mathbf{\Theta}$; (2) $\mathbf{R}$ can be obtained from $\mathbf{C}$ through a many-to-one mapping $\mathbf{C} \to \mathbf{R}(\mathbf{C})$. In practice, in communication problems, $\mathbf{C}$ is always chosen as a superset of the incomplete data [3], that is,

$$\mathbf{C} = \left[\mathbf{R}^T, \mathbf{I}^T\right]^T, \tag{2}$$

where the so-called *imputed* data $\mathbf{I}$ are properly selected to simplify the ML estimation problem [25]. In particular, when $\mathbf{\Theta}$ consists of all the transmitted channel symbols, $\mathbf{I}$ collects all the unwanted random parameters (fading, phase jitter, etc.) affecting the communication channel [3, 25]. These choices lead to *hard* detection algorithms often having an acceptable complexity and capable of incorporating the statistical properties of the channel parameters. In the following the complete data vector $\mathbf{C}$ will be always structured as in (2).

Given $\mathbf{C}$, the *auxiliary function*

$$\begin{aligned} Q_{\mathrm{EM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) &\doteq E_{\mathbf{I}}\{L_{\mathbf{c}}(\boldsymbol{\theta}) \,|\, \mathbf{R} = \mathbf{r},\ \mathbf{\Theta} = \tilde{\boldsymbol{\theta}}\} \\ &= E_{\mathbf{I}}\{\log f(\mathbf{C}|\boldsymbol{\theta}) \,|\, \mathbf{R} = \mathbf{r},\ \mathbf{\Theta} = \tilde{\boldsymbol{\theta}}\} \\ &= \int_{\mathbf{S_i}} \log f(\mathbf{r}, \mathbf{i}|\boldsymbol{\theta}) f(\mathbf{i} \,|\, \mathbf{r}, \tilde{\boldsymbol{\theta}}) d\mathbf{i} \end{aligned} \tag{3}$$

is evaluated, where $E_{\mathbf{X}}\{\cdot\}$ denotes the statistical average with respect to a random vector $\mathbf{X}$ and $\mathbf{S_i}$ is the space of $\mathbf{I}$. Then, this function is employed in the following two-step procedure generating successive approximations $\{\boldsymbol{\theta}^{(k)},\ k = 1, 2, \ldots\}$ of $\boldsymbol{\theta}_{\mathrm{ML}}$ (1):

(1) *expectation step*—$Q_{\mathrm{EM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$ in (3) is evaluated for $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_{\mathrm{EM}}^{(k)}$;
(2) *maximization step*—given $\boldsymbol{\theta}_{\mathrm{EM}}^{(k)}$, the next estimate $\boldsymbol{\theta}_{\mathrm{EM}}^{(k+1)}$ is computed as

$$\boldsymbol{\theta}_{\mathrm{EM}}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q_{\mathrm{EM}}\left(\boldsymbol{\theta}, \boldsymbol{\theta}_{\mathrm{EM}}^{(k)}\right), \quad k = 0, 1, \ldots. \tag{4}$$

An initial estimate $\boldsymbol{\theta}_{\mathrm{EM}}^{(0)}$ of $\boldsymbol{\theta}$ must be provided for the algorithm start-up. In digital communication problems, proper initialization of the EM algorithm is usually accomplished exploiting the information provided by known (pilot) symbols [3]. It can be proved that, under mild conditions, the sequence $\{\boldsymbol{\theta}_{\mathrm{EM}}^{(k)}\}$ converges to the true ML estimate $\boldsymbol{\theta}_{\mathrm{ML}}$ of (1), provided that the existence of local maxima does not prevent it. Avoiding this requires an accurate initial estimate $\boldsymbol{\theta}_{\mathrm{EM}}^{(0)}$ whose choice, for this reason, is critical [25].

[1] In the following, a random vector and its realizations are always denoted by an uppercase letter and the corresponding lowercase letter, respectively.

2.2. The BEM algorithm

The unknown vector $\mathbf{\Theta} = [\Theta_0, \Theta_1, \ldots, \Theta_{L-1}]^T$ mentioned in the previous paragraph can be also modeled as a *random* quantity, when its joint pdf $f(\boldsymbol{\theta})$ is available. In this case the MAP estimate $\boldsymbol{\theta}_{\text{MAP}}$ of $\mathbf{\Theta}$, given the observed data vector $\mathbf{r}$, can be evaluated as [36]

$$\boldsymbol{\theta}_{\text{MAP}} = \arg\max_{\tilde{\boldsymbol{\theta}}} M_{\mathbf{r}}(\tilde{\boldsymbol{\theta}}), \tag{5}$$

where $M_{\mathbf{r}}(\tilde{\boldsymbol{\theta}}) \doteq \log f(\mathbf{r}, \tilde{\boldsymbol{\theta}})$. Solving (5) may be a formidable task for the same reasons previously illustrated for the ML problem (1). In principle, however, an improved estimate of $\mathbf{\Theta}$ can be evaluated via the MAP approach since statistical information about channel uncertainty are exploited.

Since there is a strong analogy between the ML problem (1) and the MAP one (5), it is not surprising that an expectation-maximization procedure, dubbed *Bayesian* EM (BEM) [29, 37], for solving the latter, is available. The BEM algorithm evolves through the same iterative procedure as the EM, but with a different auxiliary function [29], namely,

$$\begin{aligned} Q_{\text{BEM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) &= E_{\mathbf{C}}\{M_{\mathbf{c}}(\boldsymbol{\theta}) \,|\, \mathbf{R} = \mathbf{r},\ \mathbf{\Theta} = \tilde{\boldsymbol{\theta}}\} \\ &= E\{\log f(\mathbf{C}, \boldsymbol{\theta}) \,|\, \mathbf{R} = \mathbf{r},\ \mathbf{\Theta} = \tilde{\boldsymbol{\theta}}\} \\ &= \int_{\mathbf{S_i}} \log f(\mathbf{r}, \mathbf{i}, \boldsymbol{\theta}) f(\mathbf{i} \,|\, \mathbf{r}, \tilde{\boldsymbol{\theta}}) d\mathbf{i}. \end{aligned} \tag{6}$$

A clear relationship can be established between the BEM and the EM algorithms. In fact, factoring the pdf $f(\mathbf{r}, \mathbf{i}, \boldsymbol{\theta})$ as

$$f(\mathbf{r}, \mathbf{i}, \boldsymbol{\theta}) = f(\mathbf{r}, \mathbf{i} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) \tag{7}$$

and substituting (7) into (6) produces

$$Q_{\text{BEM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) = Q_{\text{EM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) + I(\boldsymbol{\theta}), \tag{8}$$

where

$$I(\boldsymbol{\theta}) \doteq \log f(\boldsymbol{\theta}). \tag{9}$$

Equation (8) shows that the difference between $Q_{\text{BEM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$ (6) and $Q_{\text{EM}}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$ (3) is simply a *bias* term $I(\boldsymbol{\theta})$ (9) favoring the most likely values of $\mathbf{\Theta}$. It is worth noting that, if a priori information about $\mathbf{\Theta}$ were unavailable and, consequently, a uniform pdf was selected for $f(\boldsymbol{\theta})$, the contribution from $I(\boldsymbol{\theta})$ would turn into a constant in (8), that is, it could be neglected. Therefore, the BEM encompasses the EM as a special case and, since the former benefits by the statistical information about $\mathbf{\Theta}$, it is expected to provide improved accuracy with respect to the latter. For the same reason, an increase in the speed of convergence and an improved robustness against the choice of the initial conditions could be offered by the BEM.

3. SISO DATA DETECTION IN THE PRESENCE OF PARAMETRIC UNCERTAINTY VIA THE BEM TECHNIQUE

In this section we show how the BEM technique can be employed to derive SISO algorithms for detecting digital signals transmitted over channels with parametric uncertainty and memory. A single-user transmission over a *single-input single-output* channel is considered for simplicity, but, as shown in the following section, the proposed approach can be extended to an arbitrary number of users and to a *multiple-input multiple-output* (MIMO) system without any substantial conceptual problem.

Here we assume that the kth component of the received data vector $\mathbf{R}$ can be expressed as[2]

$$R_k = g_k(\mathbf{D}, \mathbf{A}) + N_k, \tag{10}$$

where $\mathbf{D} \doteq [D_0, D_1, \ldots, D_{N-1}]^T$ is a vector of independent channel symbols belonging to a constellation $\Sigma = \{s_0, s_1, \ldots, s_{M-1}\}$ of cardinality M and average energy E_s, $\mathbf{A} \doteq [A_0, A_1, \ldots, A_{L-1}]^T$ is a vector of random channel parameters *independent* of $\mathbf{D}$ and *with known statistical properties*, $\{N_k\}$ is an AWGN sequence with variance σ_N^2, and $g_k(\cdot, \cdot)$ expresses the known functional dependence of the channel on both the transmitted symbols and its parametric uncertainty. In particular, we concentrate on *conditional finite memory* channels, that is, on random channels such that

$$g_k(\mathbf{D}, \mathbf{A}) = g_k(D_k, D_{k-1}, D_{k-2}, \ldots, D_{k-L_c}, \mathbf{A}), \tag{11}$$

where L_c denotes the *channel memory*.

Our target is devising MAP SISO detection algorithms [18, 22], given the observed data $\mathbf{R} = \mathbf{r}$ and a statistically known parameter vector $\mathbf{A}$. In data detection problems involving the EM technique, two different choices have been usually suggested for the imputed data $\mathbf{I}$ (see (2)) and the parameter vector $\mathbf{\Theta}$:

(1) $\mathbf{I} = \mathbf{A}$ and $\mathbf{\Theta} = \mathbf{D}$ [3];
(2) $\mathbf{I} = \mathbf{D}$ and $\mathbf{\Theta} = \mathbf{A}$ [6, 8, 29].

It is extremely important to comment now on the meaning and the consequences of these choices.

In the first case, both EM and BEM-based algorithms aim at producing *hard* estimates of the transmitted data. The only substantial difference between these two classes of strategies is that BEM allows to exploit the data statistics, that is, their APRPs, in the detection algorithm, since $I(\boldsymbol{\theta})$ in (8) turns into (see (9))

$$I(\boldsymbol{\theta}) = I(\mathbf{D}) = \sum_{n=0}^{N-1} \log \Pr(d_n), \tag{12}$$

[2] Here we concentrate on detection algorithms processing one sample per channel symbol. The extension of the following ideas to multisampling detection is straightforward.

where $\Pr(d_n)$ denotes the probability of the event $\{D_n = d_n\}$. In other words, employing the EM (BEM) technique leads to *hard-in* (*soft-in*) *hard-output* detection algorithms.

In the second case, both EM- and BEM-based algorithms estimate the random parameters of the communication channel in a direct fashion. Nonetheless, they can be considered as SISO detectors, since they generate *soft* estimates (i.e., the APPs) of the transmitted data as a by-product of the estimation procedure and can also incorporate the data APRPs. BEM-based estimators, however, also make use of channel statistics, whereas EM-based estimators do not, that is, they operate in a *blind* fashion. Since blind detection techniques can be substantially outperformed by their counterparts exploiting channel statistics (see, e.g., [4, 38, 39]), this offers a strong motivation for preferring BEM-based strategies to EM-based ones when such statistical information are available. To further clarify these ideas, we derive now the BEM estimator of $\boldsymbol{\Theta} = \mathbf{A}$, given $\mathbf{I} = \mathbf{D}$. In (6) the joint pdf $f(\mathbf{r}, \mathbf{i}, \boldsymbol{\theta})$ can be factored as

$$f(\mathbf{r}, \mathbf{i}, \boldsymbol{\theta}) = f(\mathbf{r}, \mathbf{d}, \mathbf{a}) = f(\mathbf{r}|\mathbf{d}, \mathbf{a}) f(\mathbf{d}) f(\mathbf{a}) \tag{13}$$

as the data $\mathbf{D}$ are independent of the channel parameters $\mathbf{A}$. Here

$$f(\mathbf{d}) = \sum_{\mathbf{d}_l \in \Lambda} \Pr(\mathbf{d}_l) \boldsymbol{\delta}_N(\mathbf{d} - \mathbf{d}_l). \tag{14}$$

