

SEMI-BLIND CHANNEL ESTIMATION FOR AMPLIFY-AND-FORWARD COOPERATIVE RELAY NETWORKS

Yi Zhang¹, Mohammad Ali Khalighi², Seyed Mohammad Sajad Sadough³, Salah Bourennane²

¹ Civil Aviation Administration of China (CAAC), Chengdu, China

² École Centrale Marseille, Institut Fresnel, UMR CNRS 7249, Marseille, France

³ Faculty of Electrical and Computer Engineering, Shahid Beheshti University G. C., Tehran, Iran

Email: zy_lj2012@163.com, {Ali.Khalighi, Salah.Bourennane}@fresnel.fr, S_Sadough@sbu.ac.ir

Abstract—We consider channel estimation in amplify-and-forward wireless relay networks when coherent distributed space-time coding is performed at the relays. Employing minimum mean-square-error signal detection at the destination, we first bring clearance on the detector formulation when only partial channel state information is available. We then consider channel estimation and propose to use a semi-blind (SB) estimation method based on the expectation maximization algorithm. Implementing the SB estimator in an iterative scheme, we show that a considerable improvement can be achieved in the receiver performance after processing only few iterations. We discuss the optimality of the proposed estimation scheme and further propose modifications to it in order to obtain an unbiased estimate to improve the receiver's performance. Moreover, we show that this way, we can obtain an interesting improvement in the receiver performance for large signal constellations.

Index Terms—Wireless sensor networks, cooperative relay networks, amplify and forward, semi-blind channel estimation, expectation maximization.

I. INTRODUCTION

In wireless sensor networks (WSNs) subject to channel fading, classical diversity techniques such as the use of multiple antennas in each sensor node cannot be used due to the limitations on the nodes' size and cost issues. It is well known that cooperative diversity can provide significant performance improvement in such cases [1], [2], [3]. By cooperative (also called distributed) diversity techniques, nodes cooperate among them to exploit some amount of spatial diversity. A special case of a WSN is a wireless relay network (WRN) in which some nodes just relay the transmitted signal from a source towards a destination. We consider in this work a WRN where two or more relay nodes are inserted between a source and a destination node. Nodes are considered to be subject to a half-duplex constraint due to hardware implementation limitations, and so, dual-hop data transmission is performed. In the first hop, data is transmitted from the source to the relays. Next, in the second hop, the received signals are processed at the relays and retransmitted to the destination while the source remains idle. We consider the famous amplify-and-forward (AF) signaling mode at the relays [4], [5], [6], where the

relays just amplify the (noisy) received signal and retransmit it to the destination while having no need to channel state information (CSI). It is well known that, compared to the decode-and-forward (DF) mode [4], [7], AF relaying has the advantages of reduced transmission delay, complexity, and power consumption at the relays.

To exploit spatial diversity in WRNs, the use of distributed space time block coding (DSTBC) has extensively been studied in the literature so far (see for instance [4], [5], [8], [9], [10], [11]). Depending on whether the destination knows the CSI or not, non-coherent, differential, or coherent DSTBCs can be used. We consider in this work coherent DSTBC at the relays that allows a simpler implementation and a better system performance, and focus on the channel estimation problem. In fact, for coherent DSTBCs, in practice, we require the CSI at the destination for signal detection, which is estimated using some pilot (training) sequences [12].

Two protocols can be used for pilot transmission. In the first protocol, firstly, the source transmits pilot signals to the relays. Then, during a second phase, the relays retransmit the corresponding received signals to the destination one by one in a time-division multiple access (TDMA) mode [12]. In the second protocol, that we consider in this paper, DSTBC is performed on pilots in the same way as on data symbols in order to benefit from distributed space diversity [13]. We will later explain in Subsection V-A that when orthogonal DSTBC is performed at the relays, this second approach is equivalent to using a code-division multiple access (CDMA) protocol over synchronous relays. To obtain the same channel estimation quality, the TDMA protocol needs more channel-uses to be devoted to pilot transmission, especially for increased number of relays. Note that for both methods, as we will explain later, we do not obtain *full* CSI but a *partial* CSI estimate.

The estimated CSI is obviously imperfect. In particular, as power consumption is a critical issue in sensor networks [14], [15], we should minimize the number of pilots in order to reduce power consumption at the source and at the relays. In [16], assuming pilot-only channel estimation, we derived an improved detector for AF relaying networks. Here, we extend

our initial results provided in [16] to iterative semi-blind (SB) receivers, where we make use of the data symbols in addition to the pilots to improve the channel estimation quality. More precisely, we propose here to use SB channel estimation based on the expectation-maximization (EM) algorithm. The interest of EM is that the estimate converges to the maximum likelihood (ML) solution. We present a simple formulation of this estimator implemented in an iterative receiver, and show that with this approach, we can considerably reduce the pilot overhead while obtaining a good-quality channel estimate. Furthermore, we show that this simple formulation of EM provides a biased channel estimate. In order to remove the bias, we then propose some modifications to the classical formulation of EM to further improve the quality of the channel estimate. We discuss the appropriateness of the proposed solutions and show that they are particularly interesting for relatively large constellation sizes.

In this paper, we were inspired by a previous work on channel estimation for multiple-input multiple-output systems [17]. Compared to [17], our contributions reside in adapting the formulation to the case of AF WRNs and considerations regarding non-Gaussian noise at the destination, estimator formulation by taking DSTBC at the relays into account, and extension to QAM modulations.

SB estimation is obviously more computationally complex than the estimation based only on pilots. We suppose that, in contrary to the source and the relay nodes, we do not have stringent constraints on the energy, memory, and computational resources of the destination node.

The remainder of this paper is organized as follows. The main assumptions and definitions are presented in Section II. The formulation of signal transmission and space-time processing at the relays are specified in Section III, and signal detection at the destination is described in Section IV. Pilot-only-based (PO) channel estimation and the proposed SB approaches are presented in Section V. Some simulation results are then provided in Section VI to compare the performances of different channel estimation methods. Finally, Section VII concludes the paper.

Notations: $(\cdot)^*$ denotes complex conjugate and $(\cdot)^t$ vector transpose; $(\cdot)^\dagger$ stands for hermitian operator; bold lower- and upper-case letters are used for vectors and for matrices, respectively; $\mathbf{I}_n$ denotes the $(n \times n)$ Identity matrix, and $\mathbf{diag}(\cdot)$ a diagonal matrix with its diagonal entries given in the parentheses; a_i denotes the i th entry of the vector $\mathbf{a}$, and $\mathbf{A}(i, j)$ the (i, j) th entry of the matrix $\mathbf{A}$; $\mathbb{E}\{\cdot\}$ is the expected value, $\text{Var}\{\cdot\}$ the statistical variance, and $\|\cdot\|$ denotes the Frobenius norm; $\mathcal{CN}(\mu, \sigma^2)$ specifies a complex circular Gaussian distribution of mean μ and variance σ^2 .

II. ASSUMPTIONS AND PARAMETER DEFINITIONS

Consider a WRN as shown in Fig. 1, which is composed of one source node, R relay nodes, and one destination node. We suppose that there is no direct transmission link between the source and the destination, and assume perfect time-synchronization among all nodes. Let $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_{T_1}]^t$ be a transmitted block of symbols from the source; T_1 corresponds

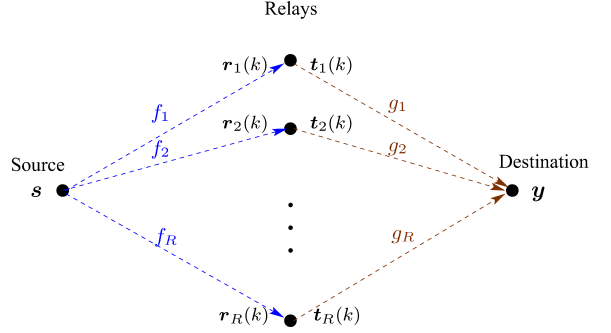


Fig. 1. General scheme of the considered relay network: f_i and g_i represent fading coefficients corresponding to the source-relays and relays-destination sub-channels, respectively.

in fact to the number of channel-uses in the first hop. We consider power-normalized transmitted symbols and impose $\mathbb{E}\{\mathbf{s}^\dagger \mathbf{s}\} = T_1$. Upon the reception of these signals, the relays amplify and retransmit them according to a DSTBC scheme, by which the received block of T_1 signals is mapped to a block of T_2 signals; with T_2 denoting the number of channel-uses during the second hop. To have the same data rate in the source-to-relays and relays-to-destination transmission steps, we choose $T_1 = T_2 = T$.

