

Article

Machine Learning for 5G MIMO Modulation Detection

Haithem Ben Chikha ^{1,*} , Ahmad Almadhor ¹  and Waqas Khalid ² 

¹ Computer Engineering and Networks Department, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia; aaalmadhor@ju.edu.sa

² Institute of Industrial Technology, Korea University, Sejong 30019, Korea; waqas283@korea.ac.kr

* Correspondence: hbchikha@ju.edu.sa

Abstract: Modulation detection techniques have received much attention in recent years due to their importance in the military and commercial applications, such as software-defined radio and cognitive radios. Most of the existing modulation detection algorithms address the detection dedicated to the non-cooperative systems only. In this work, we propose the detection of modulations in the multi-relay cooperative multiple-input multiple-output (MIMO) systems for 5G communications in the presence of spatially correlated channels and imperfect channel state information (CSI). At the destination node, we extract the higher-order statistics of the received signals as the discriminating features. After applying the principal component analysis technique, we carry out a comparative study between the random committee and the AdaBoost machine learning techniques (MLTs) at low signal-to-noise ratio. The efficiency metrics, including the true positive rate, false positive rate, precision, recall, F-Measure, and the time taken to build the model, are used for the performance comparison. The simulation results show that the use of the random committee MLT, compared to the AdaBoost MLT, provides gain in terms of both the modulation detection and complexity.

Keywords: 5G; multi-relay cooperative MIMO systems; modulation detection; random committee machine learning technique



Citation: Ben Chikha, H.; Almadhor, A.; Khalid, W. Machine Learning for 5G MIMO Modulation Detection.

Sensors **2021**, *21*, 1556.

<http://doi.org/10.3390/s21051556>

Academic Editor: Ahmet Kondo

Received: 12 January 2021

Accepted: 20 February 2021

Published: 24 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recently, the integration of 5G new radio (NR) and multiple-input multiple-output (MIMO) has received increasing attention due to its effectiveness in improving both the capacity and robustness of the wireless systems [1]. In fact, the use of multiple antenna elements in MIMO systems is considered as one of the most promising technologies in 5G NR systems that can be employed to enable beamforming and spatial multiplexing [2]. The cooperative MIMO systems also offer a considerable rate gain and improve the diversity order [3–7]. The efficient relaying mechanisms are taken into account by the standard specifications of the mobile broadband communication systems, such as LTE-advanced (LTE-A) [8]. Furthermore, the estimation of communication parameters (e.g., number of antennas, coding, and modulation) has received a great deal of attention. It has found applications in several military and civilian communication systems, such as software-defined radio and cognitive radios [9,10]. It is important to consider the effect of spatial correlation and imperfect channel state information (CSI) in the cooperative MIMO systems. The distance limitation between the antennas and scatterers existing in propagation environment affects the diversity, multiplexing, and capacity gains. The errors caused by the channel estimation, quantization, reciprocity mismatch, and delay produce the imperfect CSI [11]. It is practically unfeasible to obtain the perfect CSI at all nodes [12].

1.1. Related Works

In a MIMO destination node, the decoder or the spatial demultiplexer and the demodulator are employed to recover the transmitted binary information. In fact, the destination node is the entity that converts the received waves into a binary stream. Therefore, the

estimation of the transmitted binary information requires a prior knowledge of the communication parameters including the number of the source antennas, the noise variance, the coding, the channel matrix, and the modulation. To design powerful cognitive radio, many estimation algorithms of the communication parameters have been developed in the literature. In [13–15], the authors have proposed algorithms for estimating the number of the source antennas. Furthermore, several approaches have been proposed for the detection of the coding [16,17] and other algorithms for estimating the channel matrix are also available [18–20]. In the last few decades, there has been a great interest on modulation detection for MIMO communications systems [21–26]. In fact, the misdetection of the modulation type causes a considerable performance degradation in the estimation of source information as the modulation estimation is done prior to the demodulation phase. The current work is focusing on the modulation detection for cooperative MIMO communication systems. In order to alleviate the aforementioned concerns, the use of artificial intelligence (AI) to detect modulations, based on MLTs, is employed in existing literature as it is simple to implement and achieves a quasi-optimal performance when choosing an appropriate features set. Based on a MLT, the authors of [24] have investigated the modulation detection for the MIMO relaying system under the conditions of perfect CSI and uncorrelated channels. In [23], Lau et al. proposed an algorithm dedicated for the short-range devices that use the European 868 MHz band under the same conditions, i.e., perfect CSI and uncorrelated channels. Under the same conditions, the detection of superposed modulations for two-way relaying MIMO systems with the physical-layer network coding is studied in [22]. While there have been considerable works that assumed the conditions of the perfect CSI and uncorrelated channels, little attention has been paid to the cooperative MIMO systems with imperfect CSI and under correlated channels. In [21], we proposed an algorithm for the modulation detection dedicated to cooperative MIMO systems over the imperfectly estimated correlated channels. In this work, we consider the AdaBoost MLT; it provides good performance of the modulation detection at an acceptable signal-to-noise ratio (SNR). However, it suffers from large training time and convergence speed that can affect the design of practical modulation detection algorithm dedicated to multi-relay cooperative MIMO systems.