Λ is the set of all the M^N possible data sequences of length N, $\boldsymbol{\delta}_N(\cdot)$ is the N-dimensional Dirac delta function, and $\Pr(\mathbf{d}) = \prod_{n=0}^{N-1} \Pr(d_n)$ denotes the APRP of the channel symbol vector $\mathbf{d}$. If we define the *channel state* vector $\Delta_k \doteq (d_{k-1}, d_{k-2}, \ldots, d_{k-L_c})$, the conditional pdf $f(\mathbf{r}|\mathbf{d}, \mathbf{a})$ in (13) can be expressed as

$$f(\mathbf{r}|\mathbf{d}, \mathbf{a}) = \prod_{k=0}^{N-1} \frac{1}{\pi \sigma_N^2} \exp\left[-\frac{|r_k - g_k(d_k, \Delta_k, \mathbf{a})|^2}{\sigma_N^2} \right] \tag{15}$$

since the kth sample r_k depends on $\mathbf{d}$ through the couple (d_k, Δ_k) only, and the random variables $\{R_k\}$, conditioned on $\mathbf{D}$ and $\mathbf{A}$, are independent. Moreover, the conditional pdf $f(\mathbf{i}|\mathbf{r}, \tilde{\boldsymbol{\theta}})$ in (6) is given by

$$f(\mathbf{i} \,|\, \mathbf{r}, \tilde{\boldsymbol{\theta}}) = f(\mathbf{d} \,|\, \mathbf{r}, \tilde{\mathbf{a}}) = \sum_{\mathbf{d}_l \in \Lambda} \Pr(\mathbf{d}_l \,|\, \mathbf{r}, \tilde{\mathbf{a}}) \boldsymbol{\delta}_N(\mathbf{d} - \mathbf{d}_l), \tag{16}$$

where $\Pr(\mathbf{d}_l|\mathbf{r}, \tilde{\mathbf{a}})$ is the probability of the event $\{\mathbf{d} = \mathbf{d}_l\}$, given $\mathbf{R} = \mathbf{r}$ and $\mathbf{A} = \tilde{\mathbf{a}}$. Substituting (14) and (15) into (13) and substituting (13) and (16) into (6) and dropping the unrelevant terms produces, after some manipulations,

$$\begin{aligned} Q_{\mathrm{BEM}}&(\mathbf{a}, \tilde{\mathbf{a}}) \\ &= -\frac{1}{\sigma_N^2} \sum_{k=0}^{N-1} \sum_{\Delta_k \in \Pi} \sum_{d_k \in \Sigma} \Pr(d_k, \Delta_k \,|\, \mathbf{r}, \tilde{\mathbf{a}}) \,|\, r_k - g_k(d_k, \Delta_k, \mathbf{a}) \,|^2 \\ &\quad + \log f(\mathbf{a}), \end{aligned} \tag{17}$$

where Π denotes the set of M^{L_c} possible channel state vectors. We define now the estimate vector $\mathbf{a}[k] \doteq [a_0[k], a_1[k], \ldots, a_{L-1}[k]]^T$ generated, at the kth iteration, by the BEM estimation algorithm based on $Q_{\mathrm{BEM}}(\mathbf{a}, \tilde{\mathbf{a}})$ (17). Such an algorithm operates as follows. First, $Q(\mathbf{a}, \mathbf{a}[k])$ is evaluated (E step). The next estimate $\mathbf{a}[k+1]$ corresponds to the maximum of $Q(\mathbf{a}, \mathbf{a}[k])$ with respect to $\mathbf{a}$. Then, taking the gradient of (17) with respect to $\mathbf{a}$ and setting it to zero produces the recursive relation

$$\begin{aligned} \frac{1}{\sigma_N^2} \sum_{k=0}^{N-1} \sum_{\Delta_k \in \Pi} \sum_{d_k \in \Sigma} &\Pr(d_k, \Delta_k \,|\, \mathbf{r}, \mathbf{a}[k]) \\ &\times 2\,\mathrm{Re}\left\{ \left(g_k^*(d_k, \Delta_k, \mathbf{a}) - r_k^*\right) \right. \\ &\qquad \left. \times \nabla_{\mathbf{a}} g_k(d_k, \Delta_k, \mathbf{a}) \right\}_{\mathbf{a}=\mathbf{a}[k+1]} \\ &- \frac{1}{f(\mathbf{a})} \nabla_{\mathbf{a}} f(\mathbf{a}) \Big|_{\mathbf{a}=\mathbf{a}[k+1]} = 0 \end{aligned} \tag{18}$$

expressing a set of nonlinear equations for evaluating $\mathbf{a}[k+1]$, given $\mathbf{a}[k]$ (M-step). It is worth noting that complexity of solving (18) depends on the type of functional dependence of $g_k(\cdot)$ on $\mathbf{a}$ and on the inner structure of $\log f(\mathbf{a})$.

We us now explain why the estimator based on (18) can be also interpreted as a SISO algorithm. First of all, we note that the contribution from $\Pr(\mathbf{d}_l)$ (coming from (14)), being independent of $\mathbf{a}$, has been dropped in $Q_{\mathrm{BEM}}(\mathbf{a}, \tilde{\mathbf{a}})$ (17). The contribution from the APRPs of the channel symbols, however, has not really disappeared since such probabilities are used in the evaluation of the APPs $\{P(d_k, \Delta_k|\mathbf{r}, \tilde{\mathbf{a}})\}$. This means that, in its $(k+1)$th iteration, the BEM-based estimation algorithm requires the evaluation of the new APPs starting from the available APRPs and the last estimate $\mathbf{a}[k]$ of channel parameters. Generally speaking, on channels with memory, these APPs can be evaluated by means of a *forward-backward* recursive procedure operating on the trellis diagram of the channel states [6, 20, 40] and which can be derived as follows. To begin, we note that the couple (Δ_k, d_k) uniquely identifies a transition (Δ_k, Δ_{k+1}) in the channel state, so that $P(d_k, \Delta_k|\mathbf{r}, \tilde{\mathbf{a}}) = P(\Delta_k, \Delta_{k+1}|\mathbf{r}, \tilde{\mathbf{a}})$. Applying the Bayes' rule to the evaluation of $P(\Delta_k, \Delta_{k+1}|\mathbf{r}, \tilde{\mathbf{a}})$ gives

$$\begin{aligned} P(\Delta_k, \Delta_{k+1} \,|\, \mathbf{r}, \tilde{\mathbf{a}}) &= \frac{f(\mathbf{r}, \Delta_k, \Delta_{k+1} \,|\, \tilde{\mathbf{a}})}{f(\mathbf{r} \,|\, \tilde{\mathbf{a}})} \\ &= \frac{f(\mathbf{r}, \Delta_k, \Delta_{k+1} \,|\, \tilde{\mathbf{a}})}{\sum_{\tilde{\Delta}_k, \tilde{\Delta}_{k+1} \in \Pi} f(\mathbf{r}, \tilde{\Delta}_k, \tilde{\Delta}_{k+1} \,|\, \tilde{\mathbf{a}})}. \end{aligned} \tag{19}$$

Following [6, 20, 40] it can be proved that

$$\begin{aligned} f(\mathbf{r}, &\Delta_k, \Delta_{k+1} \,|\, \tilde{\mathbf{a}}) \\ &= \alpha_k(\Delta_k) f(r_k \,|\, \Delta_k, \Delta_{k+1}, \tilde{\mathbf{a}}) \beta_{k+1}(\Delta_{k+1}) \Pr(\Delta_{k+1} \,|\, \Delta_k) \end{aligned} \tag{20}$$

where $\mathbf{r}_j^l \doteq [r_j, r_{j+1}, \ldots, r_l]^T$, $\alpha_k(\Delta_k) \doteq f(\mathbf{r}_0^{k-1}, \Delta_k|\tilde{\mathbf{a}})$, $\beta_{k+1}(\Delta_{k+1}) \doteq f(\mathbf{r}_{k+1}^{N-1}|\Delta_{k+1}, \tilde{\mathbf{a}})$, $\Pr(\Delta_{k+1}|\Delta_k)$ is the probability of the state transition $\Delta_k \to \Delta_{k+1}$, and $f(r_k|\Delta_k, \Delta_{k+1}, \tilde{\mathbf{a}}) = [\pi \sigma_N^2]^{-1} \exp[-|r_k - g_k(d_k, \Delta_k, \tilde{\mathbf{a}})|^2 / \sigma_N^2]$. The quantities $\{\alpha_k(\Delta_k)\}$, and $\{\beta_{k+1}(\Delta_{k+1})\}$ are evaluated by means of the following recursive equations:

$$\alpha_k(\Delta_k) = \sum_{\tilde{\Delta}_{k-1}\in S(\tilde{\Delta}_{k-1},\Delta_k)} \alpha_{k-1}(\tilde{\Delta}_{k-1}) f(r_{k-1} \mid \Delta_k, \tilde{\Delta}_{k-1}, \tilde{\mathbf{a}}) \times \Pr(\Delta_k \mid \tilde{\Delta}_{k-1}), \tag{21}$$

$$\beta_{k+1}(\Delta_{k+1}) = \sum_{\tilde{\Delta}_{k+2}\in S(\Delta_{k+1},\tilde{\Delta}_{k+2})} \beta_{k+2}(\tilde{\Delta}_{k+2}) f(r_{k+1} \mid \Delta_{k+1}, \tilde{\Delta}_{k+2}, \tilde{\mathbf{a}}) \times \Pr(\tilde{\Delta}_{k+2} \mid \Delta_{k+1}), \tag{22}$$

where $S(\Delta_i, \Delta_j)$ is the subset of states Δ_i such that the transition $\Delta_i \to \Delta_j$ is admissible. The initial conditions $\{\alpha_0(\Delta_0) = \Pr(\Delta_0);\ \Delta_0 \in \Pi\}$ and $\{\beta_N(\Delta_N) = 1;\ \Delta_N \in \Pi\}$ need to be fixed before starting the forward (21) and the backward iterations (22), respectively.

After K iterations the BEM algorithm stops, producing a final estimate $\mathbf{a}_{\text{BEM}} = \mathbf{a}[K]$ and the APPs $\{\Pr(d_k, \Delta_k|\mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ of the channel symbols. The symbol APPs $\{\Pr(d_k|\mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ can be easily derived from these quantities as

$$\Pr(d_k \mid \mathbf{r}, \mathbf{a}_{\text{BEM}}) = \sum_{\Delta_k\in\Omega(d_k)} \Pr(d_k, \Delta_k \mid \mathbf{r}, \mathbf{a}_{\text{BEM}}), \tag{23}$$

where $\Omega(d_k)$ denotes the subset of all the state transitions $\Delta_k \to \Delta_{k+1}$ labeled by the channel symbol d_k. Then, decisions on the channel symbols can be taken according to the MAP decision strategy [6]

$$\hat{d}_k = \arg\max_{d_k} \Pr(d_k \mid \mathbf{r}, \mathbf{a}_{\text{BEM}}) \tag{24}$$

with $k = 0, 1, \ldots, N-1$. Alternatively, if channel coding is employed, the APPs $\{\Pr(d_k|\mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ can be delivered to soft decoding stages (see, e.g., [30, 31]) to improve the error performance of a digital receiver (see Section 4.4.3).

Finally, we note that substantial simplifications of the BEM-based procedure based on (18) can be found when the communication channel does not have memory, that is, $L_c = 1$, since in this case the forward-backward procedure is no more required. Specific examples of BEM-based algorithms for memoryless channels can be found in [30, 31, 32], where frequency-flat fading and phase jitter are considered as channel impairments.

4. SPECIFIC APPLICATIONS

In this section, three specific applications of the BEM strategy are briefly illustrated. In particular, SISO detectors are developed for the following three different scenarios: (1) a synchronous multiuser CDMA system; (2) a single-carrier system employing an orthogonal *space-time block code* (STBC); (3) an *orthogonal frequency division multiplexing* (OFDM) system using an orthogonal STBC on a subcarrier-by-subcarrier basis. For each scenario we provide a brief introduction citing a set of key references about the specific problem, a description of the signal and channel models, an analysis of the corresponding BEM-based SISO algorithm, and some numerical results.

4.1. Multiuser detection of synchronous DSSS signals over frequency-flat fading channels

4.1.1. Introduction

One of the most challenging problems in receiver design for DSSS-CDMA systems is the derivation of reduced-complexity multiuser detectors. This is due to the fact that the complexity of optimal multiuser detection grows exponentially with the number of users [41]. One of the interesting applications of the EM technique has been the derivation of multiuser detectors for synchronous DS-CDMA systems operating over frequency-flat fading channels [42, 43, 44]. However, all the solutions proposed in the cited papers produce hard estimates of the data. A BEM-based soft detector is illustrated in the following.

