Blind Digital Modulation Identification for Spatially-Correlated MIMO Systems

Kais Hassan, Iyad Dayoub, *Member, IEEE*, Walaa Hamouda, *Senior Member, IEEE*, Crépin Nsiala Nzéza, and Marion Berbineau

***Abstract*—Modulation type is one of the most important characteristics used in signal waveform identification and classification. Spatial correlation is a crucial factor for practical multiple-input multiple-output (MIMO) systems. This paper addresses the problem of blind digital modulation identification in spatially-correlated MIMO systems. The proposed algorithm is verified using higher order statistical moments and cumulants of the received signal. The purpose is to discriminate among different M-ary shift keying linear modulation schemes without any priori signal information. This study employs several MIMO techniques to identify the modulation with and without channel state information (CSI). The proposed classifier shows a high identification performance in acceptable signal-to-noise ratio (SNR) range.**

***Index Terms*—Higher order statistics, multiple-input multiple-output systems, modulation identification, spatial correlation.**

I. Introduction

Nowadays, multiple-input multiple-output (MIMO) technology is considered one of the promising technologies for developing the next generation of wireless systems. Recently, blind algorithms and techniques for MIMO signals interception have gained more attention. One essential step in the signal interception process is to blindly identify the modulation scheme of MIMO signals. Modulation identification has its roots in military applications such as; communication intelligence (COMINT), electronic support measures (ESM) and spectrum surveillance. Also recent and rapid developments in software defined radio (SDR) in the context of cognitive radio (CR) have given modulation identification more prominence in civil applications.

Many modulation identification algorithms have been developed for single-input single-output (SISO) systems [1]. These algorithms are generally divided into two categories. The first category is based on decision theoretic approach while the second on pattern recognition. The decision theoretic approach is a probabilistic solution based on a priori knowledge of probability functions and certain hypotheses [1], [2]. On the other hand, the pattern recognition approach is based on extracting some basic characteristics of the received signal called features [3]–[6]. This approach is generally divided into two subsystems: the features extraction subsystem and the classifier subsystem. However, the second approach is easier to implement and reaches a quasi-optimal performance if the proper features set is chosen. Recently, much work has been conducted on modulation identification [3]–[6]. The identification techniques, which have been employed to extract the signal features necessary for digital modulation identification, include spectral based features set [3], higher order statistics (HOS) [4], cyclostationarity signatures [5], and wavelets transforms [6]. With their efficient performance in pattern recognition problems (e.g., modulation classification), many studies have proposed the application of artificial neural networks (ANNs) as classifiers [1], [3].

In [7], Swami *et. al* proposed a simple yet very low complexity method, based on elementary fourth-order cumulants for the classification of digital modulation schemes. This method was applied in a hierarchical manner to classify various digital modulation schemes. Also classification thresholds developed by deriving the expressions for the variance of the estimates of the cumulants. Furthermore, the statistics used by the classifier can be recursively updated. The robustness of this approach comes about not only from the resistance of HOS to additive colored Gaussian noise, but also from a natural robustness to constellation rotation and phase jitter.

So far most of the work on modulation identification did not address MIMO systems. For instance, Choqueuse *et. al* [8] proposed two likelihood based modulation classifiers for uncorrelated MIMO systems. The first one, called average likelihood ratio tests (ALRT), is optimal in the Bayesian sense but requires a perfect channel state information (CSI). The second classifier, called hybrid likelihood ratio tests (HLRT), approximates the ALRT by replacing the channel matrix with its estimate. The channel is estimated in two steps by using an independent component analysis and a phase correction algorithm respectively. Simulations showed that the two classifiers perform well, for example, perfectly recognizing 2-PSK, 4-PSK, 16-PSK and 16-QAM modulations at a SNR of 5 dB.

The most commonly used MIMO channel model is the spatially independent identically distributed (i.i.d.) flat-fading channel. This corresponds to the rich scattering narrowband scenario. It was soon realized, however, that many propagation environments result in spatial correlation. Hence, spatial correlation is a crucial factor for practical MIMO systems and its effect on their performance must be evaluated. In this

Manuscript received February 7, 2011; revised August 24, 2011; accepted November 11, 2011. The associate editor coordinating the review of this paper and approving it for publication was D. Reynolds.

K. Hassan is with the Univ. Lille Nord de France, F-59000 Lille, IFSTTAR, LEOST, F-59650 Villeneuve d'Ascq, IEMN, DOAE, F-59313 Valenciennes (e-mail: kais.hassan@ifsttar.fr).

I. Dayoub is with the Univ. Lille Nord de France, F-59000 Lille, IEMN, DOAE, F-59313 Valenciennes (e-mail: iyad.dayoub@univ-valenciennes.fr).

W. Hamouda is with Concordia University, Montreal, QC, Canada H3G 1M8 (e-mail: hamouda@ece.concordia.ca).

C. N. Nzéza and M. Berbineau are with the Univ. Lille Nord de France, F-59000 Lille, IFSTTAR, LEOST, F-59650 Villeneuve d'Ascq (e-mail: {crepin.nsiala, marion.berbineau}@ifsttar.fr).

Digital Object Identifier 10.1109/TWC.2011.122211.110236

paper, we employ the widely accepted Kronecker model [9] for spatially-correlated MIMO systems. The correlation matrix is modeled by a single coefficient exponential correlation model [10].

Contrary to [8] we propose a pattern recognition approach. Also, to the best of our knowledge, no work has yet considered the problem of modulation identification in correlated MIMO channels. Moreover, in this paper we examine the effect of channel estimation error on our proposed modulation identification algorithm. In our approach, the features extraction subsystem is based on the higher order cumulants (HOC) and the higher order moments (HOM) of the processed received signal. Our proposed classifier is a multi-layer artificial neural network trained using the resilient backpropagation learning algorithm (RPROP).

Here, modulation identification is performed without any priori information of the received signal (e.g. probability functions, noise statistics, etc.). The MIMO channel mixes the modulated transmitted symbols and change their statistical properties, therefore, we propose three algorithms to identify the modulation: the first depends on the HOS of the received mixed symbols and called direct digital modulation identification (D-DMI) while the second and the third estimate the transmitted symbols before employing their original statistical properties. By employing zero-forcing (ZF) technique we propose the ZF-DMI algorithm. When blindly separating the MIMO symbols using simplified constant modulus algorithm (SCMA), we propose an SCMA-DMI algorithm. Our proposed algorithm is considered as semi-blind when assuming a perfect CSI knowledge at the receiver side. Conversely, this algorithm is completely blind when using erroneous channel estimation or when it blindly separates the source symbols. The proposed algorithm has the capability to identify different M-ary shift keying linear modulation types (ASK, QAM or PSK) and the order of the identified modulation type. The performance of our algorithm is examined through the probability of identification. Our results show high identification performance at acceptable signal-to-noise ratio (SNR) range.

The remainder of the paper is organized as follows: Section II defines the system model and introduces the different assumptions. Section III describes the proposed modulation identification algorithm while section IV focuses on the performance analysis. The results and algorithm performance evaluation are presented in section V. Finally, conclusions and perspectives of the research work are presented in section VI.

II. System Model

In this section we define first the MIMO signal model. Secondly the spatial correlation model is introduced since it is the practical case that we address. Then, two MIMO equalization algorithms are presented as they are employed to reveal some ambiguity from the received symbols.

A. MIMO Signal Model

A MIMO system with N_t transmit antennas and N_r receive antennas is considered ($N_r \geqslant N_t$). Under the assumption of a frequency flat and time invariant MIMO channel, the baseband received symbol vector at the instant k is described as:

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{y}(k) = [y_1(k), ..., y_{N_r}(k)]^T$ is an $(N_r \times 1)$ received signal vector without any time oversampling and optimum symbol timing and with perfect carrier frequency and phase estimation, $\mathbf{x}(k) = [x_1(k), ..., x_{N_t}(k)]^T$ is the $(N_t \times 1)$ vector representing the transmitted source signals, and $\mathbf{n}(k) = [n_1(k), ..., n_{N_r}(k)]^T$ is an $(N_r \times 1)$ vector corresponds to the additive zero-mean white circularly complex Gaussian noise with variance σ_n^2; i.e. $\mathbf{n}(k) \backsim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_r})$, where $\mathbf{I}_{N_r}$ is the identity matrix of size N_r. $\mathbf{H}$ corresponds to the $(N_r \times N_t)$ spatially-correlated complex matrix of the MIMO channel. These source symbols $\mathbf{x}$ are i.i.d and mutually independent. The symbols are assumed to belong to the same linear modulation scheme. We will assume, without loss of generality, that the source constellations are normalized to have zero-mean and unit energy.