Let $\mathbf{r}_i(k)$ and $\mathbf{t}_i(k)$ represent the received and the retransmitted signals at the i -th relay and corresponding to the k -th block, respectively. The corresponding received block at the destination is denoted by $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_T]^t$. Each symbol frame corresponds to N_s blocks of T symbols each. As shown in Fig. 1, we further denote the channel fading coefficient from the source to the i -th relay by f_i , and that from the i -th relay to the destination by g_i , $i = 1, \dots, R$. Without loss of generality, we consider frequency non-selective, quasi-static, and Rayleigh independent identically-distributed (IID) fading model for f_i and g_i and model them by $\mathcal{CN}(0, \sigma_{f_i}^2)$ and $\mathcal{CN}(0, \sigma_{g_i}^2)$, respectively.

Let q be the total transmit power of the whole network during each channel use, and q_1 and q_2 the corresponding transmit powers allocated to the source and each relay, respectively.¹ Optimal power distribution among the source and the relays for minimizing the pairwise error probability is considered in [8], [9], where it is shown that for the case of $T_1 = T_2$, the optimal solution is given by $q_1 = \frac{q}{2}$ and $q_2 = \frac{q}{2R}$. This solution also maximizes the expected value of the received SNR [8]. Therefore, we weight the transmitted symbols from the source by $\sqrt{q_1}$. Also, we assume that the additive noises at the relays are Gaussian-distributed. We denote the noise at the i -th relay by n_{r_i} and its variance by $\sigma_{nr_i}^2$, i.e., $n_{r_i} \sim \mathcal{CN}(0, \sigma_{nr_i}^2)$. Given that the variance of the received signals at this relay equals $(\sigma_{g_i}^2 q_1 + \sigma_{nr_i}^2)$, the retransmitted symbols from the relay should be weighted by a scalar c_i defined as:

$$c_i = \sqrt{\frac{q_2}{\sigma_{g_i}^2 q_1 + \sigma_{nr_i}^2}}. \quad (1)$$

¹We assume here that we have equal power available at all the relay nodes. Our formulation can however be generalized to the case where different powers are available at the different relays [13].

For the sake of notation simplicity, we will assume that all the channel fading coefficients are normalized and set $\sigma_{g_i}^2$ to one. As a result, we use the same weight at all the relays as follows.²

$$c = \sqrt{\frac{q_2}{q_1 + \sigma_{nr}^2}}. \quad (2)$$

III. DATA TRANSMISSION FORMULATION

We consider bit-interleaved coded modulation (BICM) at the source. This way, the information bits are first encoded using a non-recursive non-systematic convolutional (NRNSC) code, and then pseudo-randomly interleaved before signal modulation according to quadrature amplitude modulation (QAM). Remember the dual-hop transmission protocol from Section I. Each frame of *data* symbols (i.e., excluding pilots) is divided into N_d blocks of T symbols. After the first hop, upon the transmission of a block of symbols, the received signal vector at the i -th relay, corresponding to the k -th block, is:

$$\mathbf{r}_i(k) = \sqrt{q_1} \mathbf{f}_i \mathbf{s}(k) + \mathbf{n}_{r_i}(k), \quad i = 1, 2, \dots, R, \quad k = 1, \dots, N_d, \quad (3)$$

where the vectors are of size $(T \times 1)$. For the sake of notation simplicity, hereafter we will mostly not indicate the block reference k . Next, in the second hop, at the relays, we perform DSTBC on the received signals. Let $\mathbf{C}_i$ of dimension $(T \times T)$ be the precoding matrix at the i -th relay. The retransmitted $(T \times 1)$ vector $\mathbf{t}_i$ from this relay is:

$$\mathbf{t}_i = c \mathbf{C}_i \mathbf{r}_i^{(*)}, \quad (4)$$

where $\mathbf{r}_i^{(*)} = \mathbf{r}_i^*$ if complex conjugation needs to be done on $\mathbf{r}_i$ according to the DSTBC scheme, and $\mathbf{r}_i^{(*)} = \mathbf{r}_i$ otherwise.³ Finally, the vector of received signals at the destination is:

$$\mathbf{y} = \sum_{i=1}^R g_i \mathbf{t}_i + \mathbf{n}_d, \quad (5)$$

where $\mathbf{n}_d$ denotes the vector of the received noise. We denote the variance of its entries by σ_d^2 . Using equations (3), (4), and (5), we have:

$$\mathbf{y} = \sum_{i=1}^R c \sqrt{q_1} \underbrace{\mathbf{f}_i^{(*)} g_i}_{\mathbf{h}_i} \underbrace{\mathbf{C}_i \mathbf{s}^{(*)}}_{\mathbf{p}_i} + \underbrace{\sum_{i=1}^R c \mathbf{C}_i g_i \mathbf{n}_{r_i}^{(*)}}_{\mathbf{z}} + \mathbf{n}_d. \quad (6)$$

Here, the vector $\mathbf{p}_i = \mathbf{C}_i \mathbf{s}^{(*)}$ is of size $(T \times 1)$, and we have defined the $(T \times 1)$ vector $\mathbf{z}$ as the vector of the total received noise at the destination, and also, $\mathbf{h}_i = \mathbf{f}_i^{(*)} g_i$. For simplicity, we assume that we have the same noise variance at the relays and at the destination and set $\sigma_{nr}^2 = \sigma_d^2 = \sigma^2$. To write $\mathbf{y}$ in a

vector multiplication form, we also define the $(R \times 1)$ vector $\mathbf{h}$ and the $(T \times R)$ matrix $\mathbf{P}$ as follows.

$$\mathbf{h} = [f_1^{(*)} g_1, f_2^{(*)} g_2, \dots, f_R^{(*)} g_R]^t \quad (7)$$

$$\mathbf{P} = [\mathbf{C}_1 \mathbf{s}^{(*)}, \mathbf{C}_2 \mathbf{s}^{(*)}, \dots, \mathbf{C}_R \mathbf{s}^{(*)}] \quad (8)$$

Then,

$$\mathbf{y} = c \sqrt{q_1} \mathbf{P} \mathbf{h} + \mathbf{z}. \quad (9)$$

• Case studies

Let us consider the cases of two and four relay nodes. For $R = 2$, the appropriate DSTBC at the relays is the Alamouti scheme with $T = 2$ and the precoding matrix as follows.

$$\mathbf{P} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \quad (10)$$

We have:

$$\mathbf{P} = [\mathbf{C}_1 \mathbf{s}, \mathbf{C}_2 \mathbf{s}^*],$$

with

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} f_1 g_1 \\ f_2^* g_2 \end{bmatrix}.$$

For $R = 4$, we use the orthogonal time-switched Alamouti (TSw-Al) scheme which is full-rate but does not provide full-diversity [19]. For this scheme, $T = 4$ and

$$\mathbf{P} = \sqrt{2} \begin{bmatrix} s_1 & -s_2^* & 0 & 0 \\ s_2 & s_1^* & 0 & 0 \\ 0 & 0 & s_3 & -s_4^* \\ 0 & 0 & s_4 & s_3^* \end{bmatrix}. \quad (11)$$

Here, the factor $\sqrt{2}$ ensures normalized transmit power from the relays. Also,

$$\mathbf{P} = [\mathbf{C}_1 \mathbf{s}, \mathbf{C}_2 \mathbf{s}^*, \mathbf{C}_3 \mathbf{s}, \mathbf{C}_4 \mathbf{s}^*], \quad (12)$$

$$\mathbf{s} = [s_1, s_2, s_3, s_4]^t,$$

$$\mathbf{C}_1 = \sqrt{2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{C}_2 = \sqrt{2} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{C}_3 = \sqrt{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C}_4 = \sqrt{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

and

$$\mathbf{h} = [f_1 g_1, f_2^* g_2, f_3 g_3, f_4^* g_4]^t. \quad (13)$$

Note that the precoding matrices $\mathbf{C}_i$ are orthonormal but not unitary. The interest of the TSw-Al scheme is that the complexity of the optimal detector at the receiver grows linearly with R .

²If one wishes to consider the general case of different channel gains, however, the only modification that should be made in the other formulations of this paper is to consider Eq.(1) for the weight factor at different relays.

³Note that if we do not want to use such a notation, to provide a general formulation for the received signal, we should separate the real and imaginary parts of the parameters, like in [8], [18], [19]. For the sake of notation simplicity, we do not adopt this approach. As we will see, in our study, we consider some classical DSTBCs; we will simply specify the receiver formulation for them.

A. Extension to the case of OFDM signaling

The data transmission formulation presented above considers frequency non-selective fading for f_i and g_i . When this condition does not hold, usually, OFDM signaling is performed in order to simplify the channel equalization at the receiver. We explain here how to extend the presented formulation as well as the data detection and channel estimation methods to the case of OFDM.

