IJICTR
International Journal of Information &
Communication Technology Research

Technical Note

Volume 10- Number 1 – Winter 2018 (62-69)

Enhanced Iterative Detection of Hierarchically Modulated Signals using VB-EM Algorithm

Mahsa Azad
Department of Electrical Engineering,
Shahid Beheshti University G.C.,
Tehran 1983969411, Iran
ma.azad@mail.sbu.ac.ir

Seyed Mohammad Sajad Sadough*
Department of Electrical Engineering,
Shahid Beheshti University G.C.,
Tehran 1983969411, Iran
s_sadough@sbu.ac.ir

Received: 27 September, 2017 - Accepted: 13 February, 2018

Abstract— **In Hierarchical Modulation (HM), the binary data is partitioned into a "high-priority" (HP) and a "low-priority" (LP) bit stream that are separately and independently encoded before being mapped on non-uniformly spaced constellation points. Indeed, HM is a multi-level modulation scheme, proposed for enabling unequal error protection. The HP and LP streams vary in their susceptibility to noise. The basic HP signal is more robust, in other words, heavily protected against noise and interference, whereas the LP stream has much less robustness. In this paper, we propose an improved iterative receiver based on variational Bayesian expectation-maximization (VB-EM) scheme for semi-blind joint channel estimation and data detection based on bit-interleaved coded modulation (BICM), with the aim of improving the bit error rate (BER) of the noisy LP stream. Moreover, the proposed receiver allows us to reduce the pilot symbol overhead compared to the classical HM receivers based on pilot-assisted channel estimation.**

Keywords- Hierarchical modulation; BICM; Iterative Receiver; Variational Bayesian Expectation-Maximization (VB-EM) algorithm.

I. Introduction

Conventional wireless communication systems are commonly designed for fixed bit-rate applications. However, recent and emerging applications such as multimedia and web services, require supporting various data rates. With hierarchical modulation (HM) [1], users can receive signals with different priority levels, depending on the receiver's channel quality. For instance, high resolution receivers can detect HDTV signals, while receivers with lower resolution and in bad channel conditions, can receive at least a basic signals. In this regard, HM is mainly proposed for unequal error protection where the initial stream is divided into a higher priority (HP) and a lower priority (LP) stream. The HP signal is very robust to noise while the LP signal which generally contains fine details of the data, is received only by receivers in favorable channel conditions. A main challenge for practical deployment of HM, is the necessity of very high signal-to-noise ratio (SNR) in order to detect the LP stream [2].

In [3], the authors focus on using HM in modern broadcast systems to increase the transmission rate. In [4], the authors show how the constellation in HM can be dynamically adapted at the physical layer in order to reduce the necessary SNR for a target BER. Constellation points are optimized in [5] with optimal binary labels that are designed for HM working over AWGN channel. The error performance of the HP stream is improved in [6] by proposing layered offset hierarchical QAM, and by reducing ISI from superimposed symbols of the LP stream. In [7] and [8], the spectrum efficiency is optimized using hierarchical modulation for DVB systems when the receivers experience high or low SNR. In [9], the authors investigated a hierarchical modulated OFDM system for local service insertion in a single frequency

* Corresponding Author

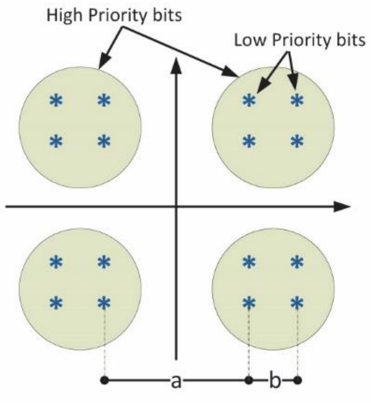

Fig 1. An instance of non-uniform hierarchical 16-QAM constellation.

network under various channel conditions. In [10], the authors investigated theoretical BER performance of HM over Rayleigh fading channels.

In order to improve the BER of the LP stream, one solution is to exploit the robustness of the HP bits as a mean to help the decoding of the LP stream, by playing the role of *soft pilot* information, as proposed in [11]. Several other reception schemes are proposed in [12] to improve the decoding of the LP stream inside an iterative receiver.

It is well known that obtaining accurate channel estimates is a main requirement to perform accurate signal detection. To perform channel estimation, one simple and widely-used method consists in sending some training (pilot) symbols from the transmitter. However, in mobile and fast fading conditions, due to channel variations, this method requires sending continuously multiple pilot sequences per frame, which however, reduces significantly the spectral efficiency.

Joint channel estimation and data detection in high mobility environments are studied in [13]. In [14], the authors propose an iterative, joint channel estimation, equalization and data detection method in the presence of high mobility for a multicarrier downlink systems operating over rapidly time-varying channels. In [15], different channel estimation methods are proposed for time-varying frequency-selective channels in multiple-input multiple-output (MIMO) relay systems. The performance of expectation-maximization (EM) algorithm for semi-blind channel estimation is investigated in [16] where the effectiveness of semi-blind channel estimation methods are compared to pilot-only based channel estimation.

As an alternative to classical pilot-based channel estimation, here, we propose a variational Bayesian EM approach (VB-EM) [17], [18] to improve the channel estimation and hence to provide a more accurate data detection. Throughout this paper, we particularly focus on reducing the necessary SNR to achieve a target BER for the LP stream.

II. Principle of Hierarchical Modulation

HM is a binary labeling technique for multiplexing and modulating data streams into one single symbol stream. Receivers with "good" channel conditions can detect both HP and LP streams, while those with poorer channel conditions may only detect the HP stream. Indeed, service providers can target two different types of receivers with two completely different services. Typically, the LP stream is of higher bit rate, but lower robustness than the HP one.

Although HM can be applied to any constellations, without loss of generality, we will limit our discussions to QPSK and 16-QAM constellations, which are widely-adopted in practical communication systems. The coded basic information bits are mapped to the QPSK constellation. The hierarchical constellation is next modified according to the coded secondary information bits, and the combined hierarchical constellation is formed as shown in Fig. 1.

This type of labeling provides an unequal protection on different contents, e.g., video, audio, text, etc. In general, HM is characterized by parameter α, where $\alpha = \frac{a}{b}$, (α is also referred to as hierarchy parameter) [19], with a being the minimum distance between two constellation points carrying different HP bits, and b is the minimum distance between any constellation point. Typically, we have $\alpha \geq 1$, where $\alpha = 1$ corresponds to the uniform and $\alpha > 1$ represents non-uniform 16-QAM.