B. Spatial Correlation Model

Here, the spatially-correlated MIMO channels are modeled by the Kronecker model [9]. This model assumes that the spatial receiver and transmitter correlations to be separable. It can be shown that, under the above assumption, the Kronecker model is given by:

$$\mathbf{H} = \mathbf{R}_r^{1/2} \mathbf{H}_w \mathbf{R}_t^{1/2} \tag{2}$$

where $\mathbf{R}_t$ and $\mathbf{R}_r$ are the transmitter and the receiver correlation matrices. $\mathbf{H}_w$ is a full rank gain matrix which the entries are i.i.d and follow a circularly symmetric complex Gaussian distribution with zero-mean and unit variance. Here, the correlation matrices are presented by the exponential correlation model which was introduced in [10] and extended in [11]. This model defines the entries of a correlation matrix $\mathbf{R}$ by a single coefficient as:

$$[\mathbf{R}]_{ij} = \begin{cases} \rho^{j-i}, & i \leq j \\ [\mathbf{R}]_{ji}{}^*, & i > j \end{cases}, \quad |\rho| < 1, \tag{3}$$

where ρ is the complex correlation coefficient of neighboring antennas and $[.]_{ij}$ denotes the entry in the ith row and the jth column of the matrix. The two matrices $\mathbf{R}_t$ and $\mathbf{R}_r$ are , respectively, defined by the transmit correlation coefficient ρ_t and the receive correlation coefficient ρ_r.

C. MIMO Equalization

In this subsection we present two equalization algorithms that are employed in our study: the zero-forcing (ZF) algorithm and simplified constant modulus algorithm (SCMA).

1) Zero-Forcing Equalizer: The ZF technique is applied to reveal some ambiguity from the received MIMO symbols. This technique consists of applying an equalizing matrix $\mathbf{W}$ on the received vector. This matrix $\mathbf{W}$ is defined by $\mathbf{W} = \mathbf{H}^\dagger = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$ where $\mathbf{H}^\dagger$ denotes the pseudo inverse operation. The transmitted symbols is estimated by:

$$\hat{\mathbf{x}}(k) = \mathbf{W}\mathbf{y}(k) \stackrel{\text{def}}{=} ZF(\mathbf{y}) = \mathbf{x}(k) + (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{n}(k). \tag{4}$$

Then the estimated vector $\hat{\mathbf{x}}(k) = [\hat{x}_1(k), ..., \hat{x}_{N_t}(k)]^T$ is the input of the modulation identifier or more precisely the input of the features extraction subsystem.

Here we assume perfect CSI at the receiver side (semi-blind classifier). If it is not the case, channel estimation has to be performed and the modulation is blindly identified. First, we are interested in investigating the impact of channel estimation error on the modulation identification. Hence, we model the estimated channel as:

$$\hat{\mathbf{H}} = \mathbf{H} + \sigma_e \mathbf{\Omega} \tag{5}$$

where the entries of $\mathbf{\Omega}$ are i.i.d with zero-mean circularly symmetric complex Gaussian variables with unity variance and σ_e^2 represents the variance of the channel estimate error.

2) Simplified Constant Modulus Algorithm: Another improvement of blind modulation classifier is to blindly separate the MIMO source symbols and then identify the modulation. Several blind source separation (BSS) algorithms exist in the literature [12], [13]. Here, the SCMA is used to blindly recover the transmitted symbols [12]. The SCMA objective is to find an $N_r \times N_t$ matrix $\mathbf{W}$ such as:

$$\mathbf{z}(k) = \mathbf{W}^H \mathbf{y}(k) \stackrel{\text{def}}{=} SCMA(\mathbf{y}) = \mathbf{W}^H \mathbf{H}\mathbf{x}(k) + \widetilde{\mathbf{n}}(k). \tag{6}$$

The purpose is to find $\mathbf{W}$ such that $\mathbf{z}(k) = \hat{\mathbf{x}}(k)$, but the transmitted symbols $\mathbf{x}$ are usually determined up to a permutation and a scalar multiple, i.e. $\mathbf{G}^H = \mathbf{W}^H\mathbf{H} = \mathbf{P}\mathbf{\Lambda}$ where $\mathbf{\Lambda}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix that introduces the arbitrary phase and permutation. In fact, the SCMA simplifies the constant modulus criterion by employing a single dimension, e.g. the real part of the signal. The SCMA attempts to minimize the following cost function:

$$\begin{cases} \mathrm{J}_{SCMA}(\mathbf{W}) = \sum_{i=1}^{N_t} E\left[(\Re(\mathbf{z}_i(k))^2 - R)^2\right] \\ \text{Subject to : } \mathbf{W}^H\mathbf{W} = \mathbf{I}_{N_t} \end{cases} \tag{7}$$

where $R = \frac{E[\Re(x(k))^4]}{E[\Re(x(k))^2]}$ is the dispersion constant. This criterion leads to a complexity reduction and ensures a non-arbitrary constellation rotation of $\pi/2$ multiples for each data stream at the output of the equalizer. This justifies our choice of such cost function among many others that exist in the literature. In fact, any multiple $\pi/2$ rotation will not affect the statistical properties of the equalized symbols since each rotated symbol still belongs to the same constellation.

Note that the SCMA is implemented using the stochastic gradient (SG) algorithm. The equalizer update equation is obtained by calculating the gradient of J_{SCMA}:

$$\widetilde{\mathbf{W}}_n(k) = \widetilde{\mathbf{W}}_n(k-1) - \mu(k)\mathbf{e}_n(k)\underline{\mathbf{y}}^*(k), \quad n \in \{1, \ldots, N_t\} \tag{8}$$

where μ is the SG step size which is updated by the time averaging adaptive step size (TAASS) mechanism [14]. The error signal e_n is given by:

$$\mathbf{e}_n(k) = (\Re(\mathbf{z}_i(k))^2 - R)\,\Re(\mathbf{z}_i(k)), \tag{9}$$

and $\underline{\mathbf{y}} = \mathbf{F}^H\mathbf{y}$ is the pre-whitened received signals. The pre-whitening method proposed in [15] is used in this paper. Also, Gram-Schmidt orthogonalization algorithm allows us to satisfy the orthogonalization constraint in (7) at each iteration [13].

Here we assume that the number of transmitting antennas is known at the receiver which is not true in most blind scenarios. However, estimating the number of sources is a well investigated problem in the literature. For instance, the authors in [16] reviewed and compared several source number detection methods. Also, detecting the used MIMO coding can be employed to improve the equalization performance and subsequently the modulation identification performance. In [17], the authors review and compare several MIMO coding detectors.

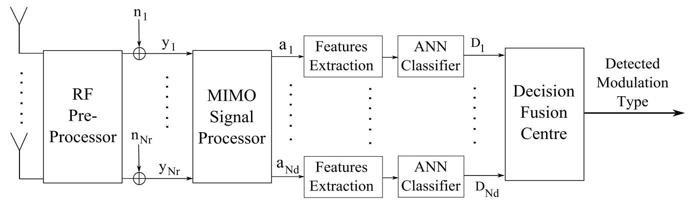

Fig. 1. A block diagram of the proposed identification scheme.

III. Proposed Modulation Identification Algorithm

The MIMO channel mixes the modulated transmitted symbols and corrupts their statistical properties which affects the modulation identification performance. Here, we propose two approaches to identify the modulation, the first depends on the HOS of the received mixed symbols and the second attempts to estimate the transmitted symbols before employing their original statistical properties which were proved to be very beneficial.