Consider OFDM signaling based on N_c subcarriers. At the transmitter, we perform inverse fast Fourier transform (IFFT) followed by cyclic prefix (CP) insertion. At the receiver, after CP removing, we perform FFT on the received signal. This way, a frequency selective channel is equivalently transformed to a set of N_c flat sub-channels. For simplicity, we assume that all impulse responses of the underlying channels in a WRN have length equal to L . Also, without loss of generality and for the sake of notation simplicity, we do not consider the CP and assume that all the subcarriers are modulated, i.e., there is no virtual subcarrier [20]. Consider the vector $\mathbf{h}_i$ of length L that contains the channel taps corresponding to the source-Relay i -destination. Corresponding to $\mathbf{h}_i$, we consider the diagonal matrix $\mathcal{H}_i$ so that its (k, k) th entry, which corresponds to the k th subcarrier, is given by:

$$\mathcal{H}_i^k = \mathcal{H}_i(k, k) = \sum_{\ell=0}^{L-1} h_{i,\ell} e^{-j2\pi k\ell/N_c}, \quad (14)$$

where $h_{i,\ell}$ is the ℓ th tap of $\mathbf{h}_i$ and $j = \sqrt{-1}$. We assume that the channel coefficients are constant during the transmission of a frame of OFDM symbols, and they change to new independent values from one frame to next. In other words, we assume quasi-static fading for the OFDM channel. At the relays, we perform DSTBC on the received OFDM symbols in the same way as we explained for the case of flat channel with single-carrier modulation. The formulation (9) will then apply to every subchannel. For example, for the case of $R = 2$, we perform Alamouti DSTBC on the two consecutive OFDM symbols. For instance, consider two consecutive OFDM symbols, corresponding to $[s_1^1, s_2^1, \dots, s_1^{N_c}]$ and $[s_2^1, s_2^2, \dots, s_2^{N_c}]$ at the source node. At the destination, after OFDM demodulation, corresponding to the k th subcarrier we have:

$$\begin{bmatrix} y_1^k \\ y_2^k \end{bmatrix} = c\sqrt{q_1} \begin{bmatrix} s_1^k & -s_2^{k*} \\ s_2^k & s_1^{k*} \end{bmatrix} \begin{bmatrix} \mathcal{H}_1^k \\ \mathcal{H}_2^k \end{bmatrix} + \begin{bmatrix} z_1^k \\ z_2^k \end{bmatrix}, \quad (15)$$

where, for instance, y_1^k denotes the received symbol corresponding to the k th subcarrier and the first channel use.

The signal detection and channel estimation formulation that will be presented in the following sections, can similarly be adapted to the case of OFDM signaling. Note that if we assume independent fading on the different subcarriers, the quasi-static fading model that we described for the case of OFDM signaling will be equivalent to the case of a flat block-fading channel. In effect, if we consider single-carrier modulation over a flat channel with $N_s T$ symbols and N_c fading blocks per frame, it will be equivalent to OFDM modulation over a frequency-selective channel with N_c subcarriers and $N_s T/N_c$

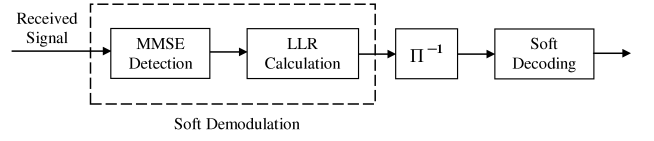


Fig. 2. Signal detection and decoding at the destination.

OFDM symbols per frame [19]. In the following sections, we come back to our assumption of a quasi-static flat channel.

IV. SIGNAL DETECTION AT THE DESTINATION

Taking into account the BICM scheme used at the source, at the destination, we perform soft signal demodulation followed by soft channel decoding, which is a rather classical approach. The block diagram of the receiver at the destination is shown in Fig. 2. For space-time decoding and signal demodulation, we first perform minimum mean-square-error (MMSE) detection and then convert the detected symbols to log-likelihood ratios (LLRs). These LLRs are next passed to the Max-Log-MAP channel decoder after being de-interleaved (the Π^{-1} block). In the following, we briefly provide details on the main blocks of the demodulator. For MMSE signal detection, we consider two cases where full or partial CSI is available at the destination.

A. MMSE detection under full CSI

Consider the case of $R = 2$ and Alamouti DSTBC. To give a simple expression for the detector, we propose to define the vectors $\mathbf{y}' = [y_1, y_2]^t$ and $\mathbf{s}' = [s_1, s_2]^t$, and the matrix $\mathbf{H}$ as follows.

$$\mathbf{H} = c\sqrt{q_1} \begin{bmatrix} h_1 & -h_2 \\ h_2^* & h_1^* \end{bmatrix} \quad (16)$$

We have, hence,

$$\mathbf{y}' = c\sqrt{q_1} \mathbf{H} \mathbf{s}' + \mathbf{z}', \quad (17)$$

where $\mathbf{z}'$ is obtained from $\mathbf{z}$ similar to $\mathbf{y}'$ and $\mathbf{s}'$. Then, MMSE detection for $\mathbf{s}'$ gives:

$$\hat{\mathbf{s}}' = \left(\mathbf{H}^\dagger \mathbf{H} + \sigma_z^2 \mathbf{I}_R \right)^{-1} \mathbf{H}^\dagger \mathbf{y}'. \quad (18)$$

Obviously, having calculated $\hat{\mathbf{s}}'$, we can obtain $\hat{\mathbf{s}}$, the estimate of $\mathbf{s}$. Similarly, for the case of $R = 4$ with TSw-AI DSTBC, we propose to define the vectors $\mathbf{y}' = [y_1, y_2^*, y_3^*, y_4]^t$ and $\mathbf{s}' = [s_1, s_2^*, s_3^*, s_4]^t$, and the matrix $\mathbf{H}$ as follows.

$$\mathbf{H} = c\sqrt{2q_1} \begin{bmatrix} h_1 & -h_2 & 0 & 0 \\ h_2^* & h_1^* & 0 & 0 \\ 0 & 0 & h_3^* & -h_4^* \\ 0 & 0 & h_4 & h_3 \end{bmatrix} \quad (19)$$

Again, the MMSE detector for $\mathbf{s}'$ is given by (18). Note that the MMSE formulation does not make any assumption on the distribution of $\mathbf{z}$ [21]. Assuming that g_i are known at the destination, which is true if we have full CSI (known f_i and g_i), from (6) we have:

$$\sigma_z^2 = T \sigma^2 \left(1 + c^2 \sum_{i=1}^R |g_i|^2 \right) \quad (20)$$

Under available full CSI, $\mathbf{z}$ in (6) is Gaussian distributed, and hence, the MMSE detector for the considered schemes is equivalent to the ML detector.

B. MMSE detection under partial CSI

In order that the destination can estimate g_i s, the relays should send some pilots independently from the source or that they should estimate f_i s and transmit them to the destination. Both solutions increase the transmission protocol complexity, energy consumption at the relays, and the overall data-rate loss. As we will explain in Subsection V-A, we can obtain an estimation of the vector $\mathbf{h}$ at the destination; what we will refer to as partial CSI. Meanwhile, we cannot obtain σ_z^2 from (20) as g_i s are unknown. Also, $\mathbf{z}$ cannot be modeled by a Gaussian random process: Assuming Rayleigh fading conditions, g_i s are Gaussian and hence in (6), we have the product of two Gaussian random variables (RVs) which is not Gaussian distributed. To calculate σ_z^2 , similar to what proposed in [22] for the case of ML detection, we approximate $\mathbf{z}_i$ with a Gaussian RV $\hat{z}_i$ verifying $E\{\hat{z}_i\} = E\{z_i\} = 0$ and $\text{Var}\{\hat{z}_i\} = \text{Var}\{z_i\}$. Using this approximation, the variance σ_z^2 is calculated as follows.⁴

$$\sigma_z^2 = T \sigma^2 \left(1 + c^2 \sum_{i=1}^R \sigma_{g_i}^2 \right) = T \sigma^2 (1 + c^2 R), \quad (21)$$

where the second equality holds because we assumed normalized channel coefficients.

C. LLR calculation

LLR calculation on the detected symbols $\hat{\mathbf{s}}$ is again done assuming Gaussian $\mathbf{z}$. Consider the example of QPSK modulation and $R = 2$ with Alamouti DSTBC for instance. Having obtained $\hat{s}_1$ by MMSE detection, the expression of LLR on the bit b_1 corresponding to the real part of s_1 is given by [19]:

$$\text{LLR}_1 = \log \frac{\exp \left(-\frac{(\hat{s}_{1\Re} - \alpha\sqrt{2}/2)^2}{\sigma_z^2} \right)}{\exp \left(-\frac{(\hat{s}_{1\Re} + \alpha\sqrt{2}/2)^2}{\sigma_z^2} \right)} = \frac{2\alpha\sqrt{2}\hat{s}_{1\Re}}{\sigma_z^2} \quad (22)$$

where $\hat{s}_{1\Re}$ denotes the real part of $\hat{s}_1$, and

$$\alpha = \frac{c^2 q_1 \|\mathbf{h}\|^2}{c^2 q_1 \|\mathbf{h}\|^2 + \sigma_z^2}, \quad \sigma_z^2 = \frac{c^2 q_1 \|\mathbf{h}\|^2}{(c^2 q_1 \|\mathbf{h}\|^2 + \sigma_z^2)^2} \quad (23)$$

After simplification, we obtain:

$$\text{LLR}_1 = 2\sqrt{2}\hat{s}_{1\Re}(c^2 q_1 \|\mathbf{h}\|^2 + \sigma_z^2), \quad (24)$$

and σ_z^2 is replaced by σ_z^2 , which is calculated from (21).