4.1.2. Channel and signal models

Multiuser detection on synchronous uplink of a J-user DS-CDMA system is considered here. In the presence of slow frequency-flat fading the output of the receiver matched filter bank in the lth symbol interval can be expressed as [42, 43]

$$\mathbf{Z}(l) = \mathbf{R}\mathbf{B}[l]\mathbf{A}[l] + \mathbf{N}[l], \tag{25}$$

where $\mathbf{Z}[l] \doteq [Z_1[l], \ldots, Z_J[l]]^T$, $\mathbf{B}[l] \doteq \text{diag}(B_1[l], \ldots, B_J[l])$ is the channel symbol matrix, $\mathbf{A}[l] \doteq [A_1[l], \ldots, A_J[l]]^T$ is the channel fading vector, $\mathbf{R} = [r_{mn}]$ $(m, n = 1, 2, \ldots, J)$ is the $J \times J$ matrix of signature cross-correlations, and $\mathbf{N}[l]$ is a complex Gaussian noise vector having zero mean and covariance matrix $\sigma_w^2\mathbf{R}$, with $\sigma_w^2 = 2N_0$. Here $B_j[l] \in \{\pm\sqrt{2E_{b,j}}\}$ ($E_{b,j}$ is the average transmitted energy per bit) is the BPSK channel symbol transmitted by the jth user in the lth signaling interval, $A_j[l]$ is the fading distortion affecting $B_j[l]$, and $r_{mn} = \int_0^{T_S} p_m(t)p_n(t)dt$ $(m, n = 1, 2, \ldots, J)$, where T_s is the symbol interval and $p_n(t)$ is the signature waveform[3] of the nth user. In the following it is assumed that the J fading processes $\{A_j[l]\}$ are independent, identically distributed and zero mean Gaussian (Rayleigh fading) with autocorrelation function $R_a[m]$ ($R_a[0] = 1$).

If $\mathbf{R}$ is positive definite, it can be Cholesky factored as $\mathbf{R} = \mathbf{\Gamma}^H\mathbf{\Gamma}$, where $\mathbf{\Gamma}$ is a lower triangular matrix. Then, premultiplying $\mathbf{Z}(l)$ (25) by $(\mathbf{\Gamma}^H)^{-1}$ produces [43]

$$\mathbf{Y}[l] = [Y_1[l], \ldots, Y_J[l]]^T \doteq (\mathbf{\Gamma}^H)^{-1}\mathbf{Z}[l] = \mathbf{C}\mathbf{B}[l]\mathbf{A}[l] + \mathbf{W}[l]. \tag{26}$$

Here the noise vector $\mathbf{W}[l] = [W_1[l], \ldots, W_J[l]]^T$ is white Gaussian since its covariance matrix is $\sigma_w^2\mathbf{I}_J$ ($\mathbf{I}_J$ is the $J \times J$ identity matrix).

Extending the one-shot model (26) to an observation interval of N consecutive symbols (with $l = 1, \ldots, N$) yields

$$\mathbf{Y} = \text{diag}(\mathbf{\Gamma})\mathbf{B}\mathbf{A} + \mathbf{W}, \tag{27}$$

[3] We assume that its support is the interval $[0, T_s]$.

where $\mathbf{Y} \doteq [\mathbf{Y}^T[1],\ldots,\mathbf{Y}^T[L]]^T$, $\mathbf{A} \doteq [\mathbf{A}^T[1],\ldots,\mathbf{A}^T[L]]^T$, $\mathbf{W} \doteq [\mathbf{W}^T[1],\ldots,\mathbf{W}^T[L]]^T$, and $\mathbf{B} \doteq \mathrm{diag}(\mathbf{B}[l],\ l = 1,2,\ldots, L)$ is an $NJ \times NJ$ block matrix having $\{\mathbf{B}[l]\}$ on its main diagonal. Following [45], we decompose the noise vector $\mathbf{W}[l]$ as $\sum_{j=1}^{J} \mathbf{W}_j[l]$, where $\{\mathbf{W}_j[l],\ l = 1,2,\ldots,N\}$ are independent Gaussian vectors having zero mean and covariance matrix $E\{\mathbf{W}_j[l]\mathbf{W}_j^H[l]\} = \sigma_{w,j}^2\mathbf{I}_J$, with $\sigma_{w,j}^2 = \beta_j\sigma_w^2$. Here $\{\beta_j\}$ are real positive coefficients satisfying the constraint $\sum_{j=1}^{J}\beta_j = 1$ in order to ensure statistical equivalence. Then, $\mathbf{Y}[l]$ (26) can be decomposed as $\sum_{j=1}^{J}\mathbf{U}_j[l]$, where

$$\mathbf{U}_j[l] = \left[U_1[l],\ldots,U_J[l]\right]^T \doteq \mathbf{\Gamma}_j b_j[l]a_j[l] + \mathbf{W}_j[l] \tag{28}$$

and $\mathbf{\Gamma}_j$ is the jth column ($j = 1,2,\ldots,J$) of $\mathbf{\Gamma}$.

4.1.3. The CDMA-BEM algorithm

We define now the vector $\mathbf{U} \doteq [\mathbf{U}^T[1],\ldots,\mathbf{U}^T[N]]^T$, with $\mathbf{U}[l] \doteq [U_1[l],\ldots,U_J[l]]^T$. Then, in applying the BEM technique, we select $\mathbf{C} = \{\mathbf{B},\mathbf{U}\}$ and $\mathbf{\Theta} = \mathbf{A}$ (see Section 2.2) as the complete and parameter vectors, respectively. This leads to the auxiliary function (further analythical details are available in [33])

$$\begin{aligned} Q(\mathbf{a},\tilde{\mathbf{a}}) &= \sum_{j=1}^{J}\sum_{l=1}^{N}\frac{1}{\sigma_{w,j}^2}\sum_{\tilde{\mathbf{b}}[l]\in\Omega} 2\,\mathrm{Re}\left\{\mathbf{\Gamma}_j^H\hat{\mathbf{u}}_j[l]a_j^*[l]\tilde{b}_j^*[l]\right\} \\ &\qquad\qquad\times \Pr\left(\tilde{\mathbf{b}}[l]\,|\,\mathbf{y},\tilde{\mathbf{a}}\right) \\ &\quad - \sum_{j=1}^{J}\sum_{l=1}^{N}\frac{2E_{b,j}}{\sigma_{w,j}^2}\left\|a_j[l]\right\|^2 - \sum_{j=1}^{J}\mathbf{a}_j^H\mathbf{C}_A^{-1}\mathbf{a}_j, \end{aligned} \tag{29}$$

where $\tilde{b}_j[l]$ is the jth component of $\tilde{\mathbf{b}}[l] = [\tilde{b}_1[l],\tilde{b}_2[l],\ldots,\tilde{b}_J[l]]^T$, $\Pr(\tilde{\mathbf{b}}[l]|\mathbf{y},\tilde{\mathbf{a}})$ is the probability of the event $\{\mathbf{b}[l] = \tilde{\mathbf{b}}[l]\}$ conditioned on $\mathbf{Y} = \mathbf{y}$ and $\mathbf{A} = \tilde{\mathbf{a}}$, and

$$\begin{aligned} \hat{\mathbf{u}}_j[l] &\doteq E\{\mathbf{u}_j[l]\,|\,\mathbf{b}[l] = \tilde{\mathbf{b}}[l],\mathbf{y},\tilde{\mathbf{a}}\} \\ &= \mathbf{\Gamma}_j\tilde{a}_j[l]\tilde{b}_j[l] + \beta_j\left(\mathbf{y}[l] - \sum_{i=1}^{J}\mathbf{\Gamma}_i\tilde{a}_i[l]\tilde{b}_i[l]\right). \end{aligned} \tag{30}$$

Given $Q(\mathbf{a},\tilde{\mathbf{a}})$ (29), the expectation-maximization can be expressed as follows [33]. Given the fading estimates $\mathbf{a}_j^k = [a_j^k[1],\ldots,a_j^k[N]]^T$, with $j = 1,2,\ldots,J$, at the kth iteration, the new estimate $\mathbf{a}_j^{k+1}$ is evaluated as

$$\mathbf{a}_j^{k+1} = \left(\mathbf{P}_j\right)^{-1}\mathbf{v}_j^k, \tag{31}$$

where

$$\mathbf{P}_j \doteq 2E_{b,j}\mathbf{I}_L + \sigma_{w,j}^2\mathbf{C}_A^{-1} \tag{32}$$

and $\mathbf{v}_j^k = [v_j^k[1],v_j^k[2],\ldots,v_j^k[L]]^T$, with

$$v_j^k[l] \doteq \sum_{\tilde{\mathbf{b}}[l]\in\Omega}\mathbf{\Gamma}_j^H\hat{\mathbf{u}}_j[l]\tilde{b}_j^*[l]\Pr\left(\tilde{\mathbf{b}}[l]\,|\,\mathbf{y},\tilde{\mathbf{a}}^k\right). \tag{33}$$

It is worth noting that the inverse of $\mathbf{P}_j$ (32) does not need to be recomputed as long as the channel statistics do not change, and that such matrix depends on j, that is, on the considered user, through $E_{b,j}$ and $\sigma_{w,j}^2$ only. The APPs $\Pr(\tilde{\mathbf{b}}[l]|\mathbf{y},\mathbf{a}^k)$ in (33) can be evaluated as

$$\begin{aligned} \Pr\left(\mathbf{b}[l] = \tilde{\mathbf{b}}[l]\,|\,\mathbf{y},\mathbf{a}^k\right) &= \frac{f\left(\mathbf{y}[l]\,|\,\tilde{\mathbf{b}}[l],\mathbf{a}^k[l]\right)\Pr\left(\tilde{\mathbf{b}}[l]\right)}{\sum_{\check{\mathbf{b}}[l]\in\Omega} f\left(\mathbf{y}[l]\,|\,\check{\mathbf{b}}[l],\mathbf{a}^k[l]\right)\Pr\left(\check{\mathbf{b}}[l]\right)}, \end{aligned} \tag{34}$$

where

$$\begin{aligned} f\left(\mathbf{y}[l]\,|\,\mathbf{b}[l],\mathbf{a}[l]\right) &= \frac{1}{\left(\pi\sigma_w^2\right)^J}\exp\left(\frac{-\left\|\mathbf{y}[l] - \mathbf{\Gamma}\mathbf{B}[l]\mathbf{A}[l]\right\|^2}{\sigma_w^2}\right). \end{aligned} \tag{35}$$

Moreover, the data APRP $\Pr(\mathbf{b}[l])$ of (34) can be expressed as

$$\Pr\left(\mathbf{b}[l]\right) = \prod_{j=1}^{J}\Pr\left(b_j[l]\right) \tag{36}$$

for the independence of the J users.

After K iterations the BEM-based algorithm based on (31)–(36) (dubbed CDMA-BEM in the following) stops producing a channel estimate $\mathbf{a}_{\mathrm{BEM}} = \mathbf{a}^{(K+1)}$ and the data APPs $\{P(b_j[l]|\mathbf{y},\mathbf{a}_{\mathrm{BEM}})\}$. Then, data decisions can be taken according to a MAP decision strategy (see (24)) or, if channel coding is used, can be delivered to soft decoding stages.

4.1.4. Numerical results

Computer simulations have been carried out in order to assess the *bit error rate* (BER) performance of the CDMA-BEM multiuser detector. In the following it is always assumed that (1) the autocovariance function of the fading process $\{A_j[l]\}$ (with $j = 1,\ldots,J$) is $R_a[m] = J_0(2\pi m B_D T_s)$ (Clarke's fading [46]), where $J_0(x)$ is the zeroth-order Bessel function of the first kind and B_D is the fading Doppler bandwidth; (2) each user continuously transmits packets containing $N = 14$ consecutive symbols; (3) each packet consists of 12 information symbols and is preceded by a couple of pilot symbols (used for channel estimation), so that the pilot symbol rate is $R_p = 1/7$; (4) Wiener filtering techniques are exploited at the receiver side in order to evaluate the channel estimates needed for the initialization of the CDMA-BEM [29]; (5) the CDMA-BEM processes a block of $(2\cdot N+2) = 30$ received signal samples corresponding to 2 consecutive packets (plus the first two samples of the next packet) and carries out $K = 3$ iterations; (6) the *signal-to-noise ratio* for the jth user (SNR_j) is defined as $E_{b,j}/N_0$, where $E_{b,j}$ is the average received energy per bit for the jth user and $N_0/2$ is the noise two-sided power spectral density; (7) the receiver is provided with an ideal estimate of the SNR for all the active users so that the parameters $\{\beta_j, j = 1,\ldots,J\}$ can be selected as [42]

$$\beta_j = \frac{E_{b,j}}{\sum_{i=1}^{J}E_{b,i}}. \tag{37}$$

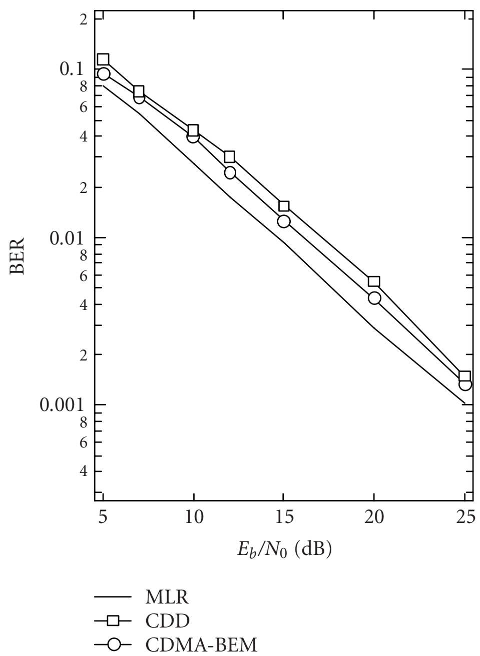

Figure 1: BER performance of the CDMA-BEM algorithm with $B_D T_s = 5 \cdot 10^{-3}$, $J = 4$, $N = 14$, and $K = 3$. The BER performance of the MLR and CDD is also shown for comparison.