The block digram of the proposed modulation identification algorithms is shown in Fig. 1. After the RF pre-processing, the N_r received MIMO streams mixed by the correlated MIMO channel are the input of MIMO signal processor block which outputs N_d streams. We propose to either identify the modulation directly from the received symbols or to estimate the transmitted symbols before the identification. Two algorithms are employed for equalization, the first is ZF algorithm which could be employed with and without CSI while the second is the blind SCMA. The features are extracted for each one of the N_d output streams of the MIMO signal processor. The modulation is identified for each one among these N_d streams separately. An ANN based classifier is employed to output N_d decision vectors. These decision vectors are combined for final decision. In this section we introduce the MIMO signal processor and three proposed algorithms before presenting the features extraction and the classification processes. Finally, the decision fusion method is explained.

A. MIMO Signal Processor

The output of the MIMO signal processor block is the signal vector $\mathbf{a}(k) = [a_1(k), ..., a_{N_d}(k)]^T$, where $N_d = (N_r$ or $N_t)$ depending on the considered algorithm. The MIMO signal processor is simply introduced to explain the three proposed algorithms as follows:

a) The received signals are directly used for identification, i.e. $\mathbf{a}(k) = \mathbf{y}(k)$. We aim to identify the modulation scheme blindly without any priori signal information despite the channel mixing effect. The influence of the channel mixing on the probability of modulation identification is examined. This algorithm is called direct digital modulation identification (D-DMI).

b) ZF equalizing precedes the identification, i.e. $\mathbf{a}(k) = ZF(\mathbf{y}(k))$ (see eq. 4). In this algorithm a perfect CSI knowledge at the receiver side is assumed when the proposed algorithm is considered as semi-blind. Conversely, this algorithm could be considered blind when using erroneous channel estimation. The influence of the channel estimation error on the probability of modulation identification is examined. This algorithm is called ZF-DMI.

c) An SCMA equalizer employed to blindly separate the MIMO sources, i.e. $\mathbf{a}(k) = SCMA(\mathbf{y}(k))$ (see eq. 6). Here, we are interested in estimating the transmitted symbols instead of considering erroneous channel estimation, i.e. our proposed algorithm is completely blind. This algorithm is called SCMA-DMI.

B. Features Extraction

One of the important aspects of modulation identification is the selection of the proper identification features. Previous works have shown that higher order cumulants (HOC) and higher order moments (HOM) of the received signal are among the best candidates for signal identification in SISO systems [4], [7]. Higher order moments of a signal x are defined by [18]:

$$M_{km}(x) = E[x^{k-m}(x^*)^m] \tag{10}$$

where k is the moment order. The cumulant of order k of the zero-mean signal x is defined by:

$$C_{km}(x) = Cum[\underbrace{x, ..., x}_{(k-m)\ times}, \underbrace{x^*, ..., x^*}_{m\ times}]. \tag{11}$$

Also, the relation between moments and cumulants can be expressed as:

$$Cum[x_1, .., x_n] = \sum_{\Phi}(\alpha - 1)!\,(-1)^{\alpha-1} \prod_{v\in\Phi} E(\prod_{i\in v} x_i) \tag{12}$$

where Φ runs through the list of all partitions of $\{1, ..., n\}$, v runs through the list of all blocks of the partition Φ, and α is the number of elements in the partition Φ. Based on (12), moments estimation leads to estimate the cumulants. That is given a signal x with N samples, one can estimate the moments as:

$$\hat{M}_{km}(x) = \frac{1}{N}\sum_{i=1}^{N} x^{k-m}(i)x^{*m}(i). \tag{13}$$

In the above we assume, without loss of generality, that the signal x is normalized to have a unity energy, i.e. $C_{21} = 1$. This will remove any scale problems in the estimators. Practically the self-normalized HOS are calculated as:

$$\tilde{M}_{km}(x) = \hat{M}_{km}(x)/\hat{M}_{21}^{k/2}(x),\ \tilde{C}_{km}(x) = \hat{C}_{km}(x)/\hat{C}_{21}^{k/2}(x). \tag{14}$$

Note that the authors in [7] used the cumulants up to the fourth order and the hierarchical classification for modulation identification in SISO systems. Based on that, the features employed in this paper consists of a combination of HOM and HOC up to order six. We consider N consecutives processed vectors $\mathbf{a}(k), \mathbf{a}(k-1), ..., \mathbf{a}(k-N-1)$. The HOS are calculated for the N_d streams; i.e. $a_1, ..., a_{N_d}$.

Note that the complexity of (13) is of order N where estimating a moment of order k requires only about N complex additions and $k \times N$ complex multiplications. Based on (12), cumulant calculation is of order N. Of course, the computational cost of the features calculation for each of the above mentioned N_d signals is of the same order. Then, the features extraction process has a very low complexity $\mathcal{O}(N)$.

C. ANN Classifier

After extracting the proper features, the modulation identification problem can be considered as a pattern recognition problem. Knowing that ANN is one of the best solutions for pattern recognition problems, many researchers have focused on ANNs to develop high performance modulation classifiers [1], [3]. In this paper, the proposed classifier is a multilayer feed-forward ANN. The HOS features are the inputs of this trained ANN. Each one among the N_d processed streams is identified with the same ANN, i.e. the neural networks in the N_d identification branches shown in Fig. 1 are identical. This totally yields to N_d decision vectors. These decision vectors are then combined to generate the final decision.

The neural network structure including the number of hidden layers, the number of nodes in each layer and the transfer function of each node has been chosen through intensive simulations. This structure is directly related to network training speed and identification precision. Speeding the learning process of the network and improving the identification accuracy can be achieved by normalizing the features set (fifteen extracted ones) and selecting the optimal subset for the discrimination process. Here, a feature subset selection based on the principal component analysis (PCA) is applied to select the best subset of the combined HOM and HOC features set.

First, the extracted set of features are normalized before subset selection to ensure that they are of zero mean and unit variance. Then, the PCA technique constructs a low-dimensional representation of the data (normalized features) that describes as much of the variance in that data as possible. The PCA is known as a linear transformation that orthogonalizes the components of the data and orders the resulting orthogonal components so that those with the largest variation come first (the principal component is the first one). This moves as much of the variance as possible into the first few components. The remaining components, therefore, tend to be highly correlated and may be dropped with minimal loss of information. Then, the selected subset is the orthogonal components with the largest variance. Here, only six orthogonal components (out of fifteen) are selected for both ANN training and testing.

After features subset selection, the training process is triggered. The initiated ANN is trained using the resilient backpropagation learning algorithm (RPROP) introduced in [19]. Beside the fast convergence, one of the main advantages of RPROP lies in the fact that no choice of parameters and initial values is needed at all to obtain optimal or at least nearly optimal convergence times [19]. Also, RPROP is known by its high performance on pattern recognition problems. After training, a test phase is launched, and the classifier is evaluated through the probability of identification.

Since the outputs of a layer in ANN are considered as linear combinations among the inputs of this layer, then the computational cost of the classifier is related to the number of nodes at each layer. Considering the static and predefined structure of ANN, and the small number of nodes at each layer, the required number of operations to obtain the classifier output is fixed and inexpensive.

D. Decision Fusion

Since there are multiple antennas at the receiver, it is possible for them to cooperate to achieve higher identification reliability. The modulation scheme of each one among the N_d processed streams is identified (or decided) independently and all the decisions vectors are jointly processed to make a final decision. We call this cooperative identification scheme as decision fusion. The final decision is made by combining the N_d decision vectors and applying the M-out-of-N_d decision fusion rule, i.e. a certain modulation scheme is identified when it is decided on M classifier among the N_d classifiers. Note that the M-out-of-N_d decision fusion rule includes the logical OR (LO) ($M=1$), the logical AND (LA) ($M=N_d$), and the majority ($M=\lceil N_d/2 \rceil$) as special cases, where $\lceil . \rceil$ denotes the ceiling function. When the decision vectors do not fit the decision fusion rule, the final decision is *rejected*, e.g. when $N_d = 4$ and there are four different decisions, the trial is rejected.

IV. Performance Analysis

In this section we present a performance analysis of our work. First the performance results of modulation identification based on HOS in SISO systems are given. Then the performance of the D-DMI algorithm is presented by studying the channel mixing effect. Also, we study the performance of the ZF-DMI algorithm when employed with or without perfect CSI. The effect of spatial correlation is introduced and finally we explain how the decision fusion improves the overall performance.