V. CHANNEL ESTIMATION

In this section, we first consider pilot-only-based channel estimation and then present the SB estimation approach that we propose.

⁴An *enhanced* Gaussian approximation has also been proposed in [23], which takes the partial channel knowledge into account. We do not use this approach here because, although more accurate, it is more computationally complex and also it does not practically result in a better performance for the specific case studies that we consider here.

A. Pilot-only-based channel estimation

As explained in the previous section, signal detection is done based on partial channel knowledge, that is, an estimate of $\mathbf{h}$. Following our discussion in Section I, we do not use TDMA signaling on the pilot symbols, as considered in [12], but rather we perform DSTBC on the pilots in the same way as on the data symbols [13]. In this way, we obtain a better channel estimate or, equivalently, we suffer from less data-rate loss for the same quality of channel estimate. We consider equal power allocation among pilot and data symbols in the two transmission hops. Our study can be generalized to the case of optimal power allocation, however [24]. Consider the $(T \times 1)$ vector $\mathbf{s}_0$ of BPSK-modulated pilot symbols transmitted from the source. Denoting the corresponding received vector at the destination by $\mathbf{y}_0$, from (9) we have:

$$\mathbf{y}_0 = c\sqrt{q_1}\mathbf{P}_0\mathbf{h} + \mathbf{z}_0, \quad (25)$$

where the $(T \times R)$ matrix $\mathbf{P}_0$ corresponds to distributed space-time processing on $\mathbf{s}_0$ at the relays:

$$\mathbf{P}_0 = [\mathbf{C}_1 \mathbf{s}_0^{(*)}, \mathbf{C}_2 \mathbf{s}_0^{(*)}, \dots, \mathbf{C}_R \mathbf{s}_0^{(*)}]. \quad (26)$$

If we consider N_p (identical) blocks of T pilot symbols, the least squares (LS) channel estimate solution gives:

$$\hat{\mathbf{h}}_{\text{LS}}^p = \frac{1}{c\sqrt{q_1}N_p} (\mathbf{P}_0^\dagger \mathbf{P}_0)^{-1} \sum_{k=1}^{N_p} \mathbf{P}_0^\dagger \mathbf{y}_0(k). \quad (27)$$

The optimal solution (in the sense of minimizing the variance of estimation errors) is obtained for diagonal $\mathbf{P}_0^\dagger \mathbf{P}_0$. This is the case when orthogonal DSTBC is done at the relays, whatever $\mathbf{s}_0$. For $R = 2$ with Alamouti and $R = 4$ with TSw-Al scheme, we just set $\mathbf{s}_0 = [1 \ 1]^t$ and $\mathbf{s}_0 = [1 \ 1 \ 1 \ 1]^t$, respectively. It can be easily verified that for $R = 2$, using the Alamouti DSTBC on $\mathbf{s}_0$ is equivalent to performing CDMA signaling on the pilots. In fact, we could alternatively consider this pilot signaling scheme as based on CDMA signaling with the Walsh-Hadamard matrix $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. In general, for $R = 2^n$, with n an integer, using Walsh-Hadamard codes is similar to using an equivalent orthogonal DSTBC.⁵

B. Semi-blind channel estimation

In order to obtain a better channel estimate without suffering from a considerable loss in the transmission data rate by increasing the pilot overhead, we can perform SB (also called data-aided) channel estimation. This way, we make use of the data symbols in addition to the pilots to improve the channel estimation quality. The scheme we propose here is based on the expectation maximization (EM) algorithm [26]

⁵For instance, for $R = 4$, we can use the Jafarkhani scheme [25] with $\mathbf{s}_0 = [1 \ 1 \ 1 \ 1]^t$ that becomes equivalent to using CDMA signaling on the pilots according to the following Walsh-Hadamard matrix:

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Here, for the sake of homogeneity, we use the TSw-Al scheme for both pilots and signals, however.

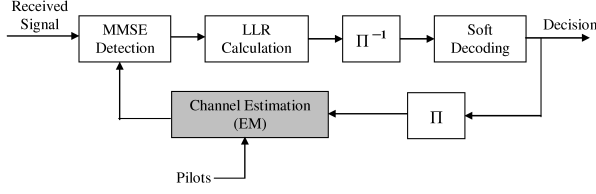


Fig. 3. Block diagram of the iterative channel estimation receiver.

and is implemented in an iterative receiver (at the destination). The block diagram of the receiver is shown in Fig. 3. At the first iteration, for signal detection, we use a primary channel estimate based on pilots only, e.g. $\hat{\mathbf{h}}_{LS}^p$ from (27). After signal demodulation and channel decoding, we use the LLRs at the output of the channel decoder to modify the channel estimate based on EM. We should take into account the DSTBC processing at the relays in the estimator formulation. Let us consider the ML estimate of $\mathbf{h}$, given the vector of observations $\mathbf{y}$:

$$\hat{\mathbf{h}}^{\text{ML}} = \arg \max_{\mathbf{h}} \log p(\mathbf{y}|\mathbf{h}). \quad (28)$$

The EM approach consists in estimating $\mathbf{h}$ from (28), given the *incomplete* data $\mathbf{y}$ and the *missing* data $\mathbf{P}$. It is implemented in two update steps of expectation and maximization that are described in the following for Iteration i [27]:

- Expectation: calculating the expected likelihood function of $(\mathbf{y}, \mathbf{P})$ conditioned on the observation $\mathbf{y}$ and the previous estimate of $\mathbf{h}$ (i.e., at Iteration $(i-1)$), denoted by $\hat{\mathbf{h}}^{(i-1)}$, that we denote by the function Q :

$$Q = \mathbb{E} \left\{ \log p(\mathbf{y}, \mathbf{P}|\mathbf{h}) \mid \mathbf{y}, \hat{\mathbf{h}}^{(i-1)} \right\}; \quad (29)$$

- Maximization: maximizing Q with respect to $\mathbf{h}$:

$$\hat{\mathbf{h}}^i = \arg \max_{\mathbf{h}} Q. \quad (30)$$

At the first iteration, PO estimation is used to initialize EM. After the i -th iteration, we have:

$$Q = \sum_{k=1}^{N_s} \sum_{u=1}^{|\chi|} \log \left[p(\mathbf{y}(k) | \mathbf{h}, \mathbf{P}(k) = \mathbf{P}_u) \right] \times p(\mathbf{P}(k) = \mathbf{P}_u | \mathbf{y}(k), \hat{\mathbf{h}}^i), \quad (31)$$

where N_s is the total number of blocks (including data and pilots), k is the block time-reference, $\mathbf{P}_u$ and $\mathbf{P}(k)$ are respectively the u -th possible and the actual space-time precoded matrices from the relays corresponding to the block k , $\hat{\mathbf{h}}^i$ is the estimated channel at the i -th iteration, and $p(\mathbf{P}(k) = \mathbf{P}_u | \mathbf{y}(k), \hat{\mathbf{h}}^i)$ is calculated using the *a posteriori* probabilities (APP) at the decoder output. Also, χ is the cardinality of $\mathbf{P}$ of size $|\chi| = 2^{BT}$, that is, the set of all possible space-time precoded matrices from the relays, with B being the number of bits per symbol. Since there is a one-to-one correspondence between a vector of symbols s_u and its corresponding DSTBC-precoded matrix $\mathbf{P}_u$, we can denote this probability by $p(s(k) = s_u | \mathbf{y}(k), \hat{\mathbf{h}}^i)$. For notation simplicity, we further denote it by $\text{APP}_k(s_u | \hat{\mathbf{h}}^i)$.