III. System Model and Main Assumptions

The considered transmitter architecture is depicted in Fig. 2. As shown, we consider a BICM scheme with an OFDM system employing N subcarriers. As shown, the binary data sequences HP and LP are independently encoded by their respective convolutional encoder module before being randomly interleaved [20]. The two interleaved streams are multiplexed into a single stream, and then mapped on a non-uniformly spaced 16-HQAM constellation. The modulated signal is then passed through an OFDM modulator and broadcasted over the multipath wireless channel. We consider an uncorrelated i.i.d. Rayleigh fading channel model where each frame of symbols corresponds to N_c fading blocks and each block contains τ_c symbols. In our model, by considering $N_c = 1$, the channel is modeled as quasi-static while $\tau_c = 1$ returns to the fast-fading channel.

The block diagram of the receiver is depicted in Fig. 3. Besides the channel estimation part, the rest of the receiver principally consists of the combination of a soft demapper and two SISO decoders [20], one for each stream. The HP and LP SISO decoders calculate the a posteriori probability (APP) and the extrinsic probability for the uncoded HP and LP streams. We consider these extrinsic probability in the form of log-likelihood ratio (LLRs). These soft informations are then re-interleaved in order to be used as *a priori* information to improve the demapping.

The source data in the frequency domain $\boldsymbol{s}_k = [s_k^0, \ldots, s_k^{N-1}]^T$ in kth time slot is modulated onto N parallel subcarriers. At the receiver, after CP removal, the received signal vector at the kth time slot $\boldsymbol{y}_k = [y_k^0, \ldots, y_k^{N-1}]^T$ can be written as

$$\boldsymbol{y}_k = \mathbf{H}_k \boldsymbol{s}_k + \boldsymbol{z}_k \tag{1}$$

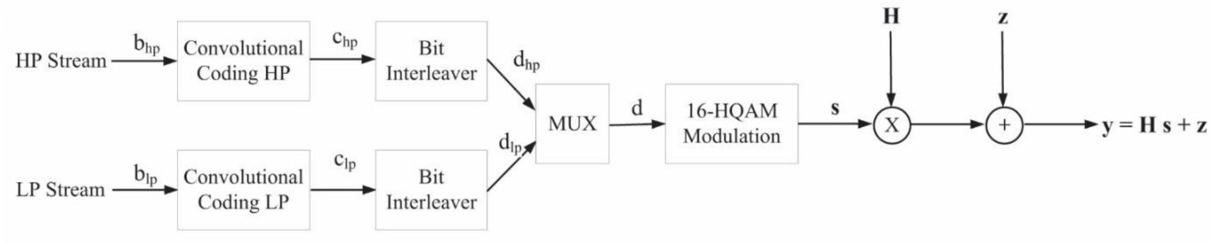

Fig 2. Block diagram of the transmission scheme using hierarchical modulation.

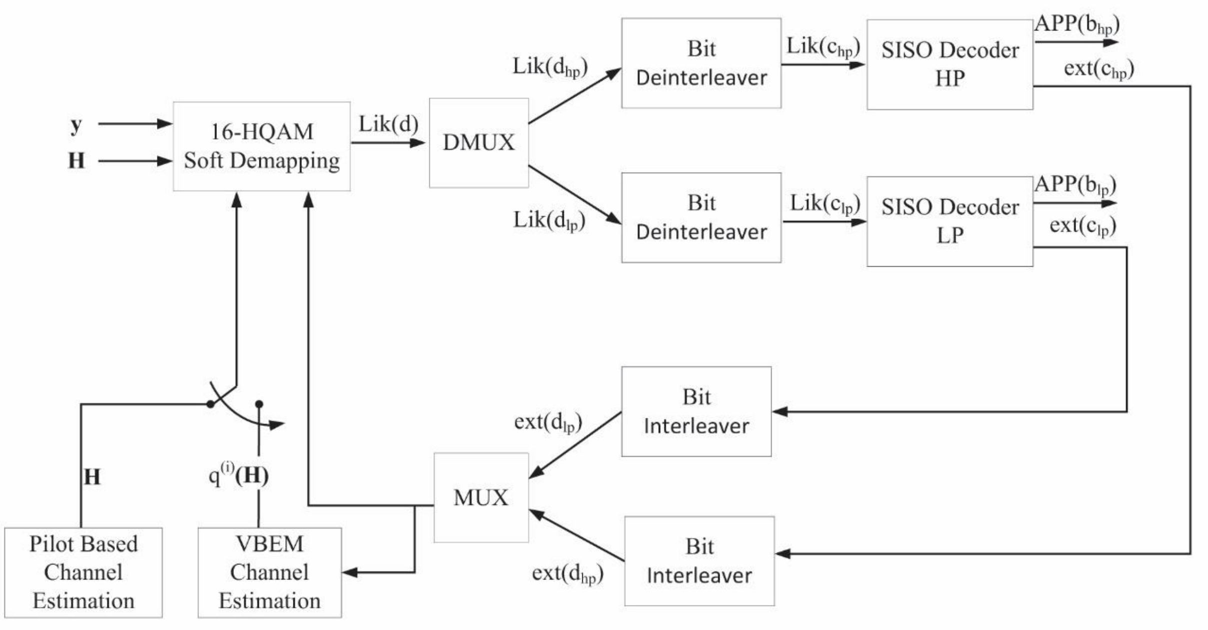

Fig 3. Block diagram of the BICM reception scheme.

where H_k is the zero mean complex Gaussian channel frequency coefficient matrix with covariance matrix $\Psi_{\mathbf{H}_k} = \mathbb{E}[\mathbf{H}_k \mathbf{H}_k^H]$. The noise vector $\boldsymbol{z}_k$ is assumed to be ZMCSCG noise with variances σ_z^2. For simplicity, hereafter, we omit the subscript k.

Let d^m be the m-th bit $m = (1, \dots, B)$, corresponding to the symbol vector $\boldsymbol{s}$. The LLR at the output of soft demapper $\gamma(d^m)$ can be obtained as [21]

$$\gamma(d^m) \propto \sum_{d^m \in \{0,1\}, n \neq m} e^{\frac{-1}{\sigma_z^2}|\mathrm{y} - \mathbf{H}\boldsymbol{s}(d^1,\dots,d^B)|^2} \prod_{n=1, n\neq m}^{B} \lambda(d^n) \tag{2}$$

where B is the size of the HM constellation and $\lambda(d^m)$ is the extrinsic probability for the bit d^m at the output of SISO decoder. Throughout receiver iterations, $\gamma(d^m)$ and $\lambda(d^m)$ are respectively *a priori* information for each other. The initial value of $\lambda(d^m)$ for the first iteration is set to $1/2$. The extrinsic probability calculated by (2) is divided into $\gamma(d_{hp})$ and $\gamma(d_{lp})$ in which d_{hp} and d_{lp} are the HP and LP bit streams, respectively. At the last iteration, the LLRs of information bits d_{hp} and d_{lp} are used for detecting the two streams.