A. Performance Results in SISO systems

The hierarchical modulation identification for SISO systems using HOS (up to four) was employed in [7]. It was shown that the probability of modulation identification is a function of SNR, number of received symbols N and the considered modulation pool. The identification performance depends on the accuracy of HOS estimation. Also the variance of HOS estimators depends on the SNR and N. It is clear that increasing N will decrease the variance and improves the overall performance. Anyway, simulations show that about 500 symbols are sufficient to produce good estimates of HOS. The probability of identification was higher than 99% when the SNR is not lower than 5dB and 10dB when considering, respectively, four and eight modulation pools. Note that the probability of identification P_i is an increasing function of SNR when considering a known modulation pool and a sufficient number of samples for SISO systems. We define the minimum required SNR for close to optimal modulation identification, SNR_{min}, by $\frac{P_i(\infty)-P_i(\text{ SNR}>\text{ SNR}_{min})}{P_i(\infty)} < \epsilon$, where normally, $\epsilon = 0.01$ and $P_i(\infty) = 100\%$. Based on our simulations, we noted that the statistics of order higher than six will not improve the probability of identification. Therefore, in our work, we consider statistics of order up to six to extract the needed features for modulation identification.

TABLE I
Some Theoretical Statistical Moments and Cumulants Values for Different Modulation Schemes of Interest [1], [5], [7]

	2-PSK	4-PSK	8-PSK	4-ASK	8-ASK	16-QAM	64-QAM
C20	1	0	0	1	1	0	0
M40	1	1	0	1.64	1.77	-0.67	-0.18
M41	1	0	0	1.64	1.77	0	0
M42	1	1	1	1.64	1.77	1.32	1.34
C40	-2	1	0	-1.36	-1.24	-0.68	-0.62
C41	-2	0	0	-1.36	-1.24	0	0
C42	-2	-1	-1	-1.36	-1.24	-0.68	-0.62
M60	1	0	0	2.92	3.62	0	0
M61	1	-1	0	2.92	3.62	-1.32	0.38
M63	1	1	1	2.92	3.62	1.96	2.08
C60	16	0	0	8.32	7.19	0	0
C61	16	-4	0	8.32	7.19	2.08	1.8
C62	16	0	0	8.32	7.19	0	0
C63	16	4	4	8.32	7.19	2.08	1.8

Theoretical values of some HOS are given in Table I to show how they can discriminate the different modulation schemes of interest. These theoretical values are calculated by averaging HOS for different digital modulation constellations under the constraints of unit variance symbols and noise free case.

B. Channel Mixing Effect

The modulation scheme is directly identified from the received signal vector and no CSI knowledge is required when considering the D-DMI algorithm. Each stream of the received MIMO signal vector $\mathbf{y}$ is given by:

$$y_j = \sum_{i=1}^{N_t} h_{ij} x_i + n_j \tag{15}$$

where h_{ij} is the complex channel gain between the i^{th} transmit antenna and the j^{th} receive antenna. The noise-free case is considered to separately study the effect of the MIMO channel. Since the transmitted symbols are i.i.d and mutually independent then the self-normalized HOS are given by $\tilde{M}_{km}(y_j) = \lambda_{km}(j)\, \tilde{M}_{km}(x^c)$ and $\tilde{C}_{km}(y_j) = \lambda_{km}(j)\, \tilde{C}_{km}(x^c)$, where x^c is an arbitary stream of the N_t source ones and $\lambda_{km}(j) = \frac{\sum_{i=1}^{N_t} h_{ij}^{k-m} h_{ij}^{*\,m}}{(\sum_{i=1}^{N_t} |h_{ij}|^2)^{k/2}}$ [4]. The identification process is done for each one of the N_r received streams before the decision fusion.

Based on the definition of the channel complex gains, we conclude that the random variables $\{\lambda_{km}(j), 1 \le j \le N_r\}$ are statistically identical and independent of j and hence we refer to them as λ_{km}. Then, the HOS of digitally modulated signals for the ideal noise-free case (Table I) are multiplied by a random factor for each received stream in the MIMO system. It is clear that $|\lambda_{km}| < 1$, i.e. the channel mixing drives the received signals to be more Gaussian. This effect will corrupt the HOS and hence the identification process which in turn degrades the overall identification percentage, requiring a higher SNR.

C. Zero-Forcing Performance

The ZF equalizer estimates the N_t transmitted streams as given in (4). The post-processing SNR of the nth stream for the ZF equalizer is given by [20]:

$$\eta_n = \frac{\eta_0}{[(\mathbf{H}^H\mathbf{H})^{-1}]_{nn}} = \eta_0\kappa_n, \quad 1 \leq n \leq N_t, \qquad (16)$$

where η_0 is the averaged SNR at each receive antenna. It was shown in [21] that κ_n is a weighted chi-square distributed random variable with $K = 2(N_r - N_t + 1)$ degrees of freedom when considering uncorrelated Rayleigh fading channel. Since the distribution of κ_n is independent of the subscript n, we denote the effective post-processing SNR by η.

It was shown above that the identification performance for each detected stream is an increasing function of the effective SNR. The effect of the ZF equalizer on the overall performance is presented by the statistical properties of η which is the actual SNR at the input of each one among the N_t classifier branches. The Cumulative distribution function $F_K(x)$ of $\eta \backsim \mathcal{X}^2(K)$ is a decreasing function of K, for any fixed $x \geq 0$ [21]; i.e. $P(\eta > x)$ increases when K increases. Then the overall performance depends on K and improves when $\Delta = N_r - N_t$ increases. Then, the performance for single-input multiple-output (SIMO) system is always better than that for MIMO system when using the same number of antennas at the receiver.

D. Channel Correlation Effect

Based on the Kronecker correlation model presented in (2) and the fact that $\mathbf{H}_w \backsim \mathcal{CN}(\mathbf{0}_{N_r \times N_t}, \mathbf{I}_{N_r} \otimes \mathbf{I}_{N_t})$, we conclude that $\mathbf{H} \backsim \mathcal{CN}(\mathbf{0}_{N_r \times N_t}, \mathbf{R}_r \otimes \mathbf{R}_t)$ where $\mathbf{0}_N$ is the zero matrix of order N. Here, we will consider the presence of only transmit correlation. Since $\mathbf{H}$ is complex normally distributed matrix then $\mathbf{Z} = \mathbf{H}^H\mathbf{H}$ is complex Wishart matrix [21], i.e. $\mathbf{Z} \backsim \mathbf{W}_{N_t}(N_r, 2\mathbf{R}_t)$. In this case the post-processing SNR of the nth ZF equalized stream is a weighted Chi-squared variable distributed as [20]:

$$f(\eta_n) = \frac{\lambda_{n,c}\exp\left(-\frac{\eta_n\lambda_{n,c}}{\eta_0}\right)}{\eta_0\Gamma(N_r - N_t - 1)}\left(\frac{\eta_n\lambda_{n,c}}{\eta_0}\right)^{(N_r - N_t)}, \; 1 \leq n \leq N_t, \qquad (17)$$

where $\Gamma(.)$ denotes the Gamma function and $\lambda_{n,c}$ is the effective SNR degradation due to transmit correlation and it is equal to $[\mathbf{R}_t^{-1}]_{nn}$. The matrix inversion is given by $\lambda_{n,c} = [\mathbf{R}_t^{-1}]_{nn} = \frac{\det[\mathbf{R}_t^{nn}]}{\det[\mathbf{R}_t]}$, where det(.) denotes the determinant of a matrix and $\det[\mathbf{R}_t^{nn}]$ is the minor of the matrix $\mathbf{R}_t$. It is easy to show that:

$$\lambda_{n,c} = \begin{cases} 1/(1 - |\rho_t|^2), & n = 1, N_t \\ (1 + |\rho_t|^2)/(1 - |\rho_t|^2), & n = 2, \ldots, N_t - 1 \end{cases} \qquad (18)$$

It is clear that when the channel is highly correlated ($|\rho_t| \to 1$), the effective SNR degradation due to correlation is more important ($\lambda_{n,c} \to \infty$). We can see that the SNR degradation of the first and last received streams is less than the remaining $(N_t - 2)$ streams. The fact that the performance improves when Δ increases is also valid in the case of correlated MIMO channels as it is clear in (17).