1.2. Contributions

In this study, we aim to propose a detection modulation algorithm with a low complexity that is dedicated to multi-relay cooperative MIMO systems over the imperfectly estimated correlated channels. More specifically, the contributions of this work are summarized as follows.

- We proposed a modulation detection using random committee MLT for multi-relay cooperative MIMO systems over the imperfectly estimated correlated channels. To the best of the authors knowledge, this is the first time that random committee MLT is used for the modulation detection. The purpose is to detect the modulation types and orders among different M-ary shift-keying linear modulations (M-PSK and M-QAM) used by broadband technologies, especially 5G NR.
- The modulation detection algorithms proposed in [21,24] are considered as benchmarks, and the comparative study is provided. The performance of the proposed modulation detection algorithm is investigated and evaluated with the number of efficiency metrics, such as the true positive rate, false positive rate, precision, recall, F-Measure, and the time taken. The superiority of the proposed modulation detection algorithm in terms of computational complexity and modulation detection is verified through the simulation results.

1.3. Outline

The rest of the paper is organized as follows. Section 2 presents the system model, including the system description and assumptions. The proposed modulation detection algorithm is provided in Section 3. The simulation results, along with the discussion and

the benefits of the proposed algorithm, are given in Section 4. Finally, the conclusions of the paper are presented in Section 5.

1.4. Notations

In this paper, we use $\text{tr}(\cdot)$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^T$, and $(\cdot)^{-1}$ to denote the trace, conjugate, conjugate transpose, transpose, and inverse, respectively. $(\cdot)_{r,c}$ represents the entry in the r^{th} row and the c^{th} column of a matrix. $\mathbb{E}[\cdot]$ denotes the statistical expectation. \mathbf{I}_N stands for a $N \times N$ identity matrix. The set of $M \times N$ matrices over complex field is denoted by $\mathbb{C}^{M \times N}$. Finally, $\mathcal{CN}(m, \Sigma)$ is a circularly symmetric complex Gaussian distribution with mean m and covariance matrix Σ . It is noted that the abbreviations used in this paper are listed in Table 1.

Table 1. A list of abbreviations.

Abbreviation	Description
AdaBoost	Adaptive Boosting
ANN	Artificial neural networks
BM	Beamforming matrix
CSI	Channel state information
DB	Learning database
FP	False positive
HOCs	Higher order cumulants
HOMs	Higher order moments
HOSs	Higher order statistics
i.i.d.	Independent, identically distributed
LTE	Long term evolution
LTE-A	LTE-advanced
MLP	Multilayer perceptron
MLT	Machine learning technique
MIMO	Multiple-input multiple-output
RPROP	resilient backpropagation
RZF	Regularized zero forcing
SNR	Signal-to-noise ratio
SM	Spatial multiplexing
TP	True positive
ZF-RZF	Zero forcing and regularized zero forcing

2. System Model and Assumptions

2.1. System Model

We consider the multi-relay cooperative MIMO system over the spatially-correlated channel as shown in Figure 1. Here, we denote the source node by S , the destination node by D , and L relay nodes by $R_l, l = 1, 2, \dots, L$ to make information transmission from S to D . We suppose that antennas N_{AS} , N_{AR} , and N_{AD} are enabled at S , each R_l , and D , respectively. We apply a two-phase transmission protocol to allow the information transmission from S to D ; via the direct link SD and the cooperative links $SR_l - R_lD, l = 1, \dots, L$. To this end, a non-regenerative and half-duplex relay technique is employed for processing and forwarding the received signals at each R_l [27]. Now, if we apply a spatial multiplexing (SM) at S and if all R_l and D are to simultaneously support all the N_{AS} independent substreams, then we must meet the requirements

$$\begin{cases} N_{AD} \geq N_{AS}, \\ N_{AR} \geq N_{AS}. \end{cases} \quad (1)$$

We presume that all nodes have an equal number of antennas for simplicity, i.e., $N_{AS} = N_{AR} = N_{AD} = N_A$.