The soft demapper requires an estimate of the channel in order to provide the probability of encoded bits [20]. Therefore, the iterations of BICM iterative decoding is naturally combined with the process of iterative semi-blind channel estimation and this leads to improved detection of the LP stream, as explained in the next Section. Notice that conventional methods based on pilot-assisted channel estimation, can not take any advantage of the *a priori* information provided by the SISO decoder.

IV. A VB-EM ALGORITHM FOR CHANNEL ESTIMATION AND DATA DETECTION

A. *Conventional Pilot-based Channel Estimation*

To estimate the transmission channel vector $\mathbf{H}$ at the receiver, which corresponds to each fading block, we send a number of pilot symbols in addition to the encoded data symbols. We devote a number of N_p channel-uses to the transmission of pilot vector $\mathbf{s}_\mathrm{p} = [\mathrm{s}_\mathrm{p}^0, \dots, \mathrm{s}_\mathrm{p}^{N-1}]^T$ for each of the fading blocks. According to the system model (1), during a given training interval with pilot vector s_p, we receive

$$\boldsymbol{y}_p = \mathbf{H}\boldsymbol{s}_p + \boldsymbol{z}_p \tag{3}$$

where $\boldsymbol{y}_p$ has the same structure as $\boldsymbol{y}$ in (1), except that it is the observation vector during pilot transmission. The definition of $\boldsymbol{z}_p$ is similar to that of $\boldsymbol{z}$ in (1). The maximum likelihood (ML) channel estimate $\widehat{\mathbf{H}}$, which is equivalent here to the least-squares solution, writes

$$\widehat{\mathbf{H}}^{\mathrm{LS}} = (\boldsymbol{s}_p^H \boldsymbol{s}_p)^{-1} \boldsymbol{s}_p^H \boldsymbol{y}_p. \tag{4}$$

It is assumed that the channel estimate $\widehat{\mathbf{H}}$ has a Gaussian distributed vector with pdf $p(\widehat{\mathbf{H}}) = \mathcal{CN}(0, \boldsymbol{\Psi}_{\widehat{\mathbf{H}}})$.

B. Variational Bayesian Iterative Channel Estimation and Data Detection

In order to improve the detection accuracy of the LP stream, in what follows, we propose to use the VB-EM approach as an alternative to estimation methods based on pilot symbols only. In fact, in addition to pilot symbols, the variational Bayesian scheme makes use of data symbols for channel estimation. In this way, a considerable performance improvement can be achieved as reported in [22]. We consider VB-EM to find an approximation of the true posterior probability $p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})$. More precisely, VB-EM looks for a parametric distribution $q(\boldsymbol{s}, \mathbf{H})$ which approximates the true posterior distribution $p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})$.

The optimal approximation is determined by minimizing the following *free energy* function

$$\mathcal{F} = \int_{H.s} q(\boldsymbol{s}, \mathbf{H}) \ln \frac{q(\boldsymbol{s}, \mathbf{H})}{p(\boldsymbol{s}, \mathbf{H}, \boldsymbol{y})} d\mathbf{H} d\boldsymbol{s} \tag{5}$$

Minimizing the free energy function is equivalent to minimizing the difference between $q(\boldsymbol{s}, \mathbf{H})$ and $p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})$. The VBA refers to an approximation of the joint distribution $q(\boldsymbol{s}, \mathbf{H})$ as a product of separable pdf, i.e., $q(\boldsymbol{s}, \mathbf{H}) = q(\boldsymbol{s})q(\mathbf{H})$, which is equivalent to assuming that s and H are independent conditioned on$y$. We have

$$\left(q^{\star}(\boldsymbol{s}), q^{\star}(\mathbf{H})\right) = \underset{q(\boldsymbol{s}), q(\mathbf{H})}{argmin} \mathcal{KL}\left(q(\boldsymbol{s})q(\mathbf{H}) || p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})\right) \tag{6}$$

subject to:

$$\int_{s} q(\boldsymbol{s}) = 1, q(\boldsymbol{s}) \geqslant 0, \forall \boldsymbol{s}, \tag{7}$$

$$\int_{H} q(\mathbf{H}) = 1, q(\mathbf{H}) \geqslant 0, \forall \mathbf{H}, \tag{8}$$

where

$$\mathcal{KL}\left(q(\boldsymbol{s}, \mathbf{H}) \parallel p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})\right) = \int_{H.s} q(\boldsymbol{s}, \mathbf{H}) \ln \frac{q(\boldsymbol{s}, \mathbf{H})}{p(\boldsymbol{s}, \mathbf{H}|\boldsymbol{y})} d\mathbf{H} d\boldsymbol{s}, \tag{9}$$

is the Kullback-Leibler divergence [23].

Classical VBA is based on an analytical solution of (6). It can easily be shown that any solution minimizing (6) follows the general form

$$\begin{aligned} q^{\star}(\boldsymbol{s}) &= K_1 \exp\{\langle \ln p(\boldsymbol{s}, \mathbf{H}, \boldsymbol{y})\rangle_{q^{\star}(\mathbf{H})}\}, \\ q^{\star}(\mathbf{H}) &= K_2 \exp\{\langle \ln p(\boldsymbol{s}, \mathbf{H}, \boldsymbol{y})\rangle_{q^{\star}(\boldsymbol{s})}\}, \end{aligned} \tag{10}$$

where K_1 and K_2 are normalization constants. From (10), we can see that the solution of (6) is not explicit unless in extremely simple cases since $q^{\star}(\boldsymbol{s})$ and $q^{\star}(\mathbf{H})$ depend on each other. Starting with an initial value, one can however iteratively compute a solution as follows

$$\begin{aligned} q^{(t)}(\boldsymbol{s}) &= K_1 \exp\left\{\langle \ln p(\boldsymbol{s}, \mathbf{H}, \boldsymbol{y})\rangle_{q^{(t-1)}(\mathbf{H})}\right\}, \\ q^{(t)}(\mathbf{H}) &= K_2 \exp\left\{\langle \ln p(\boldsymbol{s}, \mathbf{H}, \boldsymbol{y})\rangle_{q^{(t)}(\boldsymbol{s})}\right\}. \end{aligned} \tag{11}$$