E. Channel Estimation Error Effect

Here, we are interested in investigating the impact of channel estimation error on the modulation identification rather than the estimation process. An erroneous channel estimation is modeled in (5). Hence, the ZF equalizer output is given by:

$$\hat{\mathbf{x}} = \hat{\mathbf{W}}\,(\mathbf{H}\mathbf{x} + \mathbf{n}) \qquad (19)$$

where $\hat{\mathbf{W}} = \hat{\mathbf{H}}^\dagger = (\mathbf{H} + \sigma_e\mathbf{\Omega})^\dagger$. The pseudo inverse $\hat{\mathbf{H}}^\dagger$ can be approximated by Taylor expansion when assuming $\sigma_e \ll 1$ as follows:

$$\hat{\mathbf{W}} \cong \mathbf{H}^\dagger(\mathbf{I}_{N_r} - \sigma_e\mathbf{\Omega}\mathbf{H}^\dagger). \qquad (20)$$

The estimated signal vector can be written as $\hat{\mathbf{x}} = \mathbf{x} + \hat{\mathbf{n}}$ where $\hat{\mathbf{n}}$ is given by:

$$\hat{\mathbf{n}} = \mathbf{H}^\dagger\mathbf{n} - \sigma_e\mathbf{H}^\dagger\mathbf{\Omega}\,\mathbf{x} - \sigma_e\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}^\dagger\mathbf{n}. \qquad (21)$$

It was shown in [22] that the post-processing SNR for each estimated stream, when considering uncorrelated Rayleigh channel, is given by:

$$\eta_n = \frac{\eta_0}{[1 + \sigma_e^2N_t\eta_0 + \sigma_e^2 tr((\mathbf{H}^H\mathbf{H})^{-1})]\,[(\mathbf{H}^H\mathbf{H})^{-1}]_{nn}}, \quad 1 \leq n \leq N_t. \qquad (22)$$

This result is also valid for correlated MIMO channels when the correlation model is defined as in (3). Hence, the post-processing SNR of the nth stream can be approximated by [22]:

$$\eta_n = \frac{\eta_0}{(1 + \sigma_e^2N_t\eta_0)\,[(\mathbf{H}^H\mathbf{H})^{-1}]_{nn}} = \frac{\eta_0\kappa_n}{\lambda_e}, 1 \leq n \leq N_t, \qquad (23)$$

where $\lambda_e = 1 + \sigma_e^2N_t\eta_0$ is the SNR degradation owing to the imperfect CSI. It is clear that the performance does not only depend on Δ but is also a function of N_t and σ_e. The reason for the dependence on N_t is that the inter-stream interference cannot be cancelled perfectly in the presence of channel estimation error [22]. For large SNR ($\eta_0 \to \infty$), the percentage $\eta_0/\lambda_e \to 1/\sigma_e^2N_t$, which leads to an upper bound for the identification percentage contrary to the perfect CSI case.

F. Decision Fusion Performance

The final decision is made by combining the N_d decision vectors and applying the M-out-of-N_d decision fusion rule. The total probability of identification of the modulation scheme is given by:

$$\mathbb{P}_i = \sum_{k=M}^{N_d}\binom{N_d}{k}{P_i}^k(1 - P_i)^{N_d - k} \qquad (24)$$

Here, we consider that the probability of identification is identical for all the decision branches, which is realistic since all the ANN classifiers are identical and the N_d processed signals are statistically identical.

V. Results and Discussion

The proposed algorithm was verified and validated for various orders of linear digital modulation schemes. In our simulations otherwise mentioned, we consider the following antenna configuration: $N_t = 2, N_r = 4$. First, 50 realizations of testing MIMO signals with $512 \times N_t$ symbols for each considered modulation scheme are generated. The source messages and the channel matrix are randomly selected for each realization. These realizations are employed only for ANN training. The combined HOM and HOC of the processed signals are calculated to form the features set. Then, features normalization and subset selection is performed as a preparation of ANN training. Extensive simulations show that the optimal ANN structure to be used for these algorithms is a two hidden layers network (excluding the input and the output layers), where the first layer consists of 10 nodes and the second of 15 nodes.

In what follows, we will consider two modulation pools in all our simulations $\Theta_1 = \{$ 2-PSK, 4-PSK, 8-PSK $\}$ and $\Theta_2 = \{$ 2-PSK, 4-PSK, 8-PSK, 4-ASK, 8-ASK, 16-QAM, 64-QAM $\}$. Actually, Θ_1 represents the intra-class modulation identification (i.e. only identifying the order of the modulation) and Θ_2 represents the full-class modulation identification (i.e. identifying the order and the type of the modulation at the same time).

All results are based on 1000 Monte Carlo trials for each modulation scheme i.e. 3000 Monte Carlo trials in total for Θ_1 and 7000 Monte Carlo trials in total for Θ_2. For each Monte Carlo trial, N_t random testing streams of 512 i.i.d symbols are used as input messages. Also, the channel matrix is randomly generated for each trial. The Marsaglia polar method and the combined multiple recursive one are used to generate , respectively, normal and uniform random variables.

For different values of SNR, a spatially-white circularly complex Gaussian noise with variance σ_n^2 is added such that the SNR $= 10\,log_{10}(\frac{\sigma_x^2}{\sigma_n^2})$ where σ_x^2 is the average transmitted power.

The probability of identification is given in percentage and estimated by $\frac{N_c}{N_{total}} \times 100$, where N_{total} (=3000 or 7000) is the total number of trials and N_c is given by: $N_c = \sum_{\theta_i \in \Theta_1 \text{ (or } \in \Theta_2\text{)}} N_{\theta_i}$, with N_{θ_i} being the number of trials for which the modulation $\theta_i \in \Theta_1$ (or $\in \Theta_2$) is correctly identified. In what follows, we present the performance results of our three proposed algorithms D-DMI, ZF-DMI and SCMA-DMI before introducing a performance comparison study.

A. D-DMI Performance

Fig. 2 shows the D-DMI performance in different scenarios. The SNR$_{min}$ for PSK intra-class modulation identification is 11dB while the probability of modulation identification will reach its upper bound when SNR is higher than 18dB for full-class modulation identification. Note that the channel mixing drives the received signals to be more Gaussian which corrupts the HOS and requires a higher SNR to achieve the same probability of identification compared to SISO systems. As shown in Fig. 2, the presence of channel correlation will degrade the performance and drive the SNR$_{min}$ to 15dB and 22dB, respectively, for PSK intra-class and full-class identification when considering moderately correlated MIMO channels ($|\rho_t| = |\rho_r| = 0.5$).

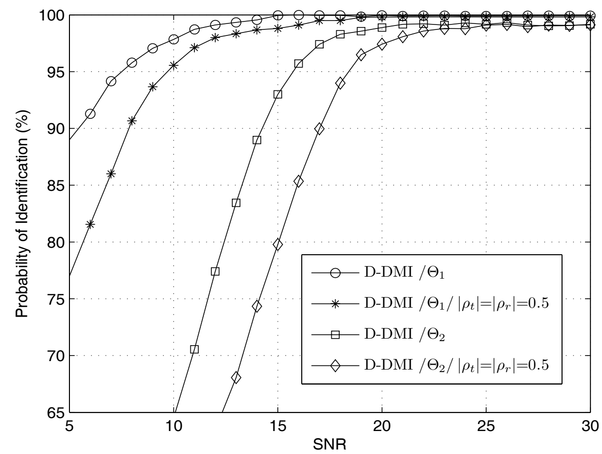

Fig. 2. Probability of identification versus SNR for D-DMI in the cases (a) Θ_1 through uncorrelated channel (b) Θ_1 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$) (c) Θ_2 through uncorrelated channel (d) Θ_2 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$).