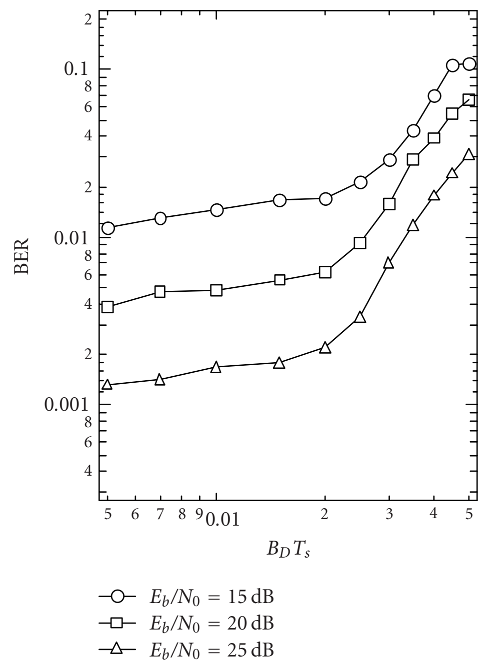

Figure 2: BER performance of the CDMA-BEM algorithm versus $B_D T_s$. $J = 4$, $E_{b,k}/N_0 = 20$ dB, $N = 14$, and $K = 3$.

In the following, we consider a four-user scenario ($J = 4$) characterized by the matrix of signature cross-correlations [43]:

$$\mathbf{R}_4 = \frac{1}{7}\begin{bmatrix} 7 & -1 & 3 & 3 \\ -1 & 7 & 3 & -1 \\ 3 & 3 & 7 & -1 \\ 3 & -1 & -1 & 7 \end{bmatrix}. \tag{38}$$

The BER performance of the CDMA-BEM receiver is illustrated in Figure 1. Here it is assumed that the normalized Doppler bandwith is $B_D T_s = 5 \cdot 10^{-3}$ and that all the users have the same SNR. In this figure the performance of the *maximum likelihood receiver* (MLR) endowed with ideal *channel state information* (CSI) and that of the *coherent decorrelator detector* (CDD) [47] are also shown for comparison. It is interesting to note that, in these scenarios, the CDMA-BEM almost achieves the same performance of the MLR and outperforms the CDD by about 1.5 dB in SNR.

Figure 2 shows the performance of CDMA-BEM versus the normalized Doppler bandwidth for $B_D T_s \in (5 \cdot 10^{-3}, 5 \cdot 10^{-2})$, under the assumption that $\text{SNR}_j = 15, 20, 25$ dB for $j = 1, \ldots, 4$. The error performance of the proposed algorithm slightly worsens as the Doppler bandwidth increases because of the poorer quality of the initial channel estimates.

Finally, the near-far resistance of the CDMA-BEM receiver is illustrated in Figure 3. The SNR of the first user (SNR_1) is set to 20 dB, whereas the other three SNRs (SNR_j, $j = 2, 3, 4$) are equal and vary in the range $(5, 25)$ dB. The performance of the MLR is also shown for comparison. These results show that, in this case, the CDMA-BEM exhibits a performance which is substantially independent of the energies of the interfering users.

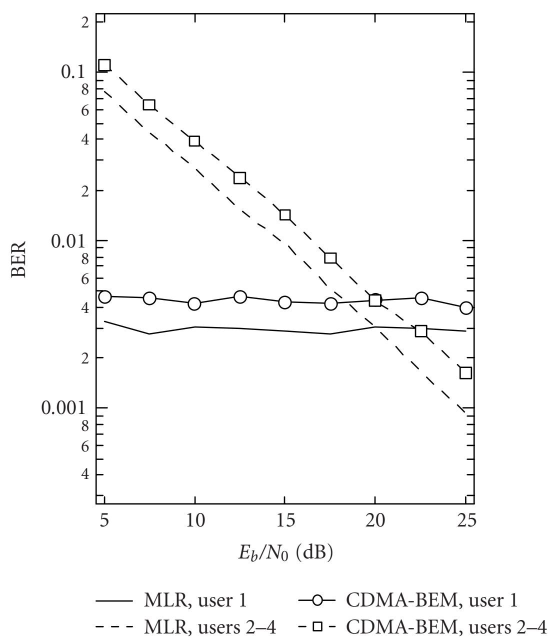

Figure 3: Near-far resistance of the CDMA-BEM algorithm. $J = 4$, $\text{SNR}_1 = 20$ dB, $\text{SNR}_k \in (5, 25)$ dB ($k = 2, 3, 4$), and $B_D T_s = 5 \cdot 10^{-3}$.

4.2. SISO detection of space-time block coded signals

4.2.1. Introduction

In the last years it has been shown that the information capacity of wireless communication systems can be substantially increased by employing antenna arrays [48], jointly with proper coding [49] and signal processing techniques [50]. One of the most promising results in this research area has been the development of new block and trellis codes for multiple antennas, known as *space-time codes* (STCs) [49, 51]. Such codes provide significant diversity gains without bandwidth expansion. Exact knowledge of the CSI is often assumed in devising space-time decoding algorithms even if channel estimation may represent a serious problem, especially in time-varying environments [52]. EM-based hard detectors for STCs have been derived in [52, 53, 54]. In this section a BEM-based soft detector for orthogonal STBCs is illustrated.

4.2.2. Signal and channel models

Here we focus on a space-time block coded system employing N_T transmit and N_R receive antennas [49]. The set of channel symbols transmitted during the nth block[4] is denoted by the $L \times N_T$ matrix $\mathbf{S}[n] = [s_{l,i}[n]]$ (with $l = 1, 2, \dots, L$, $i = 1, 2, \dots, N_T$), where L is the overall duration of the block in channel symbols and $s_{l,i}[n]$ is the channel symbol feeding the ith antenna in the symbol interval $(l + nL)$.

In the following we assume that the multiple channels involved in the communication system are (a) affected by frequency-flat Rayleigh fading and (b) *quasi-static*, that is, channel variations within each block are negligible, whereas changes from block to block are taken into account. Then the path gain $a_{i,j}[n]$ (with $i = 1, 2, \dots, N_T$ and $j = 1, 2, \dots, N_R$) from the ith transmit antenna to the jth receive antenna during the nth block is a complex Gaussian random process having zero mean and correlation function $R_a[m] \doteq E\{a_{i,j}[n+m]a_{i,j}^*[n]\}$ (with $R_a[0] = 1$). Moreover, the gain processes $\{a_{i,j}[n]\}$ are *independent* (rich scatterer environment).

Let $r_{l,j}[n]$ denote the received signal sample taken at the output of the jth receive antenna in the $(l + nL)$th symbol interval, with $j = 1, \dots, N_R$ and $l = 1, \dots, L$. Then the $L \times N_R$ received signal matrix $\mathbf{R}[n] = [r_{l,j}[n]]$ is given by [52]

$$\mathbf{R}[n] = \mathbf{S}[n]\mathbf{A}[n] + \mathbf{W}[n]. \tag{39}$$

Here $\mathbf{S}[n] \in \Omega$, where $\Omega = \{\mathbf{S}_m,\ m = 1, \dots, M\}$ is an M-ary alphabet of unitary matrices (i.e., $(\mathbf{S}_m)^H\,\mathbf{S}_m = \mathbf{I}_{N_T}$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix) [49, 51]. Moreover $\mathbf{A}[n] = [a_{i,j}[n]]$ and $\mathbf{W}[n] = [w_{l,j}[n]]$ are the $N_T \times N_R$ fading matrix and the $L \times N_R$ noise matrix, respectively. The elements $\{w_{l,j}[n]\}$ of $\mathbf{W}[n]$ are independent Gaussian random variables, all having zero mean and variance $\sigma_w^2 = 2N_0$.

[4]Throughout the section, the parameter n denotes the block index, whereas k specifies the location of a channel symbol within each block.

A set of N consecutive vectors (39) (with $n = 0, \dots, N-1$) can be grouped as $\mathbf{R} \doteq [\mathbf{R}^H[0], \mathbf{R}^H[1], \dots, \mathbf{R}^H[N-1]]^H$ ($(\mathbf{A})^T$ and $(\mathbf{A})^H$ denote transpose and conjugated transpose of $\mathbf{A}$, resp.), with

$$\mathbf{R} = \mathbf{D}(\mathbf{S})\mathbf{A} + \mathbf{W}, \tag{40}$$

where $\mathbf{A} \doteq [\mathbf{A}^H[0], \mathbf{A}^H[1], \dots, \mathbf{A}^H[N-1]]^H$ and $\mathbf{W} \doteq [\mathbf{W}^H[0], \mathbf{W}^H[1], \dots, \mathbf{W}^H[N-1]]^H$, respectively, and $\mathbf{D}(\mathbf{S}) \doteq \mathrm{diag}\{\mathbf{S}[0], \mathbf{S}[1], \dots, \mathbf{S}[N-1]\}$.

4.2.3. A BEM-based SISO algorithm for space-time block coded systems

Following the same indications illustrated in the previous application, we set $\mathbf{\Theta} = \mathbf{A}$ and $\mathbf{C} = \{\mathbf{R}, \mathbf{S}\}$ in applying the BEM technique. Then the auxiliary function is (analytical details can be found in [55])

$$\begin{aligned} Q(\mathbf{A}, \tilde{\mathbf{A}}) = &- \sum_{j=1}^{N_R} \mathbf{A}_j^H \Big[\mathbf{C}_{\mathbf{A}}^{-1} + \Big(\frac{1}{\sigma_w^2}\Big)\mathbf{I}_{NN_T}\Big]\mathbf{A}_j \\ &- \Big(\frac{2}{\sigma_w^2}\Big) \mathrm{Re}\,\{\tilde{\mathbf{V}}_j{}^H \mathbf{A}_j\}, \end{aligned} \tag{41}$$

where $\mathbf{A}_j$ is the jth column of $\mathbf{A}$, $\mathbf{C}_{\mathbf{A}} \doteq E\{\mathbf{A}_j\mathbf{A}_j^H\}$ is a fading covariance matrix, and $\tilde{\mathbf{V}}_j$ is the jth column of the matrix

$$\tilde{\mathbf{V}} \doteq \mathbf{D}^H(\tilde{\mathbf{S}})\mathbf{R} \tag{42}$$

with $\tilde{\mathbf{S}} = \{\tilde{\mathbf{S}}[n],\ n = 0, 1, \dots N-1\}$. Here

$$\tilde{\mathbf{S}}[n] = \sum_{\mathbf{S}_m \in \Omega} \mathbf{S}_m \Pr\,(\mathbf{S}[n] = \mathbf{S}_m \,|\, \mathbf{R}, \tilde{\mathbf{A}}), \tag{43}$$

where $\Pr(\mathbf{S}[n] = \mathbf{S}_m|\mathbf{R}, \tilde{\mathbf{A}})$ is the APP of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$, given $\mathbf{R}$ and $\mathbf{A} = \tilde{\mathbf{A}}$. Starting from (41), the following BEM-based recursive channel estimator can be derived. Given the channel estimate $\mathbf{A}^{(k)}$ at the kth iteration, the next estimate $\mathbf{A}^{(k+1)}$ is evaluated as

$$\mathbf{A}_j^{(k+1)} = [\mathbf{P}]^{-1}\mathbf{V}_j^{(k)}, \tag{44}$$

where $\mathbf{P} \doteq \mathbf{I}_{NN_T} + \sigma_w^2 \mathbf{C}_A^{-1}$. The APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m|\mathbf{R}, \tilde{\mathbf{A}})\}$ needed for the evaluation of (42) can be computed using the Bayes formula

$$\Pr\,(\mathbf{S}[n] = \mathbf{S}_m \,|\, \mathbf{R}, \tilde{\mathbf{A}}) = \frac{f\,(\mathbf{R}[n] \,|\, \mathbf{S}_m, \tilde{\mathbf{A}}[n])\, \Pr\,(\mathbf{S}_m)}{\sum_{\tilde{\mathbf{S}}_m \in \Omega} f\,(\mathbf{R}[n] \,|\, \tilde{\mathbf{S}}_m, \tilde{\mathbf{A}}[n])\, \Pr\,(\tilde{\mathbf{S}}_m)}, \tag{45}$$

where $\Pr(\mathbf{S}_m)$ is the probability of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$, and

$$\begin{aligned} &f\,(\mathbf{R}[n] \,|\, \mathbf{S}_m, \tilde{\mathbf{A}}[n]) \\ &\quad = \frac{1}{\det\,(\pi\sigma_w^2 \mathbf{I}_L)^{N_R}} \exp\Big[-\frac{h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{A}}[n])}{\sigma_w^2}\Big] \end{aligned} \tag{46}$$

with $h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{A}}[n]) \doteq \mathrm{tr}\{(\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])^H(\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])\}$.