To develop (31), we should take into account the distribution of $\mathbf{y}(k)$ given $\mathbf{P}_u$ (or s_u) and $\mathbf{h}$, which reduces to the distribution of $\mathbf{z}$ from (9). However, as we consider only partial CSI at the destination, $\mathbf{z}$ is not Gaussian, and as a result, obtaining a closed form expression for the estimator is quite difficult. We propose to use again the simple Gaussian approximation for $\mathbf{z}$, which results in:

$$Q = - \sum_{k=1}^{N_s} \sum_{u=1}^{|\chi|} \log \left[2 \log(\sigma_{\hat{\mathbf{z}}}^2) + \frac{\|\mathbf{y}(k) - c\sqrt{q_1} \mathbf{P}_u \mathbf{h}\|^2}{2 \sigma_{\hat{\mathbf{z}}}^2} + \kappa \right] \times \text{APP}_k(s_u | \hat{\mathbf{h}}^i), \quad (32)$$

where κ is a constant and $\sigma_{\hat{\mathbf{z}}}^2$ is calculated using (21). Now, to find the function that maximizes Q , we should differentiate it with respect to $\mathbf{h}$. We hence obtain the estimation update for the iteration $(i+1)$:

$$\hat{\mathbf{h}}^{i+1} = \mathbf{R}_{Py} \mathbf{R}_P^{-1}, \quad (33)$$

where

$$\mathbf{R}_{Py} = c\sqrt{q_1} \sum_{k=1}^{N_s} \sum_{u=1}^{|\chi|} \mathbf{P}_u^\dagger \mathbf{y}(k) \text{APP}_k(s_u | \hat{\mathbf{h}}^i), \quad (34)$$

$$\mathbf{R}_P = c^2 q_1 \sum_{k=1}^{N_s} \sum_{u=1}^{|\chi|} \mathbf{P}_u^\dagger \mathbf{P}_u \text{APP}_k(s_u | \hat{\mathbf{h}}^i). \quad (35)$$

Note that for notation simplicity, we indicate the iteration number only for $\hat{\mathbf{h}}$. Similar to what is proposed in [17], we simplify these equations whose computational complexity grows exponentially with B and T . By this approach, using the decoder soft-outputs, we first calculate soft-estimates of the transmitted symbols, $\tilde{s}(k)$:

$$\tilde{s}(k) = \sum_{u=1}^{|\chi|} s_u \text{APP}_k(s_u | \hat{\mathbf{h}}^i). \quad (36)$$

Then, we construct the estimated space-time-coded matrices $\tilde{\mathbf{P}}(k)$, and rewrite (34):

$$\mathbf{R}_{Py} = c\sqrt{q_1} \sum_{k=1}^{N_s} \tilde{\mathbf{P}}^\dagger(k) \mathbf{y}(k). \quad (37)$$

Separating pilots and data symbols, we can also write $\mathbf{R}_{Py}$ in the following form:

$$\mathbf{R}_{Py} = c\sqrt{q_1} \left[\sum_{k=1}^{N_p} \mathbf{P}_0^\dagger \mathbf{y}_0(k) + \sum_{k=N_p+1}^{N_s} \tilde{\mathbf{P}}^\dagger(k) \mathbf{y}(k) \right], \quad (38)$$

where we have used $\mathbf{P}_0(k) = \mathbf{P}_0$ since we use identical pilot blocks. Similarly, $\mathbf{R}_P$ from (35) can be written as follows:

$$\mathbf{R}_P = c^2 q_1 \left[\sum_{k=1}^{N_p} \mathbf{P}_0^\dagger \mathbf{P}_0 + \sum_{k=N_p+1}^{N_s} \sum_{u=1}^{|\chi|} \mathbf{P}_u^\dagger \mathbf{P}_u \text{APP}_k(s_u | \hat{\mathbf{h}}^i) \right]. \quad (39)$$

The two terms in the brackets in (39) are diagonal matrices as we perform orthogonal DSTBC at the relays. Then, $\mathbf{R}_P$ can be simplified as follows:

$$\mathbf{R}_P = c^2 q_1 N_p T \mathbf{I}_R + \mathbf{\Xi}, \quad (40)$$

where we have defined the diagonal matrix Ξ as follows:

$$\Xi = \text{diag}(\xi_1, \dots, \xi_R) = c^2 q_1 \sum_{k=N_p+1}^{N_s} \tilde{\mathbf{P}}^2(k), \quad (41)$$

where $\tilde{\mathbf{P}}^2(k)$ is a diagonal matrix which is defined in (42) and (43) for the cases of $R = 2$ with Alamouti and $R = 4$ with TSw-Al DSTBC, respectively:

$$\tilde{\mathbf{P}}^2(k) = \left(\sum_{i=1}^R \tilde{s}_i^2(k) \right) \mathbf{I}_R, \quad (42)$$

$$\tilde{\mathbf{P}}^2(k) = 2 \text{diag}(\tilde{s}_1^2(k) + \tilde{s}_2^2(k), \tilde{s}_3^2(k) + \tilde{s}_4^2(k)) \otimes \mathbf{I}_2. \quad (43)$$

Here, $\tilde{s}_i^2(k)$ denotes the soft-estimate of $|s_i(k)|^2$, which is calculated similar to (36), and $\otimes$ is the Kronecker product. For the case of PSK modulations, we have $\tilde{s}_i^2(k) = 1$, $\xi_i = c^2 q_1 N_d T$, and (40) simplifies to the following:

$$\mathbf{R}_P = c^2 q_1 N_s T \mathbf{I}_R. \quad (44)$$

So, overall, we calculate in each iteration $\mathbf{R}_P$ from (44) and $\mathbf{R}_{Py}$ from (38), and update the channel estimate from (33). We will call this estimator CB-EM (CB standing for classical biased) because, as we explain in the following subsection, this classical EM formulation provides a biased estimate of the channel.

C. Improving the EM-based estimate

Let us focus on $\mathbf{R}_{Py}$ in (38) by replacing $\mathbf{y}$ and $\mathbf{y}_0$ from (9) and (25), respectively. We obtain:

$$\begin{aligned} \mathbf{R}_{Py} &= \underbrace{c^2 q_1 N_p T \mathbf{I}_R \mathbf{h}}_{\mathbf{R}_{P_0}} + \underbrace{c^2 q_1 \left(\sum_{k=N_p+1}^{N_s} \tilde{\mathbf{P}}^\dagger(k) \mathbf{P}(k) \right) \mathbf{h}}_{\mathbf{R}'_P} + \\ &\quad \underbrace{c\sqrt{q_1} \sum_{k=1}^{N_p} \mathbf{P}_0^\dagger z_0(k)}_{\boldsymbol{\eta}_0} + \underbrace{c\sqrt{q_1} \sum_{k=N_p+1}^{N_s} \tilde{\mathbf{P}}^\dagger(k) \mathbf{z}(k)}_{\boldsymbol{\eta}} \\ &= (\mathbf{R}_{P_0} + \mathbf{R}'_P) \mathbf{h} + \boldsymbol{\eta}_0 + \boldsymbol{\eta}, \end{aligned} \quad (45)$$

where we have defined the matrices $\mathbf{R}_{P_0}$ and $\mathbf{R}'_P$ and the vectors $\boldsymbol{\eta}_0$ and $\boldsymbol{\eta}$. Comparing (45) with (40), we notice that $(\mathbf{R}_{P_0} + \mathbf{R}'_P) \neq \mathbf{R}_P$, and hence, from (33), the obtained channel estimate is biased. This bias that affects the data detection part, is due to the difference between $\mathbf{P}^2$ and $\tilde{\mathbf{P}}^\dagger \mathbf{P}$ (corresponding to data symbols), or in other words, due to the difference between $\tilde{s}_i^2$ and $\tilde{s}_i^* s_i$, $i = 1, \dots, R$. At very high SNR, however, we have $\tilde{s}_i \approx s_i$ and $\tilde{s}_i^2 \approx \tilde{s}_i^* s_i$, and the bias becomes negligible.

Our aim here is to propose a modified estimator in order to eliminate the bias through the receiver iterations and to improve the channel estimation quality. Note that our approach is similar to that proposed in other contexts such as channel equalization for OFDM-based transmissions subject to multipath propagation with Doppler shift in [28], where it is proposed to remove the estimation bias through iterative

processing. Concerning our biased estimation problem, we consider here three solutions to remove the bias of the channel estimate.

1) *Direct bias removing method*: One simple solution is to remove the bias by multiplying $\hat{\mathbf{h}}$ by the inverse of the matrix $[(\mathbf{R}_{P_0} + \mathbf{R}'_P) \mathbf{R}_P^{-1}]$. We call this method Direct Bias removing and denote it by DB-EM. Let us denote by $\hat{\mathbf{h}}_{\text{EM}}$ the biased estimate obtained by the classical EM formulation. We have then,

$$\hat{\mathbf{h}}_{\text{DB}} = [(\mathbf{R}_{P_0} + \mathbf{R}'_P) \mathbf{R}_P^{-1}]^{-1} \hat{\mathbf{h}}_{\text{EM}} = \mathbf{M}^{-1} \hat{\mathbf{h}}_{\text{EM}}, \quad (46)$$

where we have defined the matrix $\mathbf{M}$ for notation simplicity. The drawback of this method could be the possible amplification of the estimation noise when multiplied by $\mathbf{M}^{-1}$.