According to (11), to update distributions at the t -th iteration, the VBA method disposes of distributions from previous iteration, i.e., $q^{(t-1)}(\boldsymbol{s})$ and $q^{(t-1)}(\mathbf{H})$. Considering $p(\mathbf{H}) = \mathcal{CN}(0, \boldsymbol{\Psi}_{\mathbf{H}})$, we have a prior for $\mathbf{H}$ conjugate with the likelihood $p(\boldsymbol{y}|\boldsymbol{s}, \mathbf{H})$ which is in the Gaussian distribution family. Therefore, the optimal posterior approximation $q(\mathbf{H})$ belongs also to a Gaussian family, i.e., $q(\mathbf{H}) = \mathcal{CN}(\boldsymbol{m}, \boldsymbol{R})$. Based on (1), the likelihood function $p(\boldsymbol{y}|\boldsymbol{s}, \mathbf{H})$ writes

$$p(\boldsymbol{y}|\boldsymbol{s}, \mathbf{H}) = \frac{1}{(\pi\sigma_z^2)^N} \exp\left\{-\frac{1}{\sigma_z^2} \parallel \boldsymbol{y} - \boldsymbol{s}\mathbf{H} \parallel^2\right\}. \tag{12}$$

It is found that given $q(\boldsymbol{s})$, there exist closed-form solutions for $\boldsymbol{m}$ and $\boldsymbol{R}$. Similarly, given $\boldsymbol{m}$ and $\boldsymbol{R}$, we can derive a closed-form solution for $q(\boldsymbol{s})$. Hence, $\mathcal{F}(q(\boldsymbol{s}), \boldsymbol{m}, \boldsymbol{R})$ is minimized iteratively, starting with an initial value for $q(\boldsymbol{s})$, $\boldsymbol{m}$ and $\boldsymbol{R}$. The update rules at the t -th iteration are derived as follows.

VB-E step:

$$\begin{aligned} q^{(t)}(\boldsymbol{s}) &\propto p(\boldsymbol{s}) \exp\left\{\langle \ln p(\boldsymbol{y}|\boldsymbol{s}, \mathbf{H})\rangle_{q^{(t-1)}(\mathbf{H})}\right\} \\ &\propto p(\boldsymbol{s}) \exp\left\{\langle -\frac{||\boldsymbol{y} - \boldsymbol{s}\,\mathbf{H}||^2}{\sigma_z^2}\rangle_{q^{(t-1)}(\mathbf{H})}\right\} \end{aligned} \tag{13}$$

VB-M step:

$$\widehat{\mathbf{R}}^{(t)} = \left[\boldsymbol{\Psi}_{\widehat{\mathbf{H}}}^{-1} + \frac{\langle \boldsymbol{s}^H \boldsymbol{s}\rangle_{q^{(t)}(\boldsymbol{s})}}{{\sigma_z}^2}\right]^{-1}, \tag{14}$$

$$\widehat{\boldsymbol{m}}^{(t)} = \widehat{\mathbf{R}}^{(t)} \left[\frac{\langle \boldsymbol{s}^H\rangle_{q^{(t)}(\boldsymbol{s})} \boldsymbol{y}}{\sigma_z^2}\right]. \tag{15}$$

By respectively setting the first-order derivative of the free energy (5) with respect to $\boldsymbol{R}$ and $\boldsymbol{m}$ to zero, we have the estimate of $\widehat{\boldsymbol{R}}$ and $\widehat{\boldsymbol{m}}$. Starting from (11) and after some simplifications we get

$$\begin{aligned} q^{(t)}(\mathbf{H}) &\propto p(\mathbf{H}) \exp\left\{\langle \ln p(\boldsymbol{y}|\boldsymbol{s}, \mathbf{H})\rangle_{q^{(t)}(\boldsymbol{s})}\right\} \\ &\propto p(\mathbf{H}) \exp\left\{\langle -\frac{||\boldsymbol{y} - \boldsymbol{s}\,\mathbf{H}||^2}{\sigma_z^2}\rangle_{q^{(t)}(\boldsymbol{s})}\right\} \end{aligned} \tag{16}$$

By inserting (14) and (15) in (13), we have

$$q^{(t)}(\boldsymbol{s}) \propto p(\boldsymbol{s}) \exp\left\{\frac{-\parallel \boldsymbol{y} \parallel^2 - \left(|\widehat{\boldsymbol{m}}^{(t-1)}|^2 + \widehat{\boldsymbol{R}}^{(t-1)}\right) \boldsymbol{s}^H \boldsymbol{s}}{\sigma_z^2} + \frac{2\Re\left(\boldsymbol{y}^H \widehat{\boldsymbol{m}}^{(t-1)} \boldsymbol{s}\right)}{\sigma_z^2}\right\}. \tag{17}$$

In the above formulation, $p(\boldsymbol{s}) = \prod_{i=1}^{N} p(\boldsymbol{s}_i)$ where $p(\boldsymbol{s}_i)$ is the priori probability on symbol $\boldsymbol{s}_i$ which is coming from SISO decoder. At the first iteration, the initial value of $q(\boldsymbol{s})$ is $q^{(0)}(\boldsymbol{s}) = \delta(\boldsymbol{s} - \boldsymbol{s}_p)$ and the channel is estimated by using (4) with pilot symbols.

Table I.