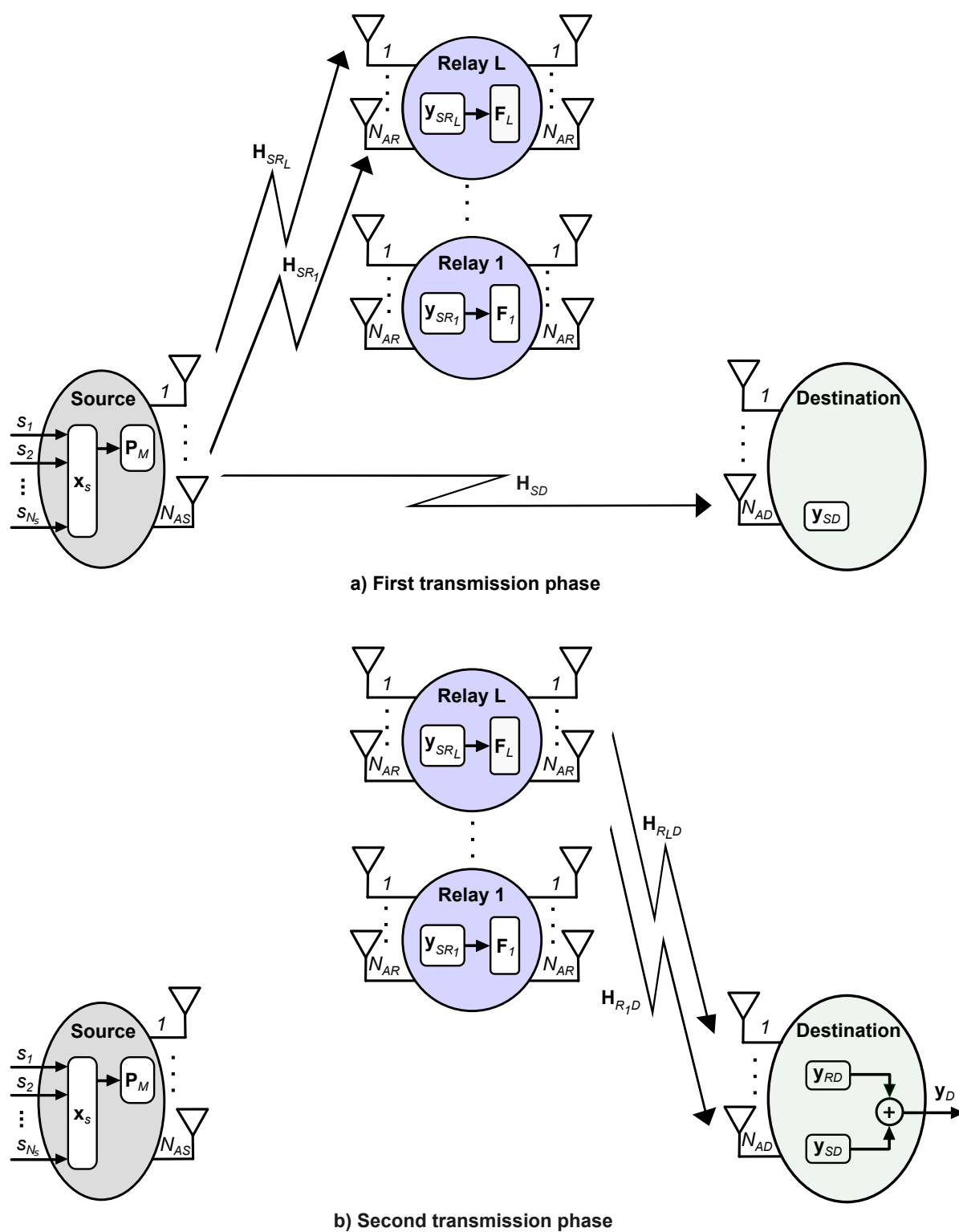


Figure 1. Multi-relay cooperative MIMO systems using one source, one destination and L relays.