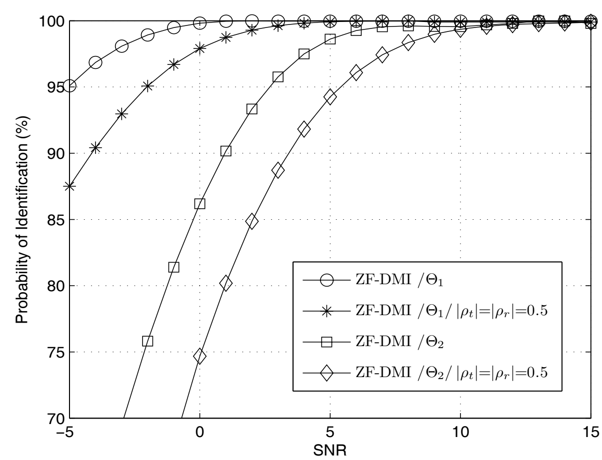

Fig. 3. Probability of identification versus SNR for ZF-DMI in the following cases (a) Θ_1 through uncorrelated channel (b) Θ_1 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$) (c) Θ_2 through uncorrelated channel (d) Θ_2 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$).

B. ZF-DMI Performance

Fig. 3 shows the ZF-DMI performance in different scenarios. The SNR$_{min}$ for PSK intra-class and full-class modulation identification is, respectively, -2dB and 6dB. It is to be noted that the ZF technique revealed some ambiguity from the received symbols which improves the identification performance. In the case of channel correlation, the performance degrades where the SNR$_{min}$ reaches 1dB and 9dB, respectively, for PSK intra-class and full-class modulation identification when $|\rho_t| = |\rho_r| = 0.5$. Also, the performance improves when

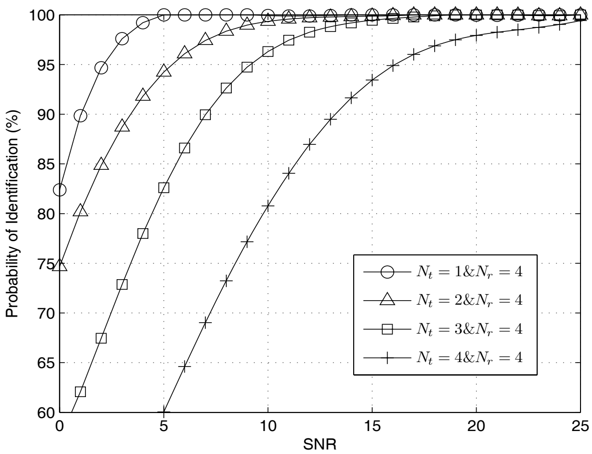

Fig. 4. Probability of identification versus SNR for different MIMO antenna configurations (ZF-DMI when considering Θ_2 with Kronecker model, $|\rho_t| = |\rho_r| = 0.5$).

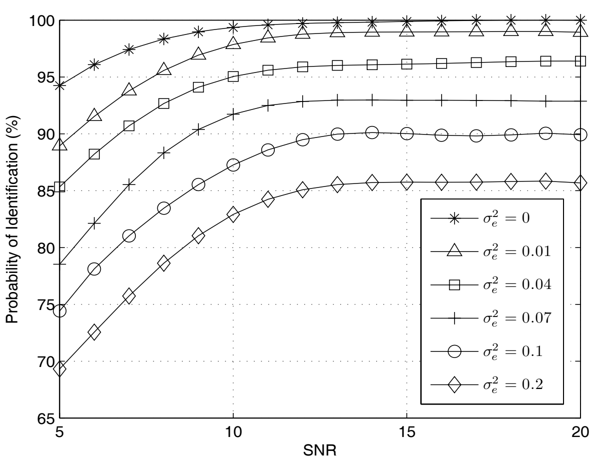

Fig. 5. Probability of identification versus SNR for different σ_e^2 (channel estimation error variance) values (ZF-DMI when considering Θ_2 with Kronecker model, $|\rho_t| = |\rho_r| = 0.5$).

the difference $\Delta = N_r - N_t$ increases as it is clear in Fig. 4. This result is expected since increasing Δ will increase the effective post-processing SNR and improve the probability of identification.

The effect of channel estimation error on modulation identification has been examined and the results are displayed in Fig. 5. As noticed, the performance will drop rapidly for an error variance $\sigma_e^2 \geq 0.1$. That is the erroneous channel estimation leads to a performance upper bound contrary to the perfect CSI case as shown in Fig. 5. This upper bound decreases as σ_e^2 increases. This upper bound problem is serious since even when the SNR is very large the identification performance will not exceed that upper bound. The proposed solution is to use a BSS technique where the channel estimation is not required.

C. SCMA-DMI Performance

Among the different BSS algorithms available in the literature we had chosen the SCMA. The performance of the blind SCMA-DMI algorithm has been examined and the results are

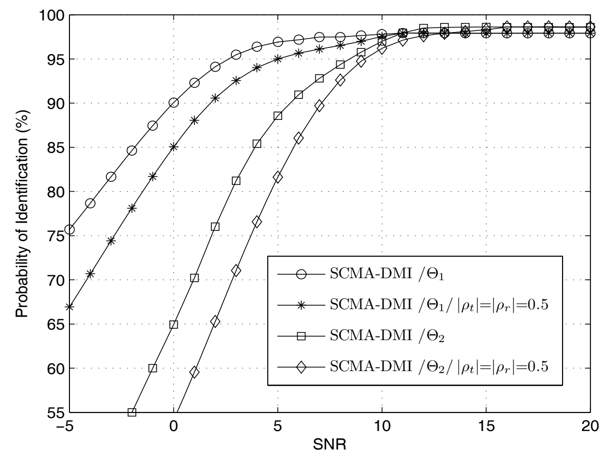

Fig. 6. Probability of identification versus SNR for SCMA-DMI in the following cases (a) Θ_1 through uncorrelated channel (c) Θ_1 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$) (b) Θ_2 through uncorrelated channel (c) Θ_2 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$) .

displayed in Fig. 6. As noticed, the SNR$_{min}$ for PSK intra-class and full-class modulation identification is, respectively, 5dB and 10dB. As shown, the presence of channel correlation will degrade the performance and drive the SNR$_{min}$ to 9dB and 13dB, respectively, for PSK intra-class and full-class identification modulation identification when $|\rho_t| = |\rho_r| = 0.5$. Obviously, the blind SCMA-DMI algorithm solves the performance upper bound problem caused by the erroneous channel estimation when employing the ZF equalizer. But the blind SCMA-DMI requires a higher SNR compared to the ZF-DMI employed when perfect CSI is assumed.

D. Channel Correlation Effect

The performance of the ZF-DMI algorithm in the presence of both transmit and receive correlations has been examined and the results are displayed in Fig. 7. As noticed, the performance degrades when the correlation increases. The SNR$_{min}$ of full-class modulation identification through uncorrelated MIMO channel is 6dB while SNR$_{min}$ is, respectively, 7dB and 9dB when $|\rho_t| = |\rho_r| = 0.3$ and $|\rho_t| = |\rho_r| = 0.5$. Also, the SNR$_{min}$ of full-class modulation identification is, respectively, 12dB and 15dB when $|\rho_t| = |\rho_r| = 0.7$ and $|\rho_t| = |\rho_r| = 0.9$. A comparison study is done to separately examine the effect of transmit and receive correlation coefficients. The results shown in Fig. 8 indicate that the transmit correlation affects the performance more than receive correlation. These results can be justified from our previous performance analysis. It is clear that when $|\rho| \to 1$, the performance degrades rapidly. Also, the degradation in SNR$_{min}$ due to increasing the correlation coefficient agrees with the SNR loss as introduced in (18).