STEPS OF ITERATIVE DETECTION USING VB-EM ALGORITHM

Start
Consider the observation vector $\boldsymbol{y}$ in (1)
Initialization $(t=0)$**:**
Set coded bits probabilities $\gamma(d^m)$ to 0.5
Consider pilot vector $\boldsymbol{s}_p$:
Set $q^{(0)}(\boldsymbol{s})=\delta(\boldsymbol{s}-\boldsymbol{s}_p)$,
Set $\widehat{\mathbf{R}}^{(0)}=\left[\boldsymbol{\Psi}_{\widehat{\mathbf{H}}}^{-1}+\frac{\langle \boldsymbol{s}^H\boldsymbol{s}\rangle_{q^{(0)}(\boldsymbol{s})}}{\sigma_z^{2}}\right]^{-1}$
Set $\widehat{\boldsymbol{m}}^{(0)}=\widehat{\mathbf{R}}^{(0)}\left[\frac{\langle \boldsymbol{s}^H\rangle_{q^{(0)}(\boldsymbol{s})}\boldsymbol{y}}{\sigma_z^2}\right]$
For $\{\boldsymbol{t}=\mathbf{1},\dots,\boldsymbol{t_{max}}\}$
Step 1: Update $q^{(t)}(\boldsymbol{s})$ using (21) for "Improved" VB-EM
Step 2: Calculate LLRs $\gamma(d^m)$ using (2) and divide into $\gamma(d_{hp})$ and $\gamma(d_{lp})$
Step 3: Deinterleave HP and LP LLRs for SISO decoder
Step 4: Calculate extrinsic probability for the bit d^m, $\lambda(d^m)$ of HP and LP streams in SISO decoders
Step 5: Interleave the HP and LP extrinsics for using in the "Improved" VB-EM-based receiver
Step 6: Update $q^{(t)}(\mathbf{H})$:
- Update the estimate $\widehat{\mathbf{R}}^{(t)}$ by using (22)
- Update the estimate $\widehat{\boldsymbol{m}}^{(t)}$by using (23)
End
Step 7: Decode the information data by thresholding the uncoded bit probabilities

These channel estimation results are injected to the soft demapper as $\mathbf{H}^{(t-1)}\triangleq\widehat{\boldsymbol{m}}^{(t-1)}$. According to Fig. 3, the soft demapper sub-block calculates the LLRs defined in Equation (2) by inserting $\mathbf{H}=\widehat{\mathbf{H}}$. After de-multiplexing and de-interleaving, we get the likelihood of coded HP and LP bits. HP and LP SISO decoders can then respectively update the extrinsic probabilities on HP and LP streams and then import them into the VB-EM channel estimation sub-block.

C. Improved VB-EM Iterative Channel Estimation and Data Detection

In the previous section, we introduced the classical VB-EM approach for channel estimation and data detection of HP and LP streams iteratively. In this method, the inaccuracies associated to the channel estimator and SISO decoder blocks are not considered. Here, we propose an improved VB-EM approach in which the channel estimated in the previous iteration is used to update the distribution probability of symbols and channel coefficients in the next iteration. Indeed, at the t th VB-EM iteration, we assume that the distribution of $q^{(t-1)}(\mathbf{H})$ is available from the $(t-1)$ th iteration and obeys a Gaussian distribution $q^{(t-1)}(\mathbf{H})=\mathcal{CN}(\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{R}})$. To improve the detection performance of the conventional VB-EM, we propose to use the modified metric $p(\boldsymbol{s}|\boldsymbol{y},\mathbf{H}^{(t-1)})$ as an alternative to $p(\boldsymbol{s}|\boldsymbol{y})$ for data detection. This is achieved by modifying properly the formulation of the VB-EM update rules (11) as

$$q^{(t)}(\boldsymbol{s})\propto\exp\left\{\langle\ln p(\boldsymbol{s},\mathbf{H},\boldsymbol{y}|\mathbf{H}^{(t-1)})\rangle_{q^{(t-1)}(\mathbf{H})}\right\},$$
$$q^{(t)}(\mathbf{H})\propto\exp\left\{\langle\ln p(\boldsymbol{s},\mathbf{H},\boldsymbol{y}|\mathbf{H}^{(t-1)})\rangle_{q^{(t)}(\boldsymbol{s})}\right\}. \tag{18}$$

We assume the independence between $\boldsymbol{s}$, $\mathbf{H}$ and $\mathbf{H}^{(t-1)}$. After some simplification, we have

$$\begin{aligned}q^{(t)}(\boldsymbol{s})&\propto\exp\left\{\langle\ln p(\boldsymbol{s},\mathbf{H},\boldsymbol{y}|\mathbf{H}^{(t-1)})\rangle_{q^{(t-1)}(\mathbf{H})}\right\}\\&\propto\exp\left\{\langle\ln\left[p(\boldsymbol{y}|\boldsymbol{s},\mathbf{H})p(\mathbf{H}|\mathbf{H}^{(t-1)})p(\boldsymbol{s})\right]\rangle_{q^{(t-1)}(\mathbf{H})}\right\}\\&\propto p(\boldsymbol{s})\exp\left\{\langle\ln p(\boldsymbol{y}|\boldsymbol{s},\mathbf{H})\rangle_{q^{(t-1)}(\mathbf{H})}\right\}\\&\propto p(\boldsymbol{s})\exp\left\{\langle-\frac{\|\boldsymbol{y}-\mathbf{H}\boldsymbol{s}\|^2}{\sigma_z^2}\rangle_{q^{(t-1)}(\mathbf{H})}\right\}.\end{aligned} \tag{19}$$

where we have omitted all terms that do not depend on $\boldsymbol{s}$.

Theorem: For a circularly symmetric complex random vector $\boldsymbol{x}=\mathcal{CN}(\mu,\boldsymbol{\Sigma})$ with mean $\mu=E[\boldsymbol{x}]$ and covariance matrix $\boldsymbol{\Sigma}=E[\boldsymbol{x}\boldsymbol{x}^\dagger]-\mu\mu^\dagger$, and Hermitian matrix $\boldsymbol{B}$ such that $\boldsymbol{I}+\boldsymbol{\Sigma}\boldsymbol{B}\geq 0$, we have [11]

$$E_{\boldsymbol{x}}[\exp\{-\boldsymbol{x}^H\boldsymbol{B}\boldsymbol{x}\}]=\frac{\exp\{\mu^H\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{\Sigma}\boldsymbol{B})^{-1}\mu\}}{\det\{\boldsymbol{I}+\boldsymbol{\Sigma}\boldsymbol{B}\}} \tag{20}$$

Let us define $\boldsymbol{x}=\boldsymbol{y}-\mathbf{H}\boldsymbol{s}$. After some algebra, we can derive the conditional pdf of x given $\boldsymbol{s}$ and $\mathbf{H}$ as $\boldsymbol{x}|(\boldsymbol{s},\mathbf{H})\sim\mathcal{CN}(\mu,\boldsymbol{\Sigma})$, where $\mu=\boldsymbol{y}-\widetilde{\boldsymbol{m}}\boldsymbol{s}$ and $\boldsymbol{\Sigma}=\boldsymbol{s}\widetilde{\boldsymbol{R}}\boldsymbol{s}^H$. By applying Theorem (20) in (19) and by considering $q(\boldsymbol{H})$ is Gaussian, the improved VB-EM updated formulas are derived as follows.