We encode the source signals \mathbf{x}_s in the first transmission phase using SM. Therefore, the signals \mathbf{x}_s is given by

$$\mathbf{x}_s = [s_1, s_2, \dots, s_{N_A}]^T, \quad (2)$$

where s_1, s_2, \dots, s_{N_A} are assumed to be independent, identically distributed (i.i.d.), and mutually independent. Given the fact that we stand the transmit power by P_{x_s} at S , \mathbf{x}_s should satisfy the power constraint given by

$$\mathbb{E}[\mathbf{x}_s \mathbf{x}_s^H] = P_{x_s}(N_A)^{-1} \mathbf{I}_{N_A}. \quad (3)$$

To achieve a near-capacity at sum-rate [28], we apply a regularized zero forcing (RZF) linear precoding technique at S . Consequently, the linear precoding matrix can be expressed as

$$\mathbf{P}_M = \hat{\mathbf{H}}_{SD}^H (\hat{\mathbf{H}}_{SD} \hat{\mathbf{H}}_{SD}^H + \alpha_1 \mathbf{I}_{N_A})^{-1}, \quad (4)$$

where $\hat{\mathbf{H}}_{SD} \in \mathbb{C}^{N_A \times N_A}$ is the matrix that estimate the transmission channel from S to D with a Gaussian distributed error, and α_1 represents the total noise variance to the total transmit power ratio expressed by [28]

$$\alpha_1 = N_A \sigma_{SD}^2 / P_{x_s}. \quad (5)$$

After performing the RZF to \mathbf{x}_s , S can transmit the precoded data in parallel to D and all R_l . At D , the received signal can be expressed as

$$\mathbf{y}_{SD} = \rho_s \mathbf{H}_{SD} \mathbf{P}_M \mathbf{x}_s + \mathbf{n}_{SD}, \quad (6)$$

where ρ_s represents the control factor of the source power, given by

$$\rho_s = \sqrt{P_{x_s} / \text{tr}(\mathbf{P}_M \mathbf{P}_M^H)}, \quad (7)$$

$\mathbf{H}_{SD} \in \mathbb{C}^{N_A \times N_A}$ is the SD channel matrix with spatial correlation and $\mathbf{n}_{SD} \sim \mathcal{CN}(0, \sigma_{SD}^2 \mathbf{I}_{N_A})$ represents an additive zero-mean spatially-white circularly complex Gaussian noise with variance σ_{SD}^2 . At $R_l, l = 1, \dots, L$, the received signal is given by

$$\mathbf{y}_{SR_l} = \rho_s \mathbf{H}_{SR_l} \mathbf{P}_M \mathbf{x}_s + \mathbf{n}_{SR_l}, \quad (8)$$

where $\mathbf{H}_{SR_l} \in \mathbb{C}^{N_A \times N_A}$ is the SR_l channel matrix with spatial correlation and $\mathbf{n}_{SR_l} \sim \mathcal{CN}(0, \sigma_{SR_l}^2 \mathbf{I}_{N_A})$.

In the second transmission phase, all R_l apply a linear beamforming matrix (BM) to received signals from the first transmission phase. We model the linear BM according to the zero forcing and regularized zero forcing (ZF-RZF) [29], denoted by \mathbf{F}_l and given by

$$\mathbf{F}_l = \hat{\mathbf{H}}_{R_l D}^H (\hat{\mathbf{H}}_{R_l D} \hat{\mathbf{H}}_{R_l D}^H + \alpha_2 \mathbf{I}_{N_A})^{-1} \times \left((\mathbf{H}_{SR_l} \mathbf{P}_M)^H \mathbf{H}_{SR_l} \mathbf{P}_M \right)^{-1} (\mathbf{H}_{SR_l} \mathbf{P}_M)^H, \quad (9)$$

where $\hat{\mathbf{H}}_{R_l D} \in \mathbb{C}^{N_A \times N_A}$ is the matrix that estimate the transmission channel from R_l to D with a Gaussian distributed error at l th relay node, and α_2 represents the total noise variance to the total transmit power ratio expressed by [28]

$$\alpha_2 = N_A \sigma_{RD}^2 / P_r. \quad (10)$$

After that, the resulting signals after ZF-RZF precoding are forwarded from all R_l , and the received signal at D can be written as

$$\mathbf{y}_{RD} = \sum_{l=1}^L \rho_{r_l} \rho_s \mathbf{H}_{R_l D} \mathbf{F}_l \mathbf{H}_{SR_l} \mathbf{P}_M \mathbf{x}_s + \sum_{l=1}^L \rho_{r_l} \mathbf{H}_{R_l D} \mathbf{F}_l \mathbf{n}_{SR_l} + \mathbf{n}_{RD}, \quad (11)$$

where ρ_{r_l} represents the control factor of the l th relay power, given by

$$\rho_{r_l} = \sqrt{P_r / \text{tr}(\rho_s^2 \mathbf{F}_l \mathbf{H}_{SR_l} \mathbf{P}_M \mathbf{P}_M^H \mathbf{H}_{SR_l}^H \mathbf{F}_l^H + \sigma_{RD}^2 \mathbf{F}_l \mathbf{F}_l^H)} \quad (12)$$

and $\mathbf{H}_{R_l D} \in \mathbb{C}^{N_A \times N_A}$ is the $R_l D$ channel matrix with spatial correlation and $\mathbf{n}_{R_l D} \sim \mathcal{CN}(0, \sigma_{R_l D}^2 \mathbf{I}_{N_A})$.