E. Identification Performance for each Modulation Scheme

The false identification probability P_f is the probability to identify certain modulation scheme when it was not transmitted. Here, we examine P_i and P_f for each modulation scheme instead of calculating the average probability for all schemes. The modulation pool and simulations conditions are

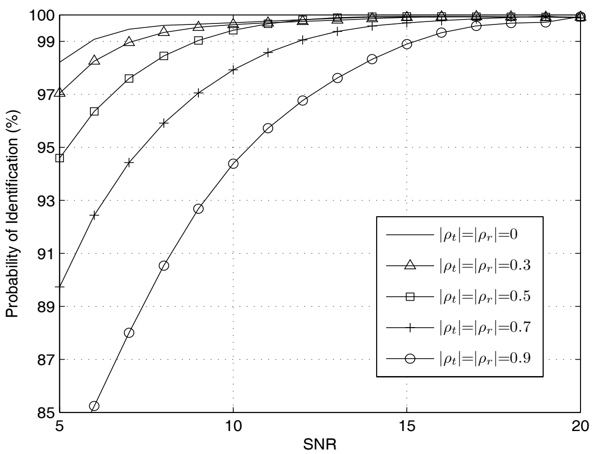

Fig. 7. ZF-DMI performance for Θ_2 with Kronecker correlation model in the following cases (a) $|\rho_t| = |\rho_r| = 0$ (b) $|\rho_t| = |\rho_r| = 0.3$ (c) $|\rho_t| = |\rho_r| = 0.5$ (d) $|\rho_t| = |\rho_r| = 0.7$ (e) and $|\rho_t| = |\rho_r| = 0.9$.

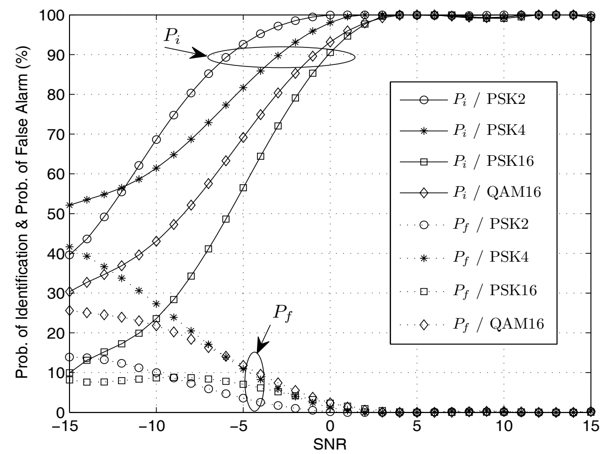

Fig. 9. ZF-DMI performance for each modulation scheme with known uncorrelated channel matrix for MIMO system using N_t = 2 and N_r = 4 antennas; 512 received samples per antenna.

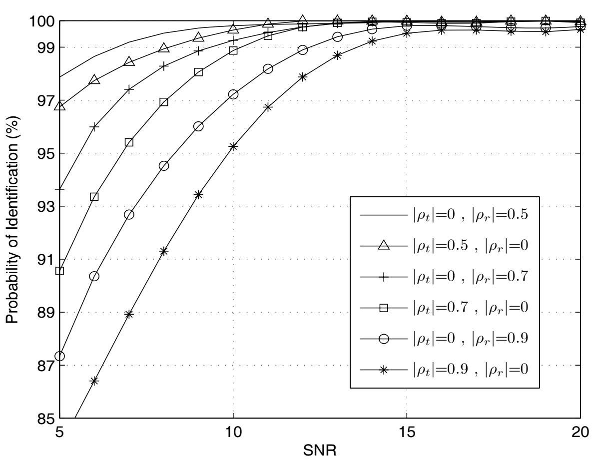

Fig. 8. Performance comparison among different values of transmit or receive correlation coefficients $|\rho_t| = 0.5, 0.7$ or 0.9 and $|\rho_r| = 0.5, 0.7$ or 0.9 (ZF-DMI performance for Θ_2 with Kronecker correlation model).

the same as that used in [8] to evaluate the performance of the ALRT and HLRT algorithms. We consider an uncorrelated MIMO system with 2×4 antennas and 512 received symbol per antenna. The considered modulation pool is: 2-PSK, 4-PSK,16-PSK and 16-QAM.

Fig. 9 shows P_i and P_f for each scheme when employing the ZF-DMI algorithm. The identification performance is the best for 2-PSK when the SNR is relatively high. On the other hand 4-PSK is the best detected and the more false identified in the low SNR region. The ZF-DMI algorithm performs well when compared to the optimal ALRT algorithm. The SCMA-DMI algorithm is used when the channel matrix is unknown. It is clear in Fig. 10 that, when SNR is relatively low, 2-PSK scheme has higher P_i and P_f relative to 16-PSK scheme which is the worst identified and the less false identified. Also as seen, the performance of SCMA-DMI is good in comparison with the HLRT algorithm. The major drawbacks of the methods in [8] are the high computational complexity and its need of perfect knowledge of the noise variance at the receiver side.

F. Performance Comparison

Finally, a comparison study among the three proposed algorithms of modulation identification introduced in this paper is done and the results are displayed in Fig. 11. As noticed, the ZF-DMI algorithm offers the best performance when perfect CSI is assumed. However, the proposed ZF-DMI algorithm is sensitive to channel estimation errors. Note that the presence of erroneous channel estimation causes a rapid performance degradation when the error variance $\sigma_e^2 \geq 0.1$. Also, this erroneous estimation leads to an upper bound of the probability of identification contrary to the perfect CSI case; i.e. SNR tends to infinity but the probability of identification does not reach 100%. To solve this problem the totally blind SCMA-DMI algorithm was proposed. This algorithm solved the upper bound problem without any CSI knowledge. However, the SCMA-DMI requires higher SNR to achieve the same performance as the ZF-DMI employed when perfect CSI is assumed. In fact the SNR$_{min}$ is, respectively, 9dB and 13dB when using ZF-DMI and SCMA-DMI algorithms for full-class identification through MIMO channels when $|\rho_t| = |\rho_r| = 0.5$. The D-DMI algorithm has the lower complexity but offers a low performance compared to the remaining two algorithms. Simulation results show a gain in SNR$_{min}$ of about 4dB when comparing the performance for SISO and uncorrelated SIMO ($N_r = 4$) systems.

VI. Conclusions

We presented three algorithms for digital modulation identification aimed for correlated MIMO systems based on HOS as features extraction subsystem and a neural network trained with resilient backpropagation learning algorithm as classifier subsystem. The less complex D-DMI algorithm extracted the features directly from the received MIMO signal and it is a good choice when the transmitted power is high. The ZF-DMI employs a zero-forcing equalizer before the features extraction process and it is the best algorithm when perfect CSI knowledge is assumed. The SCMA-DMI algorithm blindly separates the symbols and offers the best performance when

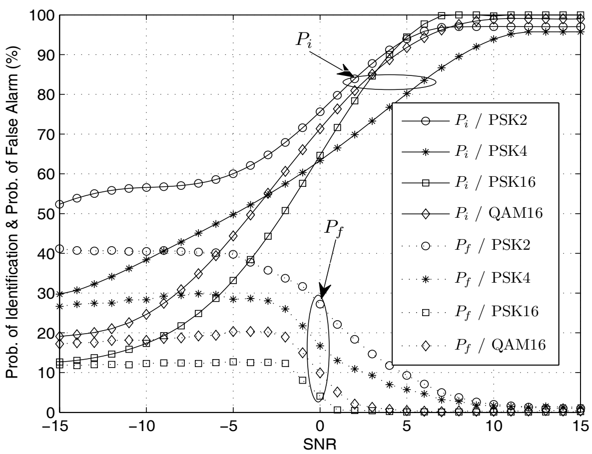

Fig. 10. SCMA-DMI performance for each modulation scheme with unknown uncorrelated channel matrix for MIMO system using N_t = 2 and N_r = 4 antennas; 512 received samples per antenna.

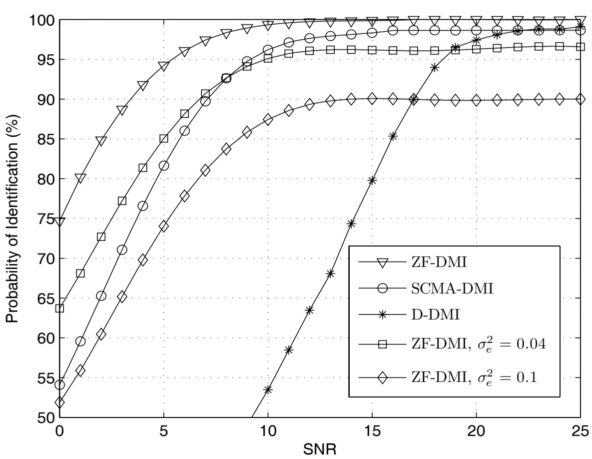

Fig. 11. Performance comparison among different algorithms for Θ_2 with Kronecker model ($|\rho_t| = |\rho_r| = 0.5$).

no CSI knowledge is assumed. The proposed algorithms are examined through correlated MIMO channels and they are shown to be capable of identifying different linear digital modulation schemes with high accuracy. The robustness of these algorithms to synchronization errors, frequency offsets and phase jitter must be studied.