2) *Simple quasi-unbiased method*: A simplified solution is to set the off-diagonal entries of $\mathbf{M}$ to zero and use the resulting matrix that we denote by $\mathbf{M}_{\text{diag}}$ in (46). We call this solution Simple Quasi-unbiased and denote it by SQ-EM. We have:

$$\hat{\mathbf{h}}_{\text{SQ}} = \mathbf{M}_{\text{diag}}^{-1} \hat{\mathbf{h}}_{\text{EM}}. \quad (47)$$

In this way, the calculation of $\hat{\mathbf{h}}_{\text{SQ}}$ will be less computationally complex than $\hat{\mathbf{h}}_{\text{DB}}$. As a matter of fact, if we perfectly know the transmitted symbols, $\mathbf{M}$ will be diagonal and also $\hat{\mathbf{h}}_{\text{EM}}$ becomes unbiased. So, in addition of reducing the estimator's computational complexity, this simplified solution may approach us more to the unbiased estimation.

3) *Optimally combined pilot- and data-based estimates*: We propose here another solution that can result in a better receiver performance. Let us consider separately the channel estimates based on pilots and data symbols that we denote by $\hat{\mathbf{h}}^p$ and $\hat{\mathbf{h}}^d$, respectively. Note that $\hat{\mathbf{h}}^p$ is unbiased, in contrary to $\hat{\mathbf{h}}^d$. Then, we can write:

$$\hat{\mathbf{h}}^p = \mathbf{h} + \boldsymbol{\eta}'_0, \quad \hat{\mathbf{h}}^d = \boldsymbol{\Lambda} \mathbf{h} + \boldsymbol{\eta}', \quad (48)$$

where the entries of the matrix $\boldsymbol{\Lambda}$ of dimension $(R \times R)$ are calculated as follows:

$$\lambda_{ij} = \boldsymbol{\Lambda}(i, j) = \frac{1}{\xi_i} \mathbf{R}'_P(i, j), \quad i, j = 1, \dots, R, \quad (49)$$

and ξ_i are defined in (41). Also, we have defined the estimation noise vectors $\boldsymbol{\eta}'_0$ and $\boldsymbol{\eta}'$. Using the definitions of $\boldsymbol{\eta}_0$ and $\boldsymbol{\eta}$ in (45), we have:

$$\boldsymbol{\eta}'_0 = \frac{1}{c^2 q_1 N_p T} \boldsymbol{\eta}_0, \quad \boldsymbol{\eta}'_i = \frac{1}{\xi_i} \boldsymbol{\eta}, \quad i = 1, \dots, R. \quad (50)$$

Denoting the variance of the entries of $\boldsymbol{\eta}'_0$ by $\sigma_{\boldsymbol{\eta}'_0}^2$, and that of $\boldsymbol{\eta}'_i$ by $\sigma_{\boldsymbol{\eta}'_i}^2$, it can be shown that:

$$\begin{cases} \sigma_{\boldsymbol{\eta}'_0}^2 = \frac{\sigma_{\mathbf{z}}^2}{c^2 q_1 N_p T} = \beta_0 \sigma_{\mathbf{z}}^2 \\ \sigma_{\boldsymbol{\eta}'_i}^2 = \frac{\mathbf{R}''_P(i, i)}{\xi_i^2} \sigma_{\mathbf{z}}^2 = \beta_i \sigma_{\mathbf{z}}^2, \end{cases} \quad (51)$$

where we have defined the coefficients $\beta_0 = 1/(c^2 q_1 N_p T)$ and $\beta_i = \mathbf{R}''_P(i, i)/\xi_i^2$, and also defined the matrix $\mathbf{R}''_P$ as follows:

$$\mathbf{R}''_P = c^2 q_1 \sum_{k=N_p+1}^{N_s} \tilde{\mathbf{P}}^\dagger(k) \tilde{\mathbf{P}}(k) \quad (52)$$

On the other hand, given (49) and the definition of $\mathbf{R}'_P$ in (45), we notice that for the orthogonal DSTBC schemes, matrix $\mathbf{\Lambda}$ in (48) has the following property:

$$\Psi = \mathbf{\Lambda}^\dagger \mathbf{\Lambda} = \text{diag}(\psi_1, \psi_2, \dots, \psi_R). \quad (53)$$

For instance, for $R = 2$ with the Alamouti scheme, from (10), we have $\psi_1 = \psi_2 = |\lambda_{11}|^2 + |\lambda_{12}|^2$. Also, for $R = 4$ with TSw-Al, from (11), we have $\psi_1 = \psi_2 = |\lambda_{11}|^2 + |\lambda_{12}|^2$ and $\psi_3 = \psi_4 = |\lambda_{33}|^2 + |\lambda_{34}|^2$.

Now, to obtain an unbiased estimator, we propose to look for an appropriate combination of $\hat{\mathbf{h}}^p$ and $\hat{\mathbf{h}}^d$. As a simple and low complexity solution, we consider each entry of $\hat{\mathbf{h}}$ as a linear combination of the corresponding entries of $\hat{\mathbf{h}}^p$ and $\hat{\mathbf{h}}^d$. To write this in a vector form, we define two weight matrices $\mathbf{A}$ and $\mathbf{B}$. Then, the unbiased channel estimate is given by:

$$\hat{\mathbf{h}} = \mathbf{A} \hat{\mathbf{h}}^p + \mathbf{B} \hat{\mathbf{h}}^d. \quad (54)$$

We determine the matrixes $\mathbf{A}$ and $\mathbf{B}$ so as to minimize the estimation error variance while satisfying the condition of unbiased estimation, that is,

$$\min \mathbb{E} \left\{ \|\mathbf{A} \boldsymbol{\eta}'_0 + \mathbf{B} \boldsymbol{\eta}'\|^2 \right\}, \text{ satisfying } \mathbf{A} + \mathbf{B} \mathbf{\Lambda} = \mathbf{I}_R. \quad (55)$$

Notice that assuming large enough N_s , we could approximate $\mathbf{R}'_P$ with a diagonal matrix and consequently consider diagonal matrixes for $\mathbf{A}$ and $\mathbf{B}$. However, this allows only a modest reduction of the computational complexity. Here, we make no special assumption on $\mathbf{R}'_P$ nor on $\mathbf{A}$ and $\mathbf{B}$. After some manipulations, we obtain the following optimized weight matrixes for the proposed linear combination.

$$\mathbf{A} = \mathbf{I}_R - \mathbf{\Gamma} \mathbf{\Psi}, \quad \mathbf{B} = \mathbf{\Gamma} \mathbf{\Lambda}^\dagger, \quad (56)$$

where $\mathbf{\Gamma}$ is an $(R \times R)$ diagonal matrix with the diagonal entries given below.

$$\gamma_{ii} = \mathbf{\Gamma}(i, i) = \frac{\beta_0}{\beta_i + \beta_0 \psi_i}, \quad i = 1, \dots, R. \quad (57)$$

We call this solution unbiased linearly-combined EM and will denote it by UL-EM.

4) Implementation of the improved estimation schemes:

The only remaining problem to obtain the unbiased channel estimate by the three proposed modified EM formulations is their dependence on $\mathbf{R}'_P$. Note that, concerning UL-EM, $\mathbf{R}'_P$ is required for calculating λ_{ij} in (49). As obviously, we do not know $\mathbf{P}$ at the destination, we cannot calculate $\mathbf{R}'_P$ from (45). To calculate $\mathbf{R}'_P$, we propose to simply construct $\mathbf{P}$ from the hard estimates of the transmitted symbols using the decoder soft outputs. Obviously, these hard estimates do not necessarily coincide with the true transmitted symbols. As we will see in the next section, this particularly limits the performance of the UL-EM method, especially for small constellation sizes.

VI. SIMULATION RESULTS

We present in this section some numerical results to study the performance of the different channel estimation methods. Concerning the transmitter, we consider pseudo-random interleaving and the rate 1/2 NRNSC code (5, 7)₈ in the BICM

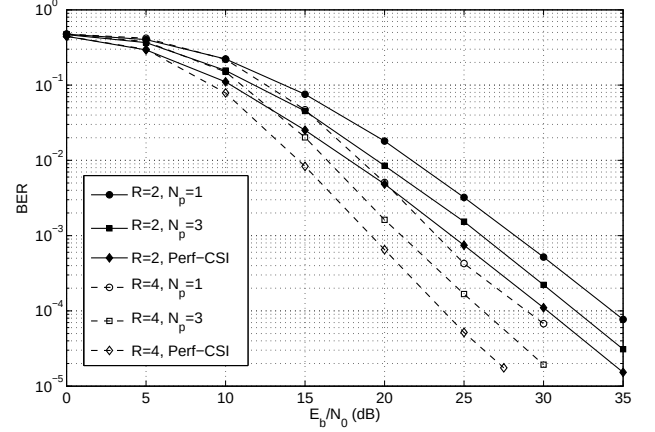


Fig. 4. LS pilot-only-based channel estimation with MMSE signal detection. N_p is the number of pilot blocks per frame, $N_s T = 64$, QPSK modulation, (5, 7)₈ channel code.

scheme. Gray-mapped QPSK and 16-QAM modulations are considered. In the quasi-static fading model that we consider, each frame of symbols corresponds to $N_s T = 64$ channel-uses. Performance evaluation is done in terms of the average bit error rate (BER) as a function of SNR which is considered in the form of E_b/N_0 ; E_b being the average received energy per information bit and N_0 the unilateral noise power spectral density. In the results we present, we will indicate the number of iterations processed at the receiver by IT. When we do not indicate it on the figures, that means that only one iteration is processed.