2.2. Spatial Correlation Model

The distance limitation between the antennas and scatterers existing in the propagation environment can affect the diversity, multiplexing, and capacity gains [30]. We propose to model spatial correlation for cooperative MIMO channels based on the Kronecker model [30]. Accordingly, the channel correlation matrices \mathbf{H}_{SD} , \mathbf{H}_{SR_l} , and $\mathbf{H}_{R_l D}$ can be expressed as

$$\begin{cases} \mathbf{H}_{SD} = \mathbf{R}_{H_{SD}, RX}^{1/2} \mathbf{H}_{w_{SD}} \mathbf{R}_{H_{SD}, TX}^{1/2} \\ \mathbf{H}_{SR_l} = \mathbf{R}_{H_{SR_l}, RX}^{1/2} \mathbf{H}_{w_{SR_l}} \mathbf{R}_{H_{SR_l}, TX}^{1/2} \\ \mathbf{H}_{R_l D} = \mathbf{R}_{H_{R_l D}, RX}^{1/2} \mathbf{H}_{w_{R_l D}} \mathbf{R}_{H_{R_l D}, TX}^{1/2} \end{cases} \quad (13)$$

where $\mathbf{H}_{w_{SD}}$, $\mathbf{H}_{w_{SR_l}}$, and $\mathbf{H}_{w_{R_l D}}$ are full rank gain matrices of which the entries are i.i.d. and follow a circularly symmetric complex Gaussian distribution with zero-mean and unit variance. $\mathbf{R}_{H_{SD}, RX}$, $\mathbf{R}_{H_{SR_l}, RX}$ and $\mathbf{R}_{H_{R_l D}, RX}$ are the receiver correlation matrices. Finally, $\mathbf{R}_{H_{SD}, TX}$, $\mathbf{R}_{H_{SR_l}, TX}$ and $\mathbf{R}_{H_{R_l D}, TX}$ are the transmitter correlation matrices. Based on the exponential correlation model defined in [31,32], we model the entries of the receiver and transmitter correlation matrices presented in (13). In fact, for a correlation matrix, denoted \mathbf{R} , the entries can be expressed as

$$(\mathbf{R})_{rc} = \begin{cases} \rho^{c-r}, & r \leq c \\ \frac{\rho^{c-r}}{(\mathbf{R})_{cr}}, & r > c \end{cases}, \quad |\rho| < 1, \quad (14)$$

where ρ denotes the amount of correlation.

2.3. Imperfect Channel Estimation Model

In practical cooperative MIMO systems, it is unfeasible to obtain perfect CSI at all nodes. In fact, errors can be produced by channel estimation, quantization, reciprocity mismatch, and delay. Consequently, it results in the presence of imperfect CSI. In this paper, we consider that the estimation of the backward \mathbf{H}_{SR_l} channels can be performed based on pilot signaling by R_l , $l = 1, \dots, L$ and thus relay nodes have a perfect knowledge of \mathbf{H}_{SR_l} [33]. However, we consider the existence of the imperfect CSI in SD and $R_l D$, $l = 1, \dots, L$ links. Therefore, we model the imperfect CSI for SD and $R_l D$, $l = 1, \dots, L$ links as [11]

$$\begin{cases} \hat{\mathbf{H}}_{SD} = \mathbf{H}_{SD} + e_{SD} \mathbf{\Omega}_{SD}, \\ \hat{\mathbf{H}}_{R_l D} = \mathbf{H}_{R_l D} + e_{R_l D} \mathbf{\Omega}_{R_l D}. \end{cases} \quad (15)$$

where the entries of $\mathbf{\Omega}_{SD}$ and $\mathbf{\Omega}_{R_l D}$ are i.i.d. zero-mean circularly symmetric complex Gaussian variables with unit variance. In addition, these matrices, i.e., $\mathbf{\Omega}_{SD}$ and $\mathbf{\Omega}_{R_l D}$, are independent of \mathbf{H}_{SD} and $\mathbf{H}_{R_l D}$, respectively. e_{SD}^2 and $e_{R_l D}^2$ denote the estimation error variances of the SD and $R_l D$ channels, respectively. At the end of the second transmission phase, two copies of source data \mathbf{x}_s are received at D through the direct link SD , i.e., \mathbf{y}_{SD} (Equation (6)), and the cooperative links $SR_l - R_l D$, i.e., \mathbf{y}_{RD} (Equation (11)). These two copies are combined in order to increase the SNR. Therefore, the received signal at D without any time oversampling and optimum symbol timing and with perfect carrier frequency and phase estimation is given by

$$\mathbf{y}_D = \mathbf{y}_{SD} + \mathbf{y}_{RD}. \quad (16)$$

After that, discriminating features will be extracted from received signals as an input to the random committee MLT [34].