It is important to note that (a) $\mathbf{P}$ does not depend on the index of the receive antenna; (b) the inverse of $\mathbf{P}$ does not need to be recomputed as long as the channel statistics do not change; (c) (44) can be simplified factoring $\mathbf{C_A}$ as

$$\mathbf{C_A} = \tilde{\mathbf{C}}_\mathbf{a} \otimes \mathbf{I}_{N_T}, \tag{47}$$

where $\tilde{\mathbf{C}}_\mathbf{a}$ is the covariance matrix of the vector $\mathbf{a}_{i,j} = [a_{i,j}[0], a_{i,j}[1], \ldots, a_{i,j}[N-1]]^T$ and $\otimes$ is the Kronecker product, so that $\mathbf{P} = (\mathbf{I}_N + \sigma_w^2 \tilde{\mathbf{C}}_\mathbf{a}^{-1}) \otimes \mathbf{I}_{N_T}$.

After K iterations the BEM algorithm stops producing a channel estimate $\mathbf{A}_{\text{BEM}} = \mathbf{A}^{(K)}$ and the APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{A}_{\text{BEM}})\}$ which can be processed exactly like in the previous application. In the following the BEM-based estimation algorithm (43)–(46) is dubbed STBC-BEM.

4.3. *Numerical results*

The error performance of the STBC-BEM algorithm has been assessed by computer simulation for the Alamouti's space-time block code [51]. Then we have

$$\mathbf{S}[n] \doteq \begin{bmatrix} s_n^1 & s_n^2 \\ -(s_n^2)^* & (s_n^1)^* \end{bmatrix}, \tag{48}$$

where the symbols $\{s_n^1, s_n^2\}$ belong to a BPSK constellation.[5] In the following we assume that (1) $R_a[m] = J_0(2\pi m L B_D T)$, where $J_0(x)$ is the zeroth-order Bessel function of the first kind, B_D is the fading Doppler bandwidth, and T is the signaling interval; (2) the SNR is defined as E_b/N_0, where E_b is the average received energy per receive antenna and information bit; (3) each *packet* of $(N_B - 1)$ consecutive information blocks is followed by one pilot block, so that the pilot symbol rate is $R_p = 1/N_B$.

The STBC-BEM algorithm processes a sample set $\mathbf{R}$ consisting of $N \cdot L$ consecutive received signal samples, corresponding to N transmitted symbol blocks. It is assumed that the first and last L samples of $\mathbf{R}$ always correspond to a pilot block. This entails that (a) $N = N_p N_B + 1$, if N_p packets are processed, and (b) the last block of each set is in common with the first of the next one. The information provided by the pilot symbols is exploited to initialize the BEM algorithm. In particular the initial channel estimate for the jth receive antenna is evaluated as $\mathbf{A}_j = \mathbf{F}\mathbf{R}_j$, where $\mathbf{R}_j$ is the jth column of $\mathbf{R}$, with $j = 1, 2, \ldots, N_R$. Here $\mathbf{F}$ is an optimal $NN_T \times NL$ matrix that can be easily derived by standard methods (Wiener filtering) [29, 36], under the assumptions that (a) the information channel symbols are independent and identically distributed and (b) the pilot symbols are exactly known.

In all the following results it is assumed that the BEM algorithm processes $N_p = 4$ consecutive packets, each consisting of $N_B = 10$ consecutive blocks.

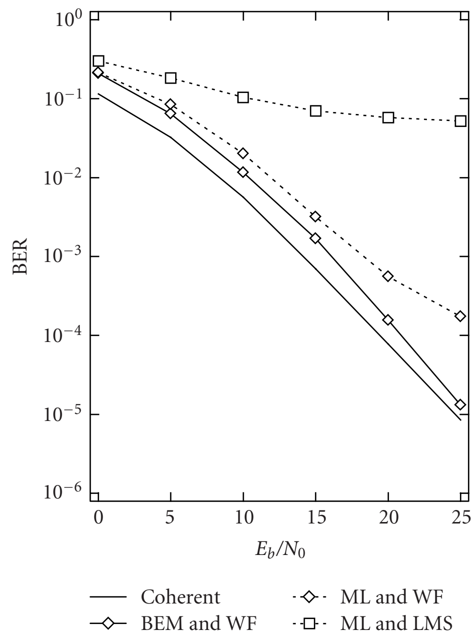

FIGURE 4: BER performance of various detection algorithms with Alamouti's STBC. $N_R = 1$ and $B_D T = 2 \cdot 10^{-2}$.

In Figure 4 the error performance of the STBC-BEM (with $K = 3$) is compared with that provided by an ML receiver using WF channel estimation[6] and an ML receiver using decision-directed *least mean square* (LMS) channel tracking with step size $\mu = 0.5$ (the tracker is initialized for each packet using the pilot block at its beginning in order to avoid runaway problems) for single receive diversity ($N_R = 1$) and $B_D T = 2 \cdot 10^{-2}$. The BER performance of a coherent receiver endowed with ideal CSI is also shown. These results evidence that (1) since the energy loss due to pilot symbols is 0.45 dB, the BEM performs very well if the fading rate is not too large; (2) the BEM substantially outperforms the other detectors. Further simulations have also shown that a blind SISO detector based on the EM-based approach illustrated in [6] and initialized by a WF does not outperform the ML detector endowed with the same channel estimator.

Figure 5 shows the error performance of the STBC-BEM with a different number of iterations, that is, with $K = 1$, 2, and 3, in the same scenario as the previous figure. These results evidence the usefulness of running three full iterations in the BEM procedure, in order to approach the performance of a coherent receiver endowed with ideal CSI. We also found, however, that negligible gains are offered by $K > 3$.

The comments already expressed about the results of Figure 4 also apply to Figure 6, referring to double receive diversity ($N_R = 2$), channel estimation based on WF and $B_D T = 5 \cdot 10^{-3}$, 10^{-2}, and $2 \cdot 10^{-2}$ for the BEM ($B_D T = 2 \cdot 10^{-2}$ only is considered for the ML detector). This figure

[5] Further results (not shown for space limitations) evidence that the comments expressed for a BPSK system also apply to larger constellations.

[6] Its error performance coincides with that offered by the BEM without iterations.

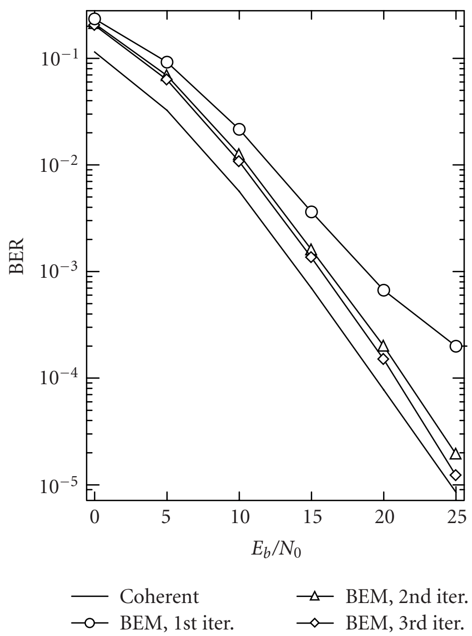

Figure 5: BER performance of the BEM detection algorithm with Alamouti's STBC. The error performance of the coherent detector is also shown for comparison. $N_R = 1$, $B_DT = 2\cdot10^{-2}$, and $K = 1$, 2, and 3.

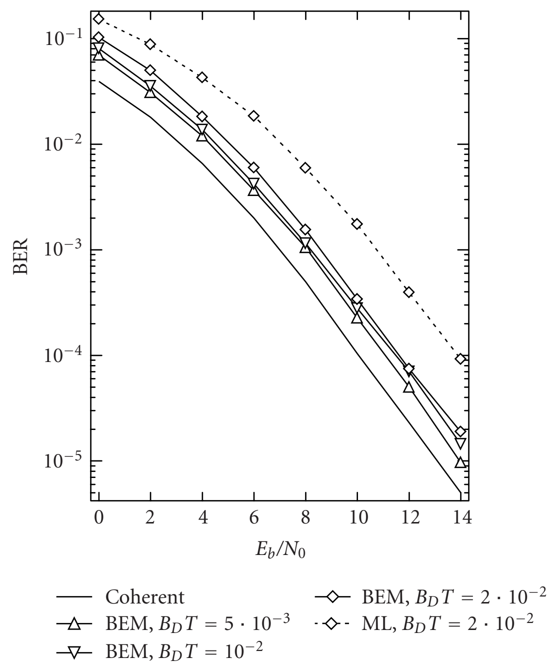

Figure 6: BER performance of various detection algorithms with Alamouti's STBC. $N_R = 2$.

also evidences that the BEM performance is not substantially affected by a change in the Doppler rate, provided that $B_DT \leq 2\cdot10^{-2}$.

In Figure 7 the BEM and the ML detector BER versus the normalized Doppler bandwidth B_DT is shown for $B_DT \in (10^{-2}, 5\cdot10^{-2})$ and $E_b/N_0 = 10$ dB (WF is used in both cases). It is worth noting that the performance degradation increases for larger Doppler bandwidths as the quality of the initial estimate of the BEM becomes poorer and this prevents BEM convergence to the global maximum, at least over some data blocks. Simulation results have also evidenced that, in this case, increasing the number of BEM iterations provides a negligible improvement.

4.4. SISO detection of space-time block coded OFDM signals

4.4.1. Introduction

The use of OFDM is often suggested to simplify channel equalization in the presence of appreciable frequency selectivity. When employed in MIMO wireless systems, the OFDM technique can be also easily combined with channel codes devised for multiple transmit antennas, that is, with *space-time* (ST) codes. A further improvement in the system performance can be achieved when conventional outer channel codes, like *convolutional codes* [56, 57] or *low-density parity-check* (LDPC) codes [58], are used in conjunction with proper ST symbol mappers.

Decoding of ST codes usually requires an accurate knowledge of CSI at the receiver. In MIMO OFDM systems, however, channel estimation may represent a serious problem, especially in time-varying environments, because of the high complexity needed to achieve a satisfying accuracy [59], even if simplified pilot-based channel estimators can be devised [60]. Recently, it has been shown that, when OFDM is combined with ST block coding [51] and a pilot-based channel estimate is available at the receiver, the EM technique can be applied to devise accurate channel estimators [61] and that such estimators can be used for *soft-in hard-output* detection [54]. In the last case, hard decisions are then converted to soft data information which can be exploited in iterative receiver architectures when outer coding is employed at the transmitter. In this part we tackle the same problem, but from a different perspective. In fact, we derive a SISO module based on the BEM technique. Preliminary simulation results suggest that this algorithm offers better performance than that derived in [54] with a lower overall computational burden.

4.4.2. Signal and channel models

In this paper we consider an ST block coded OFDM system employing N subcarriers jointly with N_T transmit and N_R receive antennas. The block diagram of the communication system is illustrated in Figure 8a. The coding scheme results from the concatenation of a convolutional or an LDPC code with an orthogonal STBC. It is worth noting that that LDPC codes have some relevant properties [62], like low decoding complexity and excellent performance, which make them a promising coding technique for ST coded OFDM systems [58].