Improved VB-E Step:

$$q^{(t)}(\boldsymbol{s})\propto\frac{p(\boldsymbol{s})}{\pi^N\left|\sigma_z^2\boldsymbol{I}_N+\boldsymbol{s}\widetilde{\boldsymbol{R}}\boldsymbol{s}^H\right|}\times$$
$$exp\left\{-(\boldsymbol{y}-\widetilde{\boldsymbol{m}}\boldsymbol{s})^H\left(\sigma_z^2\boldsymbol{I}_N+\boldsymbol{s}\widetilde{\boldsymbol{R}}\boldsymbol{s}^H\right)^{-1}(\boldsymbol{y}-\widetilde{\boldsymbol{m}}\boldsymbol{s})\right\} \tag{21}$$

Improved VB-M Step:

$$\widehat{\boldsymbol{R}}^{(t)}=\left(\frac{\langle\boldsymbol{s}^H\boldsymbol{s}\rangle_{q^{(t)}(\boldsymbol{s})}}{\sigma_z^2}+\widetilde{\boldsymbol{R}}^{-1}\right)^{-1} \tag{22}$$

$$\widehat{\boldsymbol{m}}^{(t)}=\widehat{\boldsymbol{R}}^{(t)}\left(\frac{\langle\boldsymbol{s}^H\rangle_{q^{(t)}(\boldsymbol{s})}\boldsymbol{y}}{\sigma_z^2}+\widetilde{\boldsymbol{R}}^{-1}\widetilde{\boldsymbol{m}}\right) \tag{23}$$

Starting from (18), we have

$$\begin{aligned}q^{(t)}(\mathbf{H})&\propto\exp\left\{\langle\ln p(\boldsymbol{s},\mathbf{H},\boldsymbol{y}|\mathbf{H}^{(t-1)})\rangle_{q^{(t)}(\boldsymbol{s})}\right\}\\&\propto\exp\left\{\langle\ln\left[p(\mathbf{y}|\mathbf{s},\mathbf{H})p(\mathbf{H}|\mathbf{H}^{(t-1)})p(\boldsymbol{s})\right]\rangle_{q^{(t)}(\boldsymbol{s})}\right\}\\&\propto p(\mathbf{H}|\mathbf{H}^{(t-1)})\exp\left\{\langle\ln p(\boldsymbol{y}|\boldsymbol{s},\mathbf{H})\rangle_{q^{(t)}(\boldsymbol{s})}\right\}\\&\propto p(\mathbf{H}|\mathbf{H}^{(t-1)})\exp\left\{\langle-\frac{\|\mathbf{y}-\mathbf{H}\boldsymbol{s}\|^2}{\sigma_z^2}\rangle_{q^{(t)}(\boldsymbol{s})}\right\}\\&\propto\exp(-\frac{\mathbf{H}^H\langle\boldsymbol{s}^H\boldsymbol{s}\rangle_{q^{(t)}(\boldsymbol{s})}\mathbf{H}}{\sigma_z^2}-\mathbf{H}^H\widetilde{\boldsymbol{R}}^{-1}\mathbf{H}\end{aligned}$$

$$+2\Re\left(\frac{\mathbf{H}^{H}\boldsymbol{y}\langle\boldsymbol{s}^{H}\rangle_{q^{(t)}(\boldsymbol{s})}}{\sigma_{\boldsymbol{z}}^{2}}+\mathbf{H}^{H}\widetilde{\boldsymbol{R}}^{-1}\widetilde{\boldsymbol{m}}\right)) \tag{24}$$

By considering that $q^{(t)}(\mathbf{H})$ has a Gaussian form, we have

$$q^{(t)}(\mathbf{H})=\frac{1}{\pi^{N}\,|\widehat{\boldsymbol{R}}|}\times \exp\left\{-\left(\mathbf{H}-\widehat{\boldsymbol{m}}^{(t)}\right)^{H}\left(\widehat{\boldsymbol{R}}^{(t)}\right)^{-1}\left(\mathbf{H}-\widehat{\boldsymbol{m}}^{(t)}\right)\right\} \tag{25}$$

where $\widehat{\boldsymbol{R}}^{(t)}$ and $\widehat{\boldsymbol{m}}^{(t)}$ are similar to (22) and (23), respectively.

The channel coefficients that are estimated with the improved VB-EM approach are injected to the soft demapper as $\mathbf{H}^{(t-1)} \triangleq \widetilde{\boldsymbol{m}}$ and then in this sub-block, the LLRs defined in (2) are updated by inserting $\mathbf{H}=\widehat{\mathbf{H}}$. After de-multiplexing and de-interleaving, we get the likelihood of coded HP and LP bits. HP and LP SISO decoders can then respectively update the extrinsic probabilities on HP and LP streams and then import them into the improved VB-EM channel estimation sub-block.

In the implementation of iterative receivers, there are several possible ways to jointly estimate the channel and data. More precisely, inside each VB-EM iteration t, the receiver performs several decoding iterations, keeping the channel distribution and updating $q^{(t)}(\boldsymbol{s})$. In this paper, we consider that the receiver performs only one pass through the SISO decoder inside each VB-EM iteration. The main steps of the proposed improved VB-EM-based iterative receiver are summarized in Table I.

V. Simulation Results

Here, we provide numerical results to evaluate the performance improvement provided by the proposed reception schemes in the presence of HM. For HP and LP channel encoding, we consider a non-recursive non-systematic convolutional code with rate ½ and constraint length 3, defined in octal form by $(5,7)_8$. All interleavers are pseudo-random and operate over their entire input sequence length which is composed of 40 data OFDM symbols and one pilot OFDM symbol, necessary for VB-EM initialization. Data symbols belong to the hierarchical 16-HQAM constellation. One OFDM symbol is composed of 128 complex constellation symbol. Performance evaluation is performed over the uncorrelated Rayleigh fading channel. Iterative decoding based on *maximum a posteriori* (MAP) decoding of BICM is performed at the receiver, and the BER is calculated after 3 decoding iterations.