A. Pilot-only-based channel estimation

Let us first consider the case of PO channel estimation. Figure 4 shows average BER curves versus E_b/N_0 for QPSK modulation, $N_p = 1, 3$ and the two case studies we considered, i.e., $R = 2$ with Alamouti coding (denoted by Al), and $R = 4$ with TSw-Al coding at the relays. As reference, we have also presented the performance for the case of perfect-CSI at the receiver. Note that the number of channel-uses devoted to pilot transmission equals $N_p T$, where $T = R$. The percentage of the transmission rate devoted to pilots, i.e., N_p/N_s , equals 3.125% and 9.375% for $N_p = 1$ and 3, respectively. Assuming a target BER of 10^{-4} , we notice that, the improvement in E_b/N_0 by increasing N_p from 1 to 3, is about 2.3 dB and 2.7 dB for $R = 2$ and $R = 4$, respectively. However, this improvement is obtained at the price of a relatively high pilot overhead, as specified above. Nevertheless, even with $N_p = 3$, we are still relatively far from the perfect-CSI performance.⁶

B. EM-based channel estimation

Now consider SB estimation. We have compared the performances of different estimation methods in Fig. 5 for the case of $R = 2$, QPSK modulation, and the minimum number of pilot blocks, i.e., $N_p = 1$. For SB estimation methods,

⁶Similar results are obtained for the case of 16-QAM modulation, where the pilot overhead for $N_p = 1$ and 3 equals 6.25% and 18.75%, respectively.

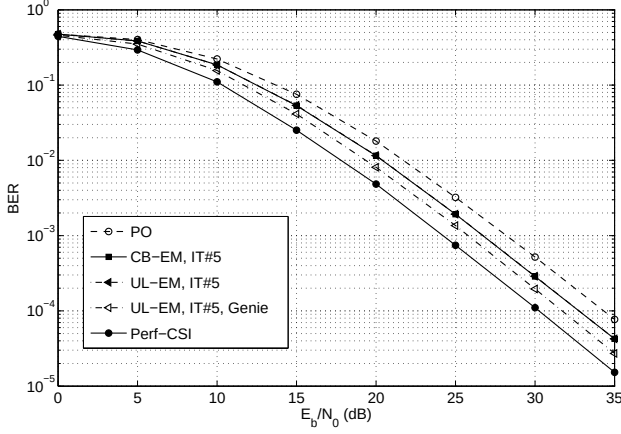


Fig. 5. Comparison of SB and PO channel estimation. IT denotes the iteration number. $R = 2$, $N_s T = 64$, $N_p = 1$, QPSK modulation, $(5, 7)_8$ channel code.

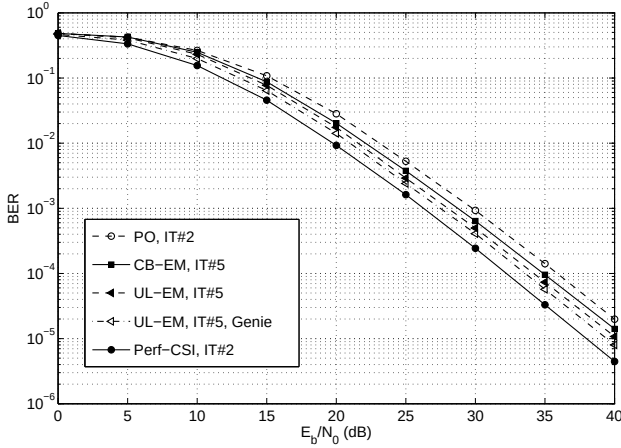


Fig. 6. Comparison of SB and PO channel estimation. IT denotes the iteration number. $R = 2$, $N_s T = 64$, $N_p = 1$, 16-QAM modulation, $(5, 7)_8$ channel code.

five receiver iterations are necessary to attain almost full receiver convergence. The least-squares PO estimate is used at the first iteration to initialize the SB algorithms. We notice an important performance improvement by SB estimation: at $\text{BER} = 10^{-4}$, we have an improvement of 1.5 dB in E_b/N_0 . The unbiased estimation methods have a very close performance and we have only shown the corresponding BER curve for UL-EM. They do not offer any advantage over CB-EM in this case, however. We have also shown in Fig. 6 the results for the case of 16-QAM modulation. We notice that CB-EM provides an improvement of about 1 dB in SNR. Again, the performance of DB-EM and QS-EM methods are not shown for the sake of clarity. Their performances are between those of CB-EM and UL-EM, with a slight advantage for QS-EM. The UL-EM approach allows 0.7 dB further improvement compared to CB-EM. Note that for increased frame size, the performance improvement by SB estimation methods becomes more significant. Also, the UL-EM takes its advantage over CB-EM for increased N_s (results are not shown for the same

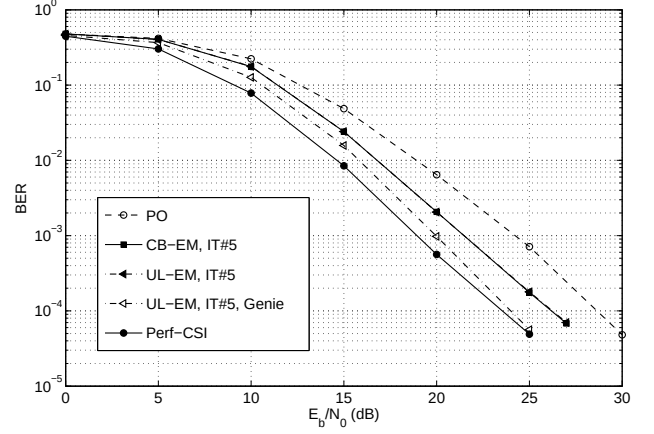


Fig. 7. Comparison of SB and PO channel estimation. $R = 4$, $N_s T = 64$, $N_p = 1$, QPSK modulation, $(5, 7)_8$ channel code.

of brevity).

To understand why the improved estimators, i.e., DB-, QS-, and UL-EM give poor improvement for small constellation sizes as for QPSK, we should come back to the calculation of $\mathbf{R}'_P$ that is required to update $\hat{\mathbf{h}}$ by this method. The calculation of $\mathbf{R}'_P$ was discussed at the last paragraph of Subsection V-C, where we explained that we use the hard-estimates of the transmitted symbols using the soft-decoder outputs to construct $\mathbf{P}$. As a matter of fact, the poor performance improvement of the modified estimators arises from the possible errors in these hard estimates. For instance, we have shown in Figs. 5 and 6 the hypothetical BER curve that is obtained if we use the exact transmitted symbols (i.e., the exact $\mathbf{P}$) in the calculation of $\mathbf{R}'_P$ using the UL-EM method. We denote this case by “UL-EM, Genie” and notice that it allows substantial improvement in the receiver performance, especially for the case of QPSK modulation. So, the limited performance improvement of UL-EM arises from the this implementation constraint. For larger constellation sizes, for the same level of BER, the performance is less affected by this hard estimation because it is done on a smaller “grid size” on the signal constellation plane.

Let us now consider the case of four relays. The corresponding performance results are shown in Fig. 7 for the case of QPSK modulation. We notice a substantial performance improvement by using SB estimation. The SNR improvement is about 2.5 dB in SNR at $\text{BER} = 10^{-4}$. Similar to $R = 2$ case, UL-EM does not allow any improvement, compared to CB-EM. It is also the case for DB-EM and QS-EM methods (results are not shown). We have also considered the case of 16-QAM modulation in Fig. 8. Note that, here, for the cases of PO estimation and perfect CSI, we process two iterations.⁷ We notice an improvement of 1.6 dB in E_b/N_0 at $\text{BER} = 10^{-4}$ by using the CB-EM method compared to PO. Also, the UL-EM method allows about 0.85 dB SNR improvement over CB-EM. Similar to the case of $R = 2$, we have shown the

⁷Two iterations are sufficient to attain the receiver convergence. Here, iterative detection helps the 16-QAM demodulator to benefit from the channel coding gain in symbol demapping.