The input bit stream is partitioned into blocks, each independently encoded by means of a channel encoder. After (optional) bit interleaving (Π) the coded bits are mapped

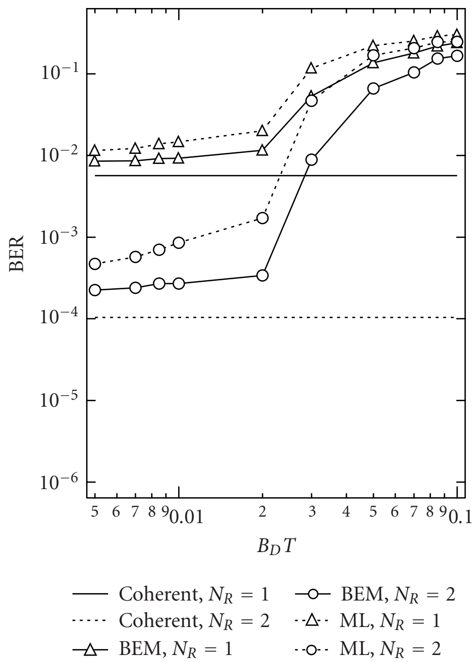

FIGURE 7: BER versus the normalized Doppler bandwidth $B_D T$ for various detection algorithms with STBC. $E_b/N_0 = 10$ dB. $N_R = 1$ and 2.

into channel symbols belonging to an M-ary PSK constellation. The resulting symbol sequence feeds an ST orthogonal block encoder. In the following, we consider, for simplicity, the Alamouti's STBC [51], even if the proposed detection algorithm can be easily extended to any orthogonal ST block code. The output sequence of the ST encoder is passed through a bank of N_T *inverse discrete Fourier transform* (IDFT) processors, which generate an ST-OFDM codeword spanning L OFDM symbol intervals. For instance, with Alamouti's STBC, we have $L = 2$ and, if $c_0[l,n]$ and $c_1[l,n]$ denote the channel symbols transmitted on the nth OFDM subcarrier (with $n = 0,\ldots,N-1$) in the lth OFDM symbol interval (with l *even*) by the first and the second transmit antenna, respectively, then $c_0[l+1,n] = -c_1^*[l,n]$ and $c_1[l+1,n] = c_0^*[l,n]$ are sent in the next symbol interval. In other words, the resulting codeword associated with the nth subcarrier is represented by the matrix

$$\mathbf{S}[n] = \begin{bmatrix} c_0[l,n] & c_1[l,n] \\ c_0[l+1,n] & c_1[l+1,n] \end{bmatrix} \tag{49}$$

belonging to an alphabet $\Omega = \{\mathbf{S}_p,\ p = 1,\ldots,P\}$ (with $P = M^2$) of unitary matrices [51].

The OFDM signal is transmitted over a *wide sense stationary uncorrelated scattering* (WSS-US) MIMO channel [63]. In the following it is assumed that (a) all the single-input single-output channels associated with different transmit/receive antenna pairs are mutually independent, identically distributed and are affected by Rayleigh fading; (b) in the propagation scenario, frequency dispersion is independent of time dispersion. Under these hypotheses a full statistical description of the MIMO channel is provided by its *power delay profile* (PDP) and its Doppler *power density spectrum* (PDS) or, equivalently, by its frequency correlation function $R_H(f)$ and its time correlation function $R_D(t)$, respectively [63]. At the receiver (see Figure 8b) a bank of N_R DFT processors (one per receive antenna) is fed by N_R distinct discrete-time signal sequences produced by matched-filtering and symbol-rate sampling. The outputs of the DFTs are processed by a BEM-based SISO detection algorithm (see the following paragraph) operating on a codeword-by-codeword basis. For this reason, in the following, we concentrate on the detection of a single ST-OFDM codeword. In particular, if $r_j[l,n]$ denotes the received signal sample taken at the output of the jth DFT for the nth subcarrier frequency in the lth OFDM symbol interval, with $j = 0,1,\ldots,N_R-1$ and $n = 0,\ldots,N-1$, we always take a couple of consecutive received signal samples for $l = 0,2,4,\ldots$. If we assume that the fading process remains constant over an ST codeword (i.e., over two adjacent OFDM symbol intervals with Alamouti's STBC), the $L \times N_R$ matrix $\mathbf{R}[l,n] = [r_j[l,n]]$ collecting the received signal samples over the observation interval for the nth subcarrier can be expressed as [54]

$$\mathbf{R}[l,n] = \mathbf{S}[l,n]\mathbf{H}[l,n] + \mathbf{W}[l,n]. \tag{50}$$

Here, $\mathbf{S}[n,l]$ is the $L \times N_T$ transmitted codeword matrix (see (49)), $\mathbf{H}[l,n] = [H_{i,j}[l,n]]$ is an $N_T \times N_R$ channel response matrix ($H_{i,j}[l,n]$ represents the complex channel gain between the ith transmit and the jth receive antenna at the nth subcarrier frequency), and $\mathbf{W}[l,n] = [w_{l,j}[l,n]]$ is an $L \times N_R$ noise matrix. The elements $\{w_{l,j}[l,n]\}$ of $\mathbf{W}[l,n]$ are independent complex zero mean Gaussian random variables with variance $\sigma_w^2 = 2N_0$. We also note that $\{H_{i,j}[l,n]\}$ are complex Gaussian random variables with zero mean and that the correlation between $H_{i,j}[l,n+m]$ and $H_{i,j}[l,n]$ is given by $R_H[m] = E\{H_{i,j}[l,n+m]H_{i,j}^*[l,n]\} = R_H(mf_\Delta)$, where f_Δ is the subcarrier spacing.

For a given l, the matrices (50) associated with all the different subcarriers ($n = 0,\ldots,N-1$) can be grouped in an $LN \times N_R$ matrix $\mathbf{R}[l] \doteq [\mathbf{R}^H[l,0], \mathbf{R}^H[l,1],\ldots,\mathbf{R}^H[l,N-1]]^H$. If the dependence on l is dropped, for simplicity, this vector can be expressed as

$$\mathbf{R} = \mathbf{D}(\mathbf{S})\mathbf{H} + \mathbf{W}, \tag{51}$$

where $\mathbf{H} \doteq [\mathbf{H}^H[0], \mathbf{H}^H[1],\ldots,\mathbf{H}^H[N-1]]^H$, $\mathbf{W} \doteq [\mathbf{W}^H[0], \mathbf{W}^H[1],\ldots,\mathbf{W}^H[N-1]]^H$, and $\mathbf{D}(\mathbf{S}) \doteq \mathrm{diag}\{\mathbf{S}[0], \mathbf{S}[1],\ldots, \mathbf{S}[N-1]\}$.

4.4.3. A BEM-based SISO algorithm for OFDM systems

Following the same approach as the previous two scenarios, we choose $\mathbf{\Theta} = \mathbf{H}$ and $\mathbf{I} = \mathbf{S}$. Then, as shown in [35], the BEM auxiliary function (6) can be expressed as

$$Q(\mathbf{H}, \tilde{\mathbf{H}}) = -\sum_{j=1}^{N_R} \mathbf{H}_j^H \mathbf{M} \mathbf{H}_j - \frac{2}{\sigma_w^2} \mathrm{Re}\left\{\tilde{\mathbf{V}}_j^H \mathbf{H}_j\right\}, \tag{52}$$

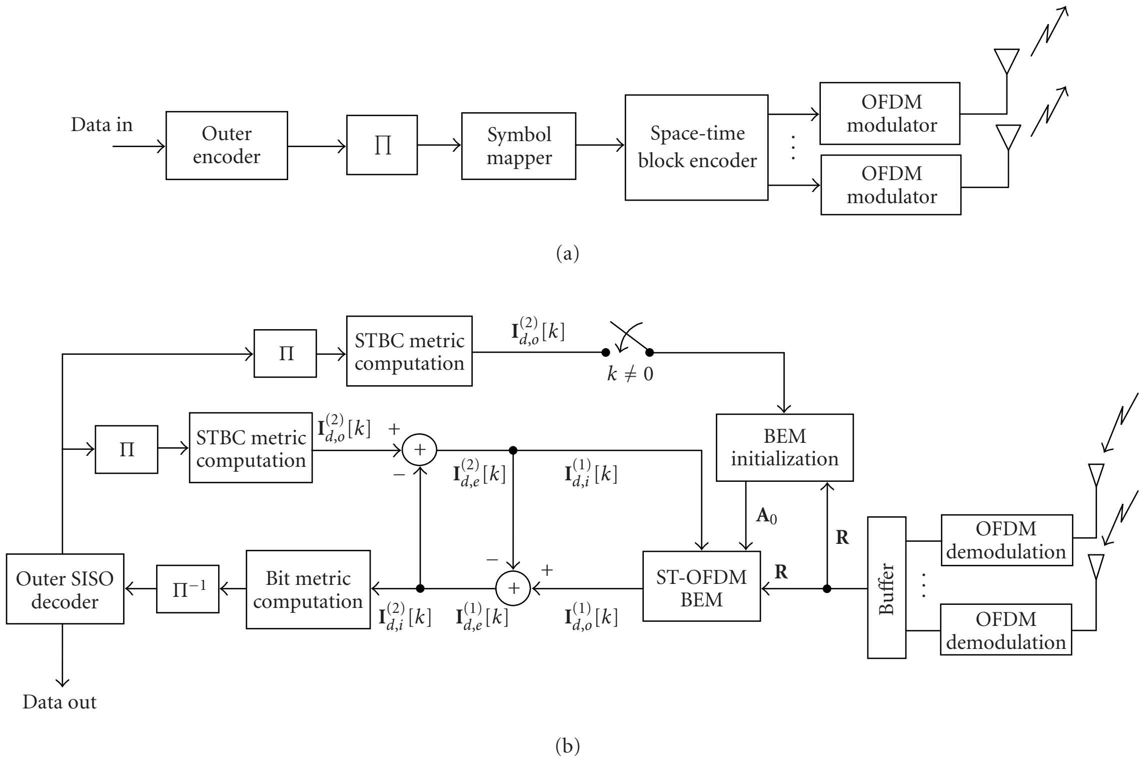

Figure 8: Block diagrams of the space-time block coded OFDM: (a) transmitter and (b) receiver.

where $\mathbf{H}_j$ is the jth column of the matrix $\mathbf{H}$, $\mathbf{M} \doteq \mathbf{C}_{\mathbf{H}}^{-1} + (1/\sigma_w^2)\mathbf{I}_{NI}$ with $\mathbf{C}_{\mathbf{H}} = E\{\mathbf{H}_j\mathbf{H}_j^H\}$, $\tilde{\mathbf{V}}_j$ is the jth column of the matrix $\tilde{\mathbf{V}} \doteq \mathbf{D}^H(\tilde{\mathbf{S}})\mathbf{R}$, and the matrix $\tilde{\mathbf{S}}$ results from the ordered concatenation of the matrices $\{\tilde{\mathbf{S}}[n],\ n = 0, 1, \ldots, N-1\}$, with

$$\tilde{\mathbf{S}}[n] \doteq \sum_{\mathbf{S}_m \in \Omega} \mathbf{S}_m \Pr\left(\mathbf{S}[n] = \mathbf{S}_m \mid \mathbf{R}, \tilde{\mathbf{H}}\right). \tag{53}$$

The APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{H}})\}$ can be evaluated using the Bayes formula

$$\Pr\left(\mathbf{S}[n] = \mathbf{S}_m \mid \mathbf{R}, \tilde{\mathbf{H}}\right) = \frac{f\left(\mathbf{R}[n] \mid \mathbf{S}_m, \tilde{\mathbf{H}}[n]\right) P(\mathbf{S}_m)}{\sum_{\tilde{\mathbf{S}}_m \in \Omega} f\left(\mathbf{R}[n] \mid \tilde{\mathbf{S}}_m, \tilde{\mathbf{H}}[n]\right) P(\tilde{\mathbf{S}}_m)}, \tag{54}$$

where

$$f\left(\mathbf{R}[n] \mid \mathbf{S}_m, \tilde{\mathbf{H}}[n]\right) = C_R \exp\left[-\frac{h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{H}}[n])}{\sigma_w^2}\right] \tag{55}$$

with $C_R \doteq \det(\pi\sigma_w^2\mathbf{I}_L)^{-N_R}$ and $h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{H}}[n]) \doteq \operatorname{tr}\{(\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])^H \cdot (\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])\}$. The BEM algorithm operates as follows. Given the channel estimate $\mathbf{H}^{(k)}$ at the kth iteration, the next estimate $\mathbf{H}^{(k+1)}$ is evaluated as

$$\mathbf{H}_j^{(k+1)} = \mathbf{P}^{-1}\mathbf{V}_j^{(k)} \tag{56}$$

with $\mathbf{P} \doteq \mathbf{I}_{NN_T} + \sigma_w^2\mathbf{C}_{\mathbf{H}}^{-1}$ and $j = 1, 2, \ldots, N_R$. It is important to note that (a) $\mathbf{P}$ does not depend on the index of the receive antenna; (b) the inverse of $\mathbf{P}$ does not need to be recomputed as long as the channel statistics do not change; (c) (56) can be simplified factoring $\mathbf{C}_{\mathbf{H}}$ as

$$\mathbf{C}_{\mathbf{H}} = \tilde{\mathbf{C}}_{\mathbf{H}} \otimes \mathbf{I}_{N_T}, \tag{57}$$

where $\tilde{\mathbf{C}}_{\mathbf{H}}$ is the covariance matrix of the vector $\mathbf{H}_{i,j} = [H_{i,j}[0], H_{i,j}[1], \ldots, H_{i,j}[N-1]]^T$ and $\otimes$ is the Kronecker product, so that $\mathbf{P} = (\mathbf{I}_N + \sigma_w^2\tilde{\mathbf{C}}_{\mathbf{H}}^{-1}) \otimes \mathbf{I}_{N_T}$. In the following the BEM-based estimation algorithm (53)–(56) is dubbed ST-OFDM BEM.