Let us now compare the BER performance of a 16-HQAM and a non-hierarchical 16-QAM constellation. Fig. 4 shows the system behavior of HM for HP and LP streams using the transceiver presented in Figs. 2 and 3. For behavior of HM for HP and LP streams using the transceiver presented in Figs. 2 and 3. For comparison, we have also provided the BER obtained with a non-hierarchical 16-QAM modulation. We observe that HM creates two layer of error protection. Compared with non HM, the HP stream is highly

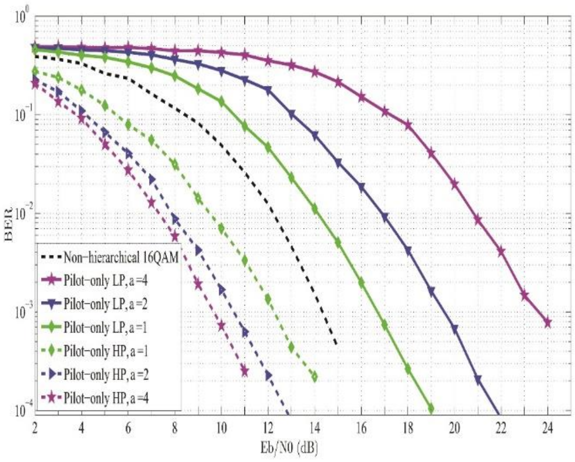

Fig 4. System behavior for HP and LP streams in the case of HM and pilot-only channel estimation for different values of parameter α.

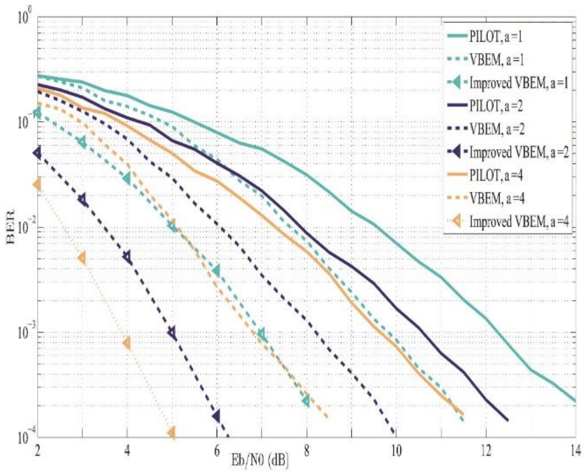

Fig 5. HP stream BER by using improved VB-EM algorithm for different values of parameter α.

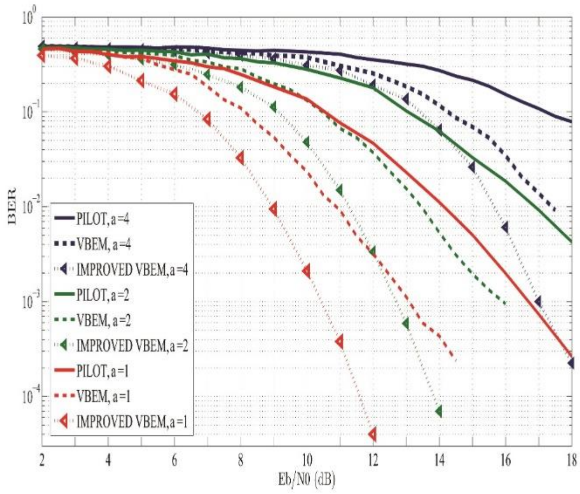

Fig 6. LP stream BER by using improved VB-EM algorithm for different values of parameter α.

protected while the LP stream BER is degraded compared to the non-hierarchical constellation. Moreover, we observe that the robustness of the HP stream is increased even further by increasing the parameter α. However, we notice that the main drawback of the receiver using pilot-based channel estimation is that when α increases, the LP stream

Table II.

BEHAVIOR COMPARISON OF 16-HQAM AND NON-HIERARCHICAL 16-QAM.

Modulation	HP Gain (dB)	LP Loss (dB)	Diff. (dB)
Classical 16-QAM	0	0	0
16-HQAM ($\alpha = 1$)	2	2.75	-0.75
16-HQAM ($\alpha = 2$)	3.5	5	-1.5
16-HQAM ($\alpha = 4$)	4.5	9.75	-5.25

Table III.

REDUCTION OF THE NECESSARY SNR TO ACHIEVE A BER OF 5×10^{-4} FOR THE HP AND LP STREAMS BY USING THE PROPOSED VB-EM CHANNEL ESTIMATION WITH 16-HQAM MODULATION.

Algorithm	Parameter	HP Gain (dB)	LP Gain (dB)
VB-EM	$\alpha = 1$	2.25	3.25
	$\alpha = 2$	2.75	4
Improved VB-EM	$\alpha = 1$	5.25	6.25
	$\alpha = 2$	6.25	7.5

requires very high SNR in order to achieve the same BER as that of the HP stream.

In Table II, we have reported the relative gain (for the HP stream) and loss (for the LP stream) in SNR (in dB) resulting from using a hierarchical 16-QAM rather than a non-hierarchical 16-QAM modulation, where the BER is equal to 5×10^{-4}. In particular, we observe that the sensitivity of the LP stream to noise is larger when α increases. However, with this type of modulation and by increasing α, the HP stream becomes more resistant, at the expense of a severe degradation of the LP stream.

Figs. 5 and 6 depict the BER versus SNR for the HP and LP streams, respectively. As shown, the VB-EM scheme proposed in this paper, leads to an SNR gain for both HP and LP streams. Also, we observe that the proposed improved VB-EM algorithm for channel estimation and data detection has improved the system performance. More precisely, the required SNR to achieve a target BER for both the LP and HP streams is effectively reduced by using VB-EM. Subsequently, with the improvement of the VB-EM algorithm and introducing an improved VB-EM scheme, less SNR is required compared to the channel estimation using pilot symbols or the semi-blind VB-EM method to achieve a target BER.

Table III shows the gain achieved by VB-EM and improved VB-EM algorithms in compared with pilot-only method where the BER is equal to 5×10^{-4}. By increasing α, further improvement of the LP streams can be achieved by using VB-EM. This improvement is very important at the bit error rate for this highly exposed stream. Because the increase in the quality of HP data receipts against the increase in LP stream noise occurs, we corrected this error rate well by designing an optimal BICM receiver along with semi-blind channel estimation algorithms.

VI. Conclusion

In this paper, we showed that when pilot-based channel estimation is adopted, the enhanced protection offered by HM for the HP stream is achieved at the price of lower noise immunity for the LP stream and hence an important performance degradation. As an alternative to pilot-only channel estimation, we proposed a *Variational Bayesian* approach and its improved algorithm for channel estimation and data detection, particularly with the aim of improving the LP detection performance. The proposed receiver led effectively to an improvement of the LP stream detection performance in terms of BER. Our numerical results confirmed the superiority of the proposed channel estimation method in terms of BER compared to pilot-only channel estimation.