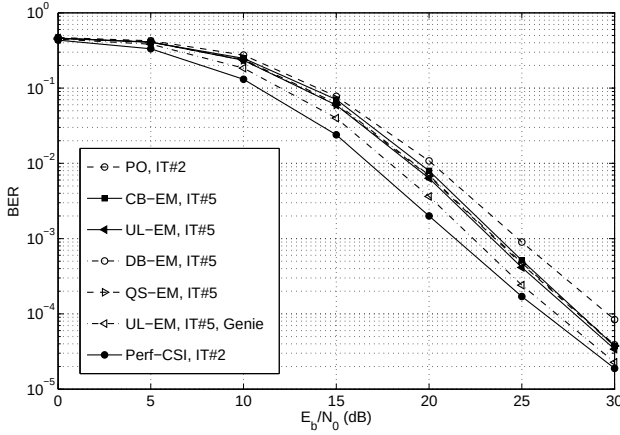


Fig. 8. Comparison of SB and PO channel estimation. $R = 4$, $N_s T = 64$, $N_p = 1$, 16-QAM modulation, $(5, 7)_8$ channel code.

BER curves for the “UL-EM, Genie” method in Figs. 7 and 8. We notice again that UL-EM suffers from the imprecision in the calculation of R'_P . On the other hand, the QS-EM outperforms DB-EM, and, interestingly, it even performs very close to UL-EM. At the target BER of 10^{-4} , we have about 0.2 dB gain by using QS-EM, compared to DB-EM, and we are about only 0.1 dB away from UL-EM. We have verified that the difference between the performances of QS- and UL-EM increases by decreasing the frame length. However, at the same time, they approach to the case of classical CB-EM estimation. It suggests that the simple and low-complexity QS-EM should be preferred to the other methods.

We have also studied the cases of block fading channels (equivalent to the case of OFDM signaling) and also when smaller power is devoted to pilot symbols, compared to the data symbols. Results showed improvement on the same order of that presented above. They are not provided for the sake of brevity.

Lastly, let us discuss the Gaussian approximation that we made in the detector formulation in (21) and (32). For the case of perfect partial CSI knowledge, we have compared the receiver performance when using this approximation with the case where we have a perfect knowledge of g_i at the destination, and noticed that there is a very small difference between the two cases. Furthermore, for the estimated CSI case, this approximation affects negligibly the performances of both CB-EM and UL-EM methods (results are not shown for the sake of brevity). Hence, using the simple Gaussian approximation is quite rational.

VII. CONCLUSIONS AND DISCUSSIONS

We considered SB channel estimation for AF cooperative relay networks based on the EM algorithm. We firstly considered the classical EM-based channel estimation and provided the estimator formulation for the AF WRN system. We called this approach CB-EM. Considering the two case studies of two and four relay nodes, we demonstrated through simulation results the substantial performance improvement that we obtain by SB estimation. This way, by devoting few

channel-uses to pilot transmission, and hence, minimizing the nodes' energy consumption and the pilot overhead, we can obtain a satisfying performance at the destination. Performance improvement with SB estimation is more significant for larger number of relays. In fact, for a larger R , benefiting from more distributed diversity, we have a better detection performance, and the contribution of the detected data symbols to channel estimation becomes more beneficial. Also, more significant performance improvement is obtained for increased frame size.

We further showed that the CB-EM formulation provides a biased channel estimate and studied three modifications to it in order to obtain an unbiased estimate in our special WRN system. We called these approaches DB-EM, QS-EM, and UL-EM. Studying the BER performance, we noticed that the modified estimators are advantageous for relatively large signal constellations. The inherent bias in the CB-EM estimate appears to be penalizing in such cases. For the simple QPSK modulation case, we obtain too small performance improvement that does not justify the increased computational complexity of the modified estimators. We showed that this is due to the difference between the transmitted symbols and their hard estimates that we need in their implementation. Under such conditions, we noticed that the simple QS-EM method would be preferable to two other methods. The reason that the UL-EM solution does not show a considerable advantage over QS-EM is that in developing the UL-EM estimator, we restricted ourselves to a linear combination of $\hat{h}^p$ and $\hat{h}^d$ in order to obtain a simple solution. This does not necessarily result in an optimum solution, however.

Finally, note that SB estimation is obviously more computationally complex than PO estimation. We assumed that there are less stringent constraints on the energy and computational resources of the destination node, compared to the source and the relays nodes. The increased computational complexity by SB estimation remains moderate and quite justified taking the significant performance improvement into account.

REFERENCES

- [1] A. Nosratinia, T. E. Hunter, and A. Hedayat, “Cooperative communication in wireless networks,” *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, Oct. 2004.
- [2] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity - Part I: system description,” *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [3] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity - Part II: implementation aspects and performance analysis,” *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.
- [4] J. N. Laneman and G. W. Wornell, “Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
- [5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [6] A. Chowdhery and R. K. Mallik, “Linear detection for the nonorthogonal amplify and forward protocol,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 826–834, Feb. 2009.
- [7] A. F. Dana and B. Hassibi, “On the power-efficiency of sensory and ad hoc wireless networks,” *IEEE Transactions on Information Theory*, vol. 52, pp. 2890–2914, July 2006.

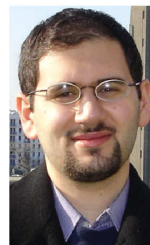
- [8] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3524–3536, Dec 2006.
- [9] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4106–4118, Nov. 2007.
- [10] Y. Jing and H. Jafarkhani, "Distributed differential space-time coding for wireless relay networks," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1092–1100, July 2008.
- [11] G. S. Rajan and B. S. Rajan, "Algebraic distributed differential space-time codes with low decoding complexity," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3962–3971, Oct 2008.
- [12] G. S. Rajan and B. S. Rajan, "Leveraging coherent distributed space-time codes for noncoherent communication in relay networks via training," *IEEE Transactions on Communications*, vol. 8, no. 2, pp. 683–688, Feb. 2009.
- [13] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1907–1915, May 2008.
- [14] S. Cui, A. J. Godsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1097, Aug 2004.
- [15] S. K. Jayaweera, "Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 984–989, May 2006.
- [16] Yi Zhang, M. A. Khalighi, M. S. Sadough, and S. Bourennane, "Signal detection for amplify-and-forward relay networks with imperfect channel estimation," *IEEE/IET International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, July 2012, Poznań, Poland.
- [17] M. A. Khalighi and J. J. Boutros, "Semi-blind channel estimation using EM algorithm in iterative MIMO APP detectors," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3165–3173, Nov. 2006.
- [18] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.
- [19] M. A. Khalighi, J.-F. Hélard, S. M. S. Sadough, and S. Bourennane, "Suitable combination of channel coding and space-time schemes for moderate-to-high spectral efficiency MIMO systems," *AEUE International Journal of Electronics and Communications*, vol. 64, no. 7, pp. 595–606, July 2010.
- [20] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for broadband multi-user communications, WLANs and broadcasting*, Wiley, 2004.
- [21] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.
- [22] M. Malkawi and I. M. Kim, "Hard/soft detection with limited CSI for multi-hop systems," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 3435–3441, July 2009.
- [23] P. Liu and I. M. Kim, "Optimum/sub-optimum detectors for multi-branch dual-hop amplify-and-forward cooperative diversity networks with limited CSI," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 78–84, Jan 2010.
- [24] B. Gedik, O. Amin, and M. Uysal, "Power allocation for cooperative systems with training-aided channel estimation," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4773–4783, Sept. 2000.
- [25] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 1–4, Jan. 2001.
- [26] A. Dempster, N. Laird, and D. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *Journal of Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, Jan. 1977.
- [27] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, Nov. 1996.
- [28] K. Fang, L. Rugini, and G. Leus, "Low-complexity block turbo equalization for OFDM systems in time-varying channels," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5555–5566, Nov. 2008.



Yi Zhang received his Ph.D. degree from École Centrale de Marseille, France, in 2012. He is currently employed as engineer in the Second Research Institute of CAAC, China. His research interests include channel estimation, signal detection, and ADS-B.



Mohammad Ali Khalighi received his Ph.D. in Electrical Engineering (Telecommunications) from Institut National Polytechnique de Grenoble (INPG), France, in 2002. From 2002 to 2005, he has been with GIPSA-lab, Tlcom Paris-Tech (ENST), and IETR-lab, as a post-doctoral research fellow. He joined Ecole Centrale Marseille and Institut Fresnel in 2005 as an Assistant Professor. His main research areas of interest include optical wireless communication and wireless sensor networks.



Seyed Mohammad Sajad Sadough received his 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 Suplec, 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, Suplec-CNRS, where he was involved in the European research project TVMSL with Alcatel-Lucent France. Since October 2008, he has been an Assistant Professor in the Department of Telecommunication, Shahid Beheshti University. His main areas of research include communication theory, digital communication and cognitive radio.



Salah Bourennane received his Ph.D. degree from Institut National Polytechnique de Grenoble (INPG), France, in signal processing. Currently, he is a full Professor at cole Centrale Marseille and he is the head of the Multidimensional Signal Processing team at Institut Fresnel, Marseille, France. His research interests include statistical signal processing, array processing, image processing, multidimensional signal processing, and telecommunications.