3. Modulation Detection Algorithm

3.1. Features Extraction

To correctly estimate the modulation from a received signal, an appropriate choice of key features is mandatory. The higher-order statistics (HOSs) that include the higher-order moments (HOMs) and higher-order cumulants (HOCs) are considered as promising features allowing to offer a good detection of modulation types [35,36]. For that reason, we choose the HOMs and HOCs statistics up to order eight for modulation detection purposes [36].

The m th-order HOM of a received signal vector at the a th antenna, denoted by $\mathbf{y}_D^{(a)} = (y_{D,1}^{(a)}, \dots, y_{D,N}^{(a)})$ can be written as [37]

$$M_{mk}(\mathbf{y}_D^{(a)}) = \mathbb{E} \left[\left(\mathbf{y}_D^{(a)} \right)^{m-k} \left(\overline{\mathbf{y}_D^{(a)}} \right)^k \right], \quad a = 1, \dots, N_{AD}. \quad (17)$$

The HOMs can be also expressed as

$$\hat{M}_{mk}(\mathbf{y}_D^{(a)}) = \frac{1}{N} \sum_{n=1}^N \left(y_{D,n}^{(a)} \right)^{m-k} \left(\overline{y_{D,n}^{(a)}} \right)^k. \quad (18)$$

The m th-order HOC of $\mathbf{y}_D^{(a)}$ is given by

$$C_{mk}(\mathbf{y}_D^{(a)}) = \text{Cum} \left[\underbrace{\mathbf{y}_D^{(a)}, \dots, \mathbf{y}_D^{(a)}}_{(m-k) \text{ times}}, \underbrace{\mathbf{y}_D^{(a)}, \dots, \mathbf{y}_D^{(a)}}_{(k) \text{ times}} \right]. \quad (19)$$

where m th-order HOC can be written based on equal and lower ordered HOMs as [37]

$$\text{Cum} [\mathbf{y}_{D1}^{(a)}, \dots, \mathbf{y}_{Dm}^{(a)}] = \sum_{\Psi} (-1)^{\beta-1} (\beta-1)! \prod_{\varphi \in \Psi} E \left[\prod_{c \in \varphi} \mathbf{y}_{Dc}^{(a)} \right], \quad (20)$$

where Ψ runs through the list of all partitions of $1, \dots, j$, φ runs through the list of all blocks of the partition Ψ and β is the elements number of the partition Ψ . The interested reader can refer to the Appendix A for further details. We raise each HOC to the power $2/m$ as the magnitude of HOCs increases with their order [38].

To improve the performance of the proposed algorithm in terms of modulation detection and decrease the computational cost, a reduced set of features is chosen based on the principal component analysis (PCA) technique [39]. In fact, this latter technique allows building a low-dimensional representation of the extracted features that describes as much as possible the variance in that features. It represents a linear transformation that transforms the components of the extracted features to orthogonal components. Thereafter, it ranks the resulting orthogonal components in a manner that those with the largest variation are placed in the top of the list. Consequently, the selected subset of features is the orthogonal components with the largest variance, while the remaining components are those that present high correlations and thus can be removed with a minimal loss of information. Simulations show that only ten orthogonal components, i.e., $N_{feat} = 10$, among twenty-eight, are chosen in the training and test phases.

To detect the modulation type of an unknown signal, denoted by \mathbf{y}_D , a training phase should be launched. It involves building a classifier from a learning database (DB). Based on the built classifier, the test phase is done to detect the modulation type. In Figure 2, we present the modulation detection of a given \mathbf{y}_D signal.