References

[1] Jiang, H., Wilford, P. "*A Hierarchical Modulation for Upgrading Digital Broadcast Systems*". IEEE Trans. Broadcasting, vol. 51, no. 2, 2005, pp. 223-229.

[2] Digital Video Broadcasting (DVB), Implementation guidelines for DVB terrestrial services, Transmission aspects. ETSI Technical Report, TR 101 190, December 1997.

[3] Meric, H., Lacan, J., Arnal, F., Lesthievent, G., Boucheret, M. L. "*Improving broadcast channel rate using hierarchical modulation*". EW. 18th European Wireless Conference, 2012, pp. 1–5.

[4] Yongheng, L., Conor, H. "*Optimization of Hierarchical Modulation for Use of Scalable Media*". EURASIP Journal on Advances in Signal Processing, 2010, pp. 1-12.

[5] Xiao, B., Xiao, K., Chen, Z., Xia, B., Liu, H., Wang, L., "*Joint design for modulation and constellation labels in non-orthogonal transmission*", IEEE Wireless Communications and Networking Conference (WCNC), 2018.

[6] Zamkotsian, M., Peppas, K.P., Lazarakis, F. And Cottis, P.G., "*Layered offset hierarchical QAM modulation for intersymbol interference reduction*", IEEE Communication Letters, vol. 17, no. 11, 2013, pp. 2176-2179.

[7] Méric, H., Piquer, J.M., "*DVB-S2 spectrum efficiency improvement with hierarchical modulation*", in Communications (ICC), IEEE International Conference on, 2014, pp. 4331-4336.

[8] Méric, H., "*Optimising DVB-S2 spectral efficiency with hierarchical modulation for broadband applications*", International Journal of Satellite Communications and Networking, vol. 35, no. 2, 2015, pp. 151-161.

[9] Chandrakala, K., Jayakumari, J., "*An Efficient Hierarchical Modulation based Orthogonal Frequency Division Multiplexing Transmission Scheme for Digital Video Broadcasting*", vol. 31, no. 4, 2018, pp. 574-579.

[10] Evans, A.M., Quazi, T. and Xu, H., "*Ber performance of an mrt-hqam system over rayleigh fading channels*", IET Communications, vol. 10, no. 17, 2016, pp. 2473-2479.

[11] Sadough, S. M. S, Duhamel, P. "Improved iterative detection and achieved throughputs of OFDM systems under imperfect channel estimation". IEEE Trans. Wireless Commun. 2008.

[12] Sadough, S. M. S. "Improved Reception Schemes for Digital Video Broadcasting Based on Hierarchical Modulation". Radioengineering, vol. 20, no. 1, April 2011.

[13] Jan, Y. H., "*Iterative Joint Channel Estimation and Signal Detection for OFDM System in Double Selective Channels*". Wireless Personal Communications, vol. 99, no. 3, 2017, pp. 1279–1294.

[14] Kabaoglu, N., Erdogan, E., Aydin, E. "*Iterative Receiver with Low Complexity for Downlink Multicarrier Communications over Rapidly Time-Varying Channels*". Wireless Communications and Mobile Computing, 2018, pp. 1–11.

[15] Kaur, H., Khosla, M., Sarin, R. K. "*Channel estimation in a MIMO relay system: Challenges and approaches channel estimation in MIMO relay system: A review*". 2nd International Conference on Inventive Systems and Control (ICISC), 2018.

[16] Nayebi, E., Rao, B. D. "*Semi-blind Channel Estimation for Multiuser Massive MIMO Systems*". IEEE Transactions on Signal Processing, vol. 66, no. 2, 2018, pp. 540–553.

[17] Christensen LPB, Larsen J. "*On data and parameter estimation using the variational Bayesian EM-algorithm for block-fading frequency-selective MIMO channels*". Proceedings of the IEEE international conference on acoustics, speech and signal processing, ICASSP. 2006.

[18] Beal MJ. "*Variational algorithms for approximate Bayesian inference*". PhD dissertation. University College London, May 2003.

[19] Schertz, A., Weck, C. "*Hierarchical modulation- the transmission of two independent DVB-T multiplexes on a single frequency*". EBU Technical Review, 2003.

[20] P. Magniez, B. Muquet, P. Duhamel, V. Buzenac, and M. Decourville, "*Optimal decoding of bit-interleaved modulations: Theoretical aspects and practical algorithms*". Proc Int. Symp. on Turbo Codes and related topics, September 2000, pp. 169–172.

[21] Sadough, S. M. S., Khalighi, M. A., Duhamel, P. "*Improved iterative MIMO signal detection accounting for channel estimation errors*". IEEE Transactions on Vehicular Technology, 2009, vol. 58, no. 7, pp. 3154 - 3167.

[22] Sadough, S. M. S., Chamideh, Z. "*Efficient Variational Bayesian Method for Joint Channel Estimation and Signal Detection in OFDM-Based AF Relay Networks*". IEEE Comm. Letters, vol. 19, no. 10, 2015, pp. 1786 - 1789.

[23] Kullback, S. "*Information Theory and Statistics*". New York: Wiley, 1959.

Mahsa Azad received her B.Sc. and M.Sc. degrees in electrical engineering (telecommunication) from Shahid Beheshti University, Tehran, Iran, in 2012 and 2016, respectively. Her research interest include semi-blind channel estimation, iterative decoding and cooperative communications.

Seyed Mohammad Sajad Sadough received his B.Sc. degree in electrical engineering (electronics) from Shahid Beheshti University, Tehran, I.R. Iran in 2002 and the M.Sc. and his Ph.D. degrees in electrical engineering (telecommunication) from Paris-Sud 11 University, Orsay, France, in 2004 and 2008, respectively. From 2004 to 2007, he held a joint appointment with the National Engineering School in Advanced Techniques (ENSTA), Paris, France, and the Laboratory of Signals and Systems (LSS), at Supelec, Gif-sur-Yvette, France. He was a lecturer in the Department of Electronics and Computer Engineering (UEI), ENSTA, where his research activities were focused on improved reception schemes for ultra-wideband communication systems. From December 2007 to September 2008, he was a postdoctoral researcher with the LSS, Supelec-CNRS, where he was involved in the European research project TVMSL with Alcatel-Lucent France. Since October 2008, he has been a member of the Faculty of Electrical Engineering, Shahid Beheshti University, where he is currently an Associate Professor. Dr. Sadough's areas of research include signal processing, communication theory, digital and optical communication and cognitive radio.