ACKNOWLEDGMENT

This work was performed in the framework of the I-Trans cluster and the regional CISIT (Campus International Sécurité et Intermodalité des Transports) project. The authors would like to thank the North Region and the ERDF (European Regional Development Fund) for financial support.

REFERENCES

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Commun.*, vol. 1, pp. 137–156, 2007.

[2] J. L. Xu, W. Su, and M. C. Zhou, "Likelihood ratio approaches to automatic modulation classification," *IEEE Trans. Systems, Man, and Cybernetics: Part C*, vol. 41, pp. 455–469, July 2011.

[3] M. L. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," *Signal Process.*, vol. 84, no. 2, pp. 351–365, Feb. 2004.

[4] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3098–3105, Aug. 2008.

[5] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Cyclostationarity-based modulation classification of linear digital modulations in flat fading channels," *Wireless Personal Commun. J.*, vol. 54, no. 4, pp. 699–717, Sep. 2010.

[6] P. Prakasam and M. Madheswaran, "Modulation identification algorithm for adaptive demodulator in software defined radios using wavelet transform," *Int'l J. Signal Process.*, vol. 5, pp. 74–81, Jan. 2009.

[7] A. Swami and B. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[8] V. Choqueuse, S. Azou, K. Yao, L. Collin, and G. Burel, "Blind modulation recognition for MIMO systems," *J. ATM Bucharest*, vol. XIX, no. 2, pp. 183–196, June 2009.

[9] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[10] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[11] A. van Zelst and J. Hammerschmidt, "A single coefficient spatial correlation model for MIMO radio channels," in *2002 General Assembly of the International Union of Radio Science*.

[12] A. Ikhlef and D. L. Guennec, "A simplified constant modulus algorithm for blind recovery of MIMO QAM and PSK signals: a criterion with convergence analysis," *EURASIP J. Wireless Commun. and Networking*, article ID 90401, 2007.

[13] C. Papadias, "Globally convergent blind source separation based on a multiuser Kurtosis maximization criterion," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3508–3519, Dec. 2000.

[14] L. Wang, R. C. de Lamare, and Y. L. Cai, "Low-complexity adaptive step size constrained constant modulus SG algorithms for adaptive beamforming," *Signal Process.*, vol. 89, pp. 2503–2513, Dec. 2009.

[15] S. Douglas, "Combined subspace tracking, prewhitening, and contrast optimization for noisy blind signal separation," in *2000 Workshop Indept. Component Analysis Source Separation*, pp. 579–584.

[16] O. Somekh, O. Simeone, Y. Bar-Ness, and W. Su, "Detecting the number of transmit antennas with unauthorized or cognitive receivers in MIMO systems," in *2007 IEEE MILCOM*.

[17] V. Choqueuse, M. Marazin, L. Collin, K. Yao, and G. Burel, "Blind recognition of linear space-time block codes: a likelihood-based approach," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1290–1299, Mar. 2010.

[18] P. McCullagh, *Tensor Methods in Statistics*. Chapman Hall, 1987.

[19] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *Proc. 1993 IEEE International Conference on Neural Networks*.

[20] D. A. Gore, R. W. Heath Jr., and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.

[21] R. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley, 2005.

[22] C. Wang, E. K. S. Au, R. D. Murch, W. H. Mow, R. S. Cheng, and V. Lau, "On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 805–810, Mar. 2007.

Kais Hassan was born in 1975. He received the M.Sc. degree in telecommunications systems from University of Valenciennes, Valenciennes, France, in 2009. He is currently working towards the Ph.D. degree in telecommunications.

From 2000 to 2008, he was a research engineer at the Higher Institute for Applied Science and Technology (HIAST), Syria. He is now a member of the Institute of Electronics, Microelectronics and Nanotechnology (IEMN), Valenciennes, France, and the French Institute of Sciences and Technology for Transport, Development and Networks (IFSTTAR), Lille, France. His research interests focus on wireless communications, blind signal interception, and cognitive radio. He also teaches digital communications, electronics, and signal processing at the University of Valenciennes.

Iyad Dayoub received the M.A.Sc. in electrical engineering from the University of Nancy (Polytechnic Institute), Nancy, France, in 1997, and the Ph.D. Degree in telecommunications from the University of Valenciennes, France, in 2001. In Sept. 2003, he joined the Department of Opto Acoustic and Electronics at the Institute of Electronics, Microelectronics and Nanotechnology (IEMN/DOAE), University of Valenciennes, France, where he is currently Associate Professor. Since 2007, he has been member of the National Council of Universities, France, in the area of electrical engineering, electronics, photonics, and systems. His research activities are developed with the Digital Communications Group of the IEMN, including wireless communications, emerging wireless access technologies, signal detection, hybrid radio optic networks and optical MIMO. Dr. Dayoub is currently Adjunct Professor at the Department of Electrical and Computer Engineering of Concordia University, Montreal, Canada.

Walaa Hamouda (SM'06) received his M.A.Sc. and Ph.D. degrees from Queen's University, Kingston, Ontario, Canada, in 1998 and 2002, respectively, both in Electrical and Computer Engineering. In July 2002, he joined the Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada, where he is now an Associate Professor. From June 2006, he has been appointed as a Concordia University Research Chair in Communications and Networks. He served as the technical co-chair of the Wireless Networks Symposium, IEEE Globecom 2012, the Ad-hoc-Sensor and Mesh Networking Symposium of the ICC 2010, the 25th Queen's Biennial Symposium on Communications, track co-chair of the Transmission Techniques, VTC-Fall 2012, Quebec, Canada, and track chair of the Radio Access Techniques of the IEEE VTC-Fall 2006 Montreal, Canada. From Sept. 2005 to Nov. 2008, has been the Chair of the IEEE Montreal chapter in Communications and Information Theory. He received many awards including the best paper award [WNS] of the ICC 2009, and the IEEE Canada Certificate of Appreciation in 2007, 2008. He serves as Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE COMMUNICATIONS LETTERS, and the *IET Wireless Sensor Systems*. His current research interests are in wireless networks, MIMO space-time processing, multiuser communications, cross-layer design, source and channel coding.

Crépin Nsiala Nzéza received the M.Sc. degree in 2002 and the Ph.D. degree in 2006 both in electrical engineering from the University of Brest, Brest, France. From 2006 to October 2008, he has been with the Centre National de la Recherche Scientifique (CNRS) in the Analog/RF IC Design Group at the Institute of Electronics, Microelectronics and Nanotechnology (IEMN) in Lille, France, as a Postdoctoral Fellow. From October 2008 to February 2011, he has been with the Institute of Sciences and Technology for Transport, Development and Networks (IFSTTAR), Lille, France, as Research Fellow. Since March 2011, he is with the Research and Development Department of the Segula Matra technologie Group as Senior Researcher in embedded electronic systems in automotive. His research interests lie in the areas of digital communications, signal processing, blind CR approaches for both C2C and C2I communications, blind signal interception, reconfigurable IC design, and software development for Telematics and infotainment systems.

Marion Berbineau was born in Toulouse, France, on September 18, 1962. She received the Engineer degree in electronics, automatic and metrology from Polytech'Lille (France) and the PhD in electronics from the University of Lille respectively in 1986 and 1989. She joined IFSTTAR (ex INRETS) as a full time researcher in telecommunications in 1989. She is currently Research Director and Director of the LEOST laboratory. Dr. Berbineau field of expertise are EM propagation and modeling, channel characterization and modeling for transport and complexes environments, particularly in tunnels, signal processing for wireless communication systems in multipath environments, MIMO systems and wireless systems for ITS and for railway applications including control-command, video surveillance, passenger information. She is involved in several national and European projects. She is author and co-author of several publications and patents. Dr Berbineau is an IEEE member, affiliated to the VTS society.