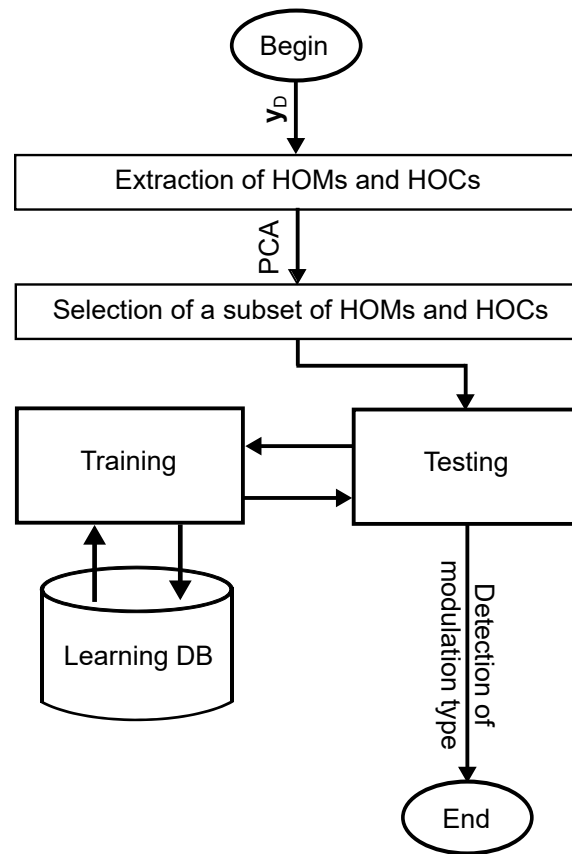


Figure 2. Modulation detection of a given y_D signal.

3.2. Random Committee Operating with Random Tree MLT

To detect the modulation type used by the source node based on the received signals at the destination node, we deploy the random committee MLT [34] using the random tree as a base MLT [40]. Let T represents the number of the training subsets (i.e., T_1, \dots, T_T). These latter construct an ensemble of C_1, \dots, C_T random tree classifiers, where each random tree classifier is formed based on a various random number seed using the same training data. With the use of the random tree classifier, a set of features is randomly selected in each node to construct the classifier and the final detection decision, denoted by D_{final} , is an average of the received predictions given by the individual random tree classifiers.

3.3. Adaptive Boosting (AdaBoost) Operating with Decision Tree

In this work, we compare the random committee with the AdaBoost (Adaptive Boosting) MLT [41]. In fact, AdaBoost MLT produces a set of sequential decision tree (J48) classifiers. Thanks to classifiers that were previously built, it adjusts the weights of the training samples. Here, the goal is to force the J48 classifier to reduce expected errors under different input distributions. In fact, the training samples, that are wrongly detected by former classifiers, will play an essential role in the training of later ones. Using AdaBoost MLT, a number of T weighted training subsets T_1, \dots, T_T are created in sequence and T classifiers C_1, \dots, C_T are construct. Then, the final decision, denoted by D_{final} , is made based on the decision of C_1, \dots, C_T classifiers through a weighted voting rule. We notice that the weight of each classifier is set based on its performance on the training subset employed to construct it. In Figure 3, we present the random committee and AdaBoost processes.