After K iterations the BEM algorithm stops producing a channel estimate $\mathbf{H}_{\text{BEM}} = \mathbf{H}^{(K)}$ and the APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})\}$. These can be exploited to take MAP decisions or for soft decoding of an outer code in a concatenated scheme. In our work, we have considered the iterative receiver structure as shown in Figure 8b. This structure operates as follows. After OFDM demodulation, the ST-OFDM BEM module takes as input the received signal vector $\mathbf{R} \doteq [\mathbf{R}^H[0], \mathbf{R}^H[1], \ldots, \mathbf{R}^H[N-1]]^H$, an initial channel estimate matrix $\mathbf{H}^{(0)}$ (consisting of $N \cdot N_T \times N_R$ matrices $\mathbf{H}^{(0)}[n]$) and the $N \times P$ a priori information matrices $\{\mathbf{I}_{d,i}^{l(1)}[k] = [(I_{d,i}^{l(1)}[k])_{n,m}]\}$. Here $(I_{d,i}^{l(1)}[k])_{n,m} = \log \Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m)$, where $\Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m)$ denotes the APRP that $\mathbf{S}[n]$ is equal to the mth codeword of the alphabet Ω at the kth step. After K iterations the BEM algorithm produces the $N \times P$ output matrices $\{\mathbf{I}_{d,o}^{l(1)}[k] = [(I_{d,o}^{l(1)}[k])_{n,m}]\}$ with $(I_{d,o}^{l(1)}[k])_{n,m} = \log \Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})$, where $\Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})$ represents the APP of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$ at the kth step. Then the *extrinsic information* matrices $\{\mathbf{I}_{d,e}^{l(1)}[k]\}$ are evaluated as $\mathbf{I}_{d,e}^{l(1)}[k] = \mathbf{I}_{d,o}^{l(1)}[k] - \mathbf{I}_{d,i}^{l(1)}[k]$. Since interleaving is performed at the bit level, before sending the extrinsic information to the deinterleaver (Π^{-1}) and to the SISO decoder, the evaluation of the *soft* bit metrics is needed (see [64, Section II-C]). The channel decoder produces the a posteriori *bit* information matrices and, after bit interleaving and probability recombination, the a posteriori *symbol* information matrices $\{\mathbf{I}_{d,o}^{l(2)}[k]\}$ (in log form). Finally, at the last iteration, the SISO decoder computes the APP matrix $\{\mathbf{P}_b\}$ together with a bit estimate vector. Subtracting $\{\mathbf{I}_{d,i}^{l(2)}[k]\}$ from $\{\mathbf{I}_{d,o}^{l(2)}[k]\}$ produces the extrinsic information matrices $\{\mathbf{I}_{d,e}^{l(2)}[k]\}$ of the channel symbols which are fed back as input to the ST-OFDM BEM decoder.

In our simulations both convolutional and LDPC codes have been employed. With convolutional codes the bit APRPs produced by the ST-OFDM BEM feed a *Bahl Cocke Jelinek Raviv* (BCJR) algorithm [20] implemented in its log MAP form [65]. With LDPC codes bit *log-likelihood ratios* (LLRs) are evaluated on the basis of the bit APRPs and sent to an LDPC decoder based on the *belief propagation* (BP) algorithm [62, 66]. It is important to point out that (a) the parity check matrices of the LDPC codes employed in our work have been generated in a random fashion [67], avoiding cycles of length 4 in the code graph in order to improve the code distance properties; (b) due to the random generation of the encoding matrix, no external interleaver (deinterleaver) is needed at the output (input) of the LDPC encoder (decoder) [58].

Finally, we note that, in the proposed receiver structure, the APPs $\{\mathbf{I}_{d,o}^{l(2)}[k]\}$, after interleaving, are also used to evaluate the estimate $\mathbf{H}^{(k+1)}$ needed for the initialization of the ST-OFDM BEM in the $(k+1)$th iteration of the receiver. At the beginning of the first iteration, however, no *a priori* information on the channel symbols is available. For this reason the initial fading estimate $\mathbf{H}^{(0)}$ of the ST-OFDM BEM is evaluated by means of the pilot-based channel estimation algorithm derived in [60].

4.5. Numerical results

In this paragraph some BER results are illustrated. In our computer simulations the reduced complexity model for WSS-US channels proposed in [68] has been used for the generation of a MIMO multipath fading channel. In particular, for a given Doppler bandwidth B_D, the Doppler PDS has been defined as $S_D(f) = 1 - 1.72 f_0^2 + 0.785 f_0^4$ for $f_0 \le 1$ and $S_D(f) = 0$ for $f_0 > 1$, where $f_0 = f/B_D$ [69]. Moreover, the multipath MIMO channel has been modeled as a 3-tap delay line approximating an exponential PDP $P_h(\tau) = \tau_0^{-1} \exp(-\tau/\tau_0) u(\tau)$, with $\tau_0 = 1.56$ microseconds (the corresponding frequency correlation function is $R_H(f) = 1/(1 + j2\pi f \tau_0)$). Then, in accordance with the OFDM physical layer specifications for the broadband radio access networks (BRAN) in [70], the following parameters have been selected for the ST block coded OFDM system: (a) the DFT order is $N = 256$; (b) the number of useful OFDM subcarriers is equal to 192, since the total number of subcarriers N includes 27 suppressed carriers on the upper frequencies, 28 suppressed carriers on the lower frequencies, 8 BPSK pilot symbols, and 1 DC carrier set to 0; (c) the OFDM symbol interval is $T_S = 0.125$ microseconds; (d) the length of the cyclic prefix in the OFDM modulator has been set to 64; (d) with convolutional codes, a 4-state rate 1/2 convolutional code with generators $g_1 = (5)_8$ and $g_2 = (7)_8$ has been adopted, when used; (e) with LDPC codes, a regular (3,6) code with rate $R = 1/2$ and a BP algorithm with a maximum number of iterations equal to 10 have been adopted, when used; (f) QPSK modulation has been employed for both uncoded and coded transmission; (g) a single frame consists of 9 ST block coded OFDM information codewords plus one pilot ST block coded OFDM codeword appended at its beginning. Moreover a single receive antenna, that is, $N_R = 1$ and a Doppler bandwidth $B_D = 200$ Hz have been chosen for our simulations.

In addition, the following assumptions have been made at the receive side: (a) the SNR is defined as E_b/N_0, where E_b is the average captured energy per receive antenna and information bit; (b) the BEM algorithm processes a block consisting of 192 Alamouti's space-time block codewords, and accomplishes $K = 3$ complete iterations; (c) the last channel estimate generated by the the BEM algorithm for each ST-OFDM codeword is used as an initial estimate of the same algorithm for the next codeword.

Figure 9 shows the ST-OFDM BEM algorithm performance without outer channel coding. Comparison is made with an ML detector endowed with ideal CSI (*genie bound*) and with an ML detector endowed with the same pilot-based *channel estimator* (CE) as the BEM [60]. These results evidence that the ST-OFDM BEM algorithm substantially outperforms a realistic ML detector. We also note that the energy loss due to pilot symbol insertion is 0.45 dB, so that the energy gap between the genie bound and the ST-OFDM BEM is about 1.5 dB [71].

Some simulation results referring to a convolutionally encoded system are shown in Figure 10, comparing the BER performance provided by the iterative receiver described in the previous paragraph (with 0, 1, and 2 iterations) with that offered by a BCJR decoder endowed with ideal CSI. We have also considered a receiver structure in which the likelihoods produced by the above-mentioned ML detector with pilot-based CE are exploited to generate soft data information feeding, after deinterleaving, the SISO outer decoder. The proposed iterative architecture substantially outperforms the latter and, if the energy loss due to pilot symbols is neglected, it approaches closely the genie bound. It is also worth noting that, in this scenario, carrying out global iterations provides a very small gain. This result can be explained as follows. The ST-OFDM BEM, starting from a pilot-based channel estimate, produces a good channel estimate and a good estimate of the data APPs since the beginning, that is, even in the absence of the APRPs produced by the BCJR, despite

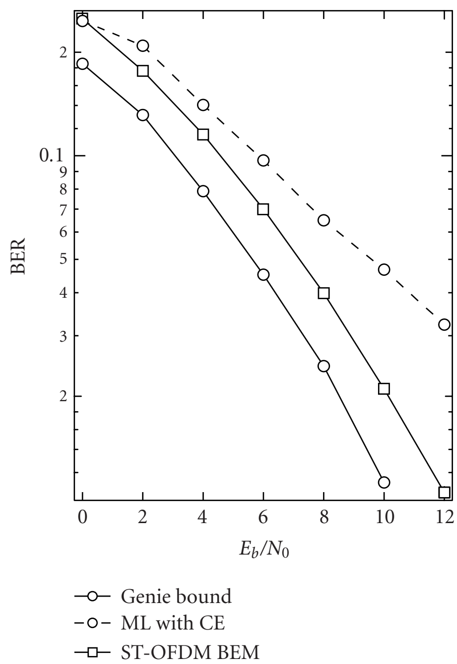

FIGURE 9: BER performance of the ST-OFDM BEM algorithm without outer coding. $N_R = 1$ and $K = 3$.

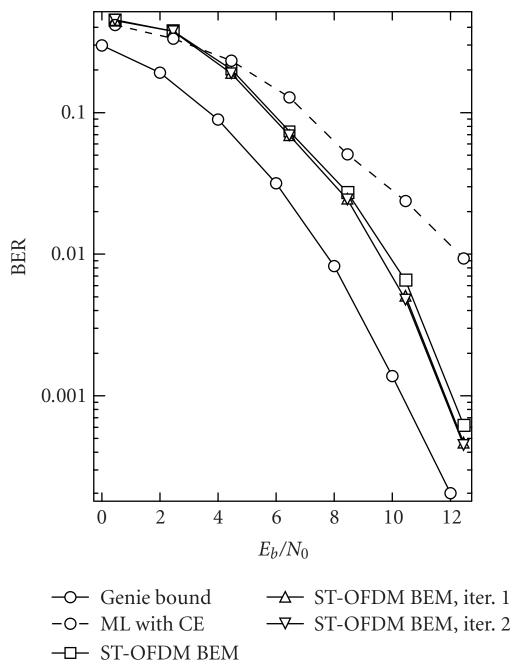

FIGURE 10: BER performance of the ST-DFDM BEM iterative receiver. Convolutional coding, $N_R = 1$, and $K = 3$.

the appreciable Doppler rate. These results are substantially different than those illustrated in [54, page 223], evidencing, for instance, a strong gap between the performance in the absence of iterations and that achieved after one iteration and suggesting the use of 3–5 global iterations. On the basis of these preliminary results, since the complexity (per iteration)

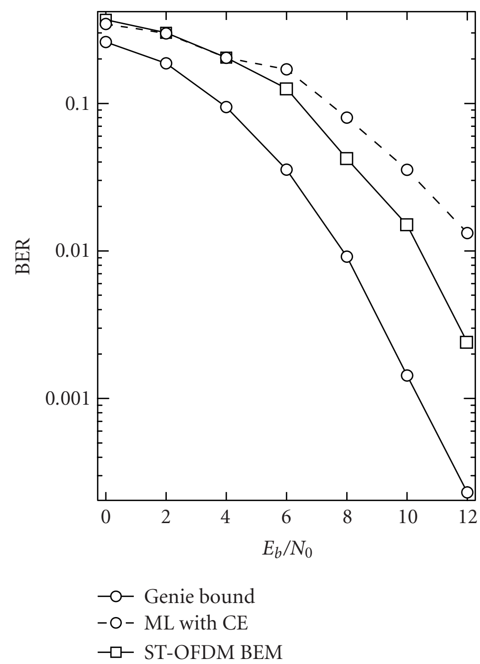

FIGURE 11: BER performance of the ST-OFDM BEM receiver. LDPC coding, $N_R = 1$, and $K = 3$.

of the ST-OFDM BEM and that of the EM algorithm derived in [54] are comparable, the use of the former should be preferred to the latter, since it ensures faster convergence, that is, a smaller overall complexity.

Finally, in Figure 11 the performance of the ST-OFDM BEM receiver for LDPC-coded signals is illustrated. The BER performance of the proposed algorithm is compared with that obtained by a BP algorithm endowed with perfect CSI. The curve labeled as “ML with CE” represents the BER performance of an ML detector endowed with pilot-based CE and followed by the LDPC decoder. Even without turbo decoding , the ST-OFDM BEM algorithm brings a substantial gain against the ML-based symbol detection approach. It is worth noting that, in this scenario, the BER performances given by the LDPC and convolutional coding schemes are widely comparable. This poor behavior obtained by LDPC coding is mainly due to the small dimension of the parity-check matrix employed in our simulations.

5. CONCLUSIONS

In this paper the BEM technique has been proposed to solve MAP estimation problems. In particular, we have shown that it represents a useful tool to derive novel SISO detectors for communication channels with random parametric uncertainty and memory. As an application of these concepts, SISO modules for the iterative detection of coded digital signals transmitted over fading channels have been derived in three specific scenarios and their error performance has been assessed. Applications of the BEM technique to other communication scenarios are the subject of ongoing research activities.

ACKNOWLEDGMENTS

This work has been performed in the framework of the project STINGRAY IST-2000-30173, which is funded by the European Community. The authors would like to acknowledge the contributions of their colleagues from the Intracom Hellenic Telecommunications and Electronics Industry S.A., the University of Modena and Reggio Emilia, the Institute of Accelerating Systems and Applications, the Technical Research Centre of Finland, and the National Technical University of Athens.

REFERENCES

[1] D. P. Taylor, G. M. Vitetta, B. D. Hart, and A. Mämmelä, "Wireless channel equalization," *European Transactions on Telecommunications*, vol. 9, no. 2, pp. 117–143, 1998.

[2] R. Raheli, A. Polydoros, and C.-K. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environment," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 354–364, 1995.

[3] C. N. 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, 1997.

[4] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1181–1193, 1999.

[5] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1638–1649, 2000.

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

[7] G. K. Kaleh, "Joint carrier phase estimation and symbol decoding of trellis codes," *European Transactions on Telecommunications*, vol. 4, no. 2, pp. 157–163, 1993.

[8] C. A. Haro, J. A. R. Fonollosa, and J. Fonollosa, "Blind channel estimation and data detection using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 241–247, 1997.

[9] W. Turin, "Map decoding in channels with memory," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 757–763, 2000.

[10] L. Davis, I. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, 2001.

[11] P. Hoeher and J. Lodge, "Turbo DPSK: iterative differential DPSK demodulation and channel decoding," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 837–843, 1999.

[12] I. D. Marsland and P. T. Mathiopoulos, "Multiple differential detection of parallel concatenated convolutional (turbo) codes in correlated fast Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 265–274, 1998.

[13] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 788–799, 1997.

[14] J. Hagenauer, "The turbo principle: tutorial introduction and state of the art," in *Proc. International Symposium on Turbo Codes & Related Topics*, pp. 1–11, Brest, France, September 1997.

[15] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 429–445, 1996.

[16] B. Sklar, "A primer on turbo code concepts," *IEEE Commun. Magazine*, vol. 35, no. 12, pp. 94–102, 1997.

[17] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, 1996.

[18] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *IEEE Communications Letters*, vol. 1, no. 1, pp. 22–24, 1997.

[19] Y. Lin, B. Vucetic, and Y. Sato, "Optimum soft-output detection for channels with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 704–713, 1995.

[20] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[21] V. Franz and J. B. Anderson, "Concatenated decoding with a reduced-search BCJR algorithm," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 186–195, 1998.

[22] S. Benedetto, G. Montorsi, D. Divsalar, and F. Pollara, "Soft-input, soft-output modules for the construction and distributed iterative decoding of code networks," *European Transactions on Telecommunications*, vol. 9, no. 2, pp. 155–172, 1998.

[23] C. Douillard, A. Picart, M. Jézéquel, P. Didier, C. Berrou, and A. Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Transactions on Telecommunications*, vol. 6, no. 5, pp. 507–511, 1995.

[24] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[25] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, no. 6, pp. 47–60, 1996.

[26] H. Z. Jafarian and S. Pasupathy, "Em-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1297–1302, 1999.

[27] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, 2004.

[28] N. Noels, C. Herzet, A. Dejonghe, et al., "Turbo synchronization: an EM algorithm interpretation," in *Proc. IEEE International Conference on Communications (ICC '03)*, pp. 2933–2937, Anchorage, Alaska, USA, May 2003.

[29] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation in the presence of random parameters via a generalized EM algorithm," in *Proc. IEEE International Conference on Communications (ICC '01)*, vol. 4, pp. 1057–1061, Helsinki, Finland, June 2001.

[30] E. Chiavaccini and G. M. Vitetta, "A BEM-based detector for CPM signals transmitted over frequency-flat fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 409–412, 2003.

[31] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency-flat Rayleigh fading channels via a Bayesian EM algorithm," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1869–1872, 2001.

[32] E. Chiavaccini and G. M. Vitetta, "A per-survivor phase-estimation algorithm for detection of PSK signals," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2059–2061, 2001.

[33] A. S. Gallo, E. Chiavaccini, and G. M. Vitetta, "BEM-based multiuser detection for synchronous DS-CDMA systems over frequency flat fading channels," in *Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication Conference (PIMRC '02)*, vol. 3, pp. 1285–1289, Lisbon, Portugal, September 2002.

[34] E. Chiavaccini, A. S. Gallo, and G. M. Vitetta, "BEM-Based SISO detection of space-time block coded signals transmitted over frequency-flat fading channels," in *Proc. IEEE Global*

Telecommunications Conference (GLOBECOM '02), vol. 1, pp. 374–378, Taipei, Taiwan, November 2002.

[35] A. S. Gallo and G. M. Vitetta, "A novel soft-in soft-out detection algorithm for space-time concatenated coding schemes operating over multipath fading channels," submitted to the *IEEE Global Telecommunications Conference (GLOBECOM '04)*, 2004.

[36] C. W. Helstrom, *Elements of Signal Detection & Estimation*, Prentice-Hall, Englewwod Cliffs, NJ, USA, 1995.

[37] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, UK, 1995.

[38] G. M. Vitetta, D. P. Taylor, and U. Mengali, "Double-filtering receivers for PSK signals transmitted over Rayleigh frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 44, no. 6, pp. 686–695, 1996.

[39] G. M. Vitetta, U. Mengali, and D. P. Taylor, "Blind detection of CPM signals transmitted over frequency-flat fading channels," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 961–968, 1998.

[40] R. W. Chang and J. C. Hancock, "On receiver structures for channels having memory," *IEEE Trans. Inform. Theory*, vol. 12, no. 4, pp. 463–468, 1966.

[41] S. Verdù, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.

[42] A. Kocian and B. H. Fleury, "Iterative joint symbol detection and channel estimation for DS/CDMA via the sage algorithm," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '00)*, vol. 2, pp. 1410–1414, London, UK, September 2000.

[43] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: an EM-based approach," *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1700–1710, 1996.

[44] D. Raphaeli, "Suboptimal maximum-likelihood multiuser detection of synchronous CDMA on frequency-selective multipath channels," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 875–885, 2000.

[45] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.

[46] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell System Technical Journal*, vol. 47, no. 6, pp. 957–1000, 1968.

[47] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 2nd edition, 1989.

[48] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[49] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility and multiple paths," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 199–207, 1999.

[50] A. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, no. 6, pp. 49–83, 1997.

[51] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.

[52] C. Cozzo and B. L. Hughes, "Joint channel estimation and data symbol detection in space-time communications," in *Proc. IEEE International Conference on Communications (ICC '00)*, vol. 1, pp. 287–291, New Orleans, La, USA, June 2000.

[53] Y. Li, C. N. Georghiades, and G. Huang, "Iterative maximum-likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948–951, 2001.

[54] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Transactions on Wireless Communications*, vol. 1, no. 2, pp. 213–225, 2002.

[55] A. S. Gallo, E. Chiavaccini, and G. M. Vitetta, "A novel BEM-based multiuser detector for co-channel signals transmitted over frequency-flat fading channels," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '03)*, pp. 241–246, San Francisco, Calif, USA, December 2003.

[56] Y. Gong and K. B. Letaief, "Concatenated space-time block coding with trellis coded modulation in fading channels," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 580–590, 2002.

[57] T. H. Liew and L. Hanzo, "Space-time codes and concatenated channel codes for wireless communications," *Proc. IEEE*, vol. 90, no. 2, pp. 187–219, 2002.

[58] B. Lu, X. Wang, and K. R. Narayan, "LDPC-based space-time coded OFDM systems over correlated fading channels," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, 2002.

[59] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 461–471, 1999.

[60] W. G. Jeon, K. H. Paik, and Y. S. Cho, "An efficient channel estimation technique for OFDM systems with transmitter diversity," *IEICE Transactions Communications*, vol. E84-B, no. 4, pp. 967–974, 2001.

[61] E. Panayirci and H. A. Cirpan, "Channel estimation for space-time block coded OFDM systems in the presence of multipath fading," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 2, pp. 1157–1161, Taipei, Taiwan, November 2002.

[62] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, 2001.

[63] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions Communication Systems*, vol. 11, no. 4, pp. 360–393, 1963.

[64] G. Taricco, G. Caire, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927–946, 1998.

[65] P. Robertson, E. Villebrun, and P. A. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. IEEE International Conference on Communications (ICC '95)*, vol. 2, pp. 1009–1013, Seattle, Wash, USA, June 1995.

[66] R. G. Gallager, "Low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[67] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, 1999.

[68] E. Chiavaccini and G. M. Vitetta, "GQR Models for multipath Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 1009–1018, 2001.

[69] V. Erceg, K. V. S. Hari, M. S. Smith, et al., "Channel models for fixed wireless applications," *IEEE 802.16.3 Task Group Contributions, Doc. IEEE 802.16.3.c-01/29r4*, 2001, available at http://www.ieee802.org/16/tg3.

[70] ETSI, "Broadband Radio Access Networks (BRAN)—HIPERMAN, OFDM Physical (PHY) Layer," draft, 2002.

[71] The STINGRAY Consortium, "Benchmarking and Performance Criteria for STC-OFDM," *Deliverable 3D3* (IST-2000-30173 STINGRAY), 2003.

A. S. Gallo was born in Nettuno, Rome, Italy, in 1974. He received the Dott. Ing. degree (cum laude) in electrical engineering from the University of Rome "La Sapienza," Italy, in 2000, defending a thesis on optimisation techniques for neuro-fuzzy networks. In 2004 he received the Ph.D. degree in information engineering from the University of Modena and Reggio Emilia, Italy. His research interests lie in the areas of wireless communication and signal processing, with main emphasis on low-cost detection/equalization techniques, space-time coding for wideband communication systems, and neuro-fuzzy networks.

G. M. Vitetta was born in Reggio Calabria, Italy, in April 1966. He received the Dr.-Ing. degree in electronic engineering (cum laude) in 1990 and the Ph.D. degree in 1994, both from the University of Pisa, Italy. In 1992/1993, he spent a period at the University of Canterbury, Christchurch, New Zealand, doing research for digital communications on fading channels. From 1995 to 1998, he was a Research Fellow at the Department of Information Engineering of the University of Pisa. From 1998 to 2001, he held the position of Associate Professor of telecommunications at the University of Modena and Reggio Emilia. He is now a Full Professor of telecommunications in the same university. His main research interests lie in the broad area of communication theory, with particular emphasis on coded modulation, synchronization, statistical modeling of wireless channels and channel equalization. He is serving as an Editor of both the IEEE Transactions on Communications and the IEEE Transactions on Wireless Communications.