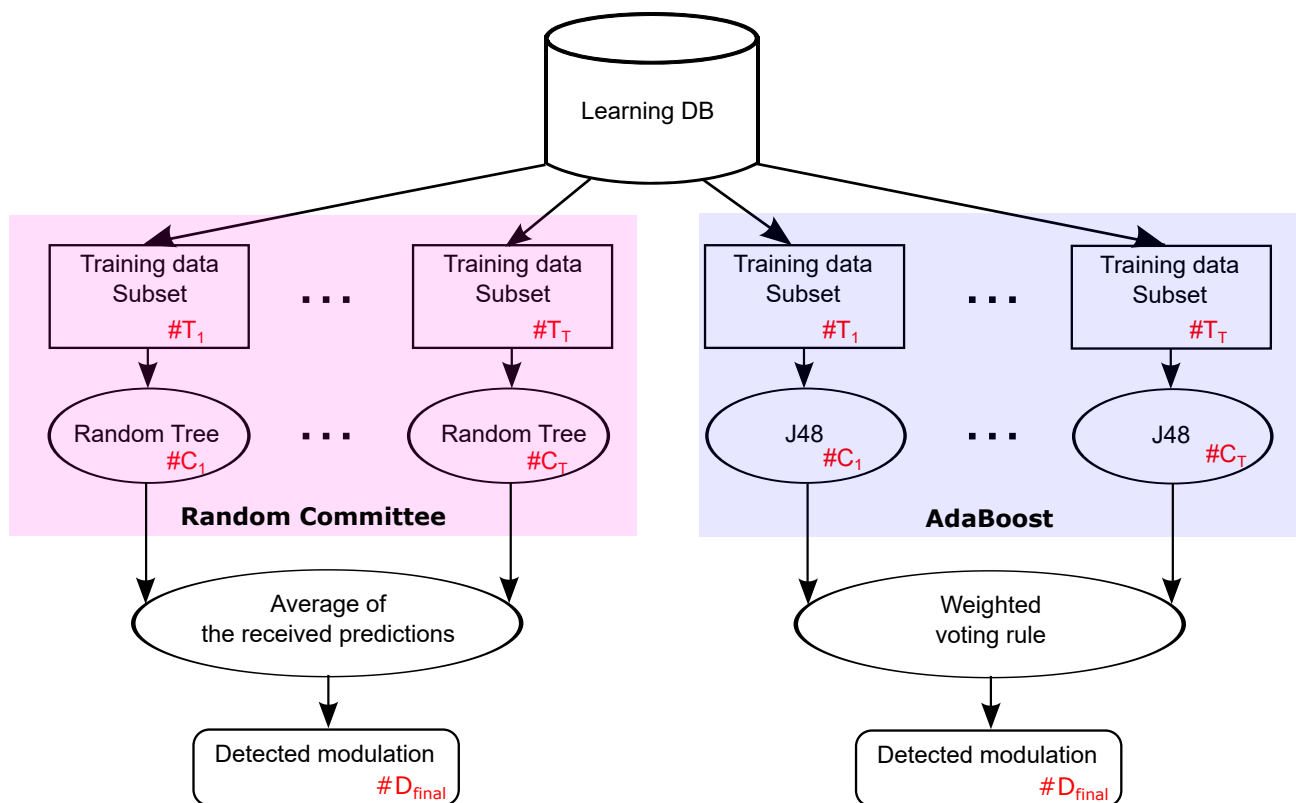


Figure 3. Random committee vs. AdaBoost processes.

Recall that T is the number of the generated training subsets, N_{feat} is the number of selected features, and let N_{samp} is the size of the learning DB, the time complexities of random committee using random tree MLT and AdaBoost using J48 MLT are given in Table 2.

Table 2. Time complexities of random committee using random tree classifier and AdaBoost using J48 MLT.

MLTs	Time Complexities
Random committee	$\mathcal{O}\left(T \times N_{samp} \times \log_2(N_{feat})\right)$
AdaBoost	$\mathcal{O}\left(T \times N_{samp}^2 \times N_{feat}\right)$

3.4. Multilayer Perceptron (MLP)

We also compare the proposed algorithm using the random committee MLT with the MLP MLT used in [24]. In fact, MLP is one of the most widely used artificial neural networks (ANN). It uses the resilient backpropagation (RPROP) proposed in [42], which is known for its good performance on pattern recognition algorithms. Note that the structure of the MLP contains one input layer, one or more hidden layers and one output layer. Here, each neuron of a layer is linked to all the neurons of the next layer. The intuition behind the introduction of this hidden layer is to enable the network to model the functions of complex nonlinear decision between any input and output layers. The optimal MLP structure to be employed in this work is determined using intensive simulations. In fact, we show that MLP with two hidden layers excluding the input and the output layers, where the first layer contains 10 nodes and the second has 15 nodes, provides a good trade-off between modulation detection and training time. Consequently, we use MLP with these settings in our simulation.

3.5. Metrics Used for Performance Evaluation of MLTs

In this paper, we compare between Random committee and AdaBoost MLTs using true positive (TP) rate, false positive (FP) rate, precision, recall, and F-Measure metrics. The precision, recall, and F-measure are given, respectively, as

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (21)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (22)$$

$$\text{F-measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (23)$$

4. Simulation Results

The performance of the proposed algorithm was verified for multi-relay cooperative MIMO systems over spatially correlated channels through numerical simulations. The simulated modulations set contains the following list: $\mathcal{M} = \{16QAM, 64QAM, BPSK, QPSK \text{ and } 8PSK\}$. A training set is built for each modulation type based on 50 realizations of signals with $512 \times N_A$ symbols, where the messages transmitted by the source node and MIMO channels are randomly generated in each realization. We assume that all sub-channels, i.e., SD , SR_l , and R_lD , $l = 1, \dots, L$, have the same correlation coefficient, i.e., $|\rho| = |\rho_{H_{ch}, RX}| = |\rho_{H_{ch}, TX}|$, where $ch = (SD, SR_l, R_lD)$. We also assume that the sub-channels SD and R_lD have the same variance of estimation error, i.e., $e^2 = e_{SD}^2 = e_{R_lD}^2$. In all results, we consider that all nodes are equipped with four antennas, i.e., $N_A = 4$. Added spatially white circularly complex Gaussian noises with variances σ_{ch}^2 are considered. Without loss of generality, we suppose that the sub-channels SR_l have the same SNR, i.e., $SNR_{SR} = SNR_{SR_l} = 20$ dB, $l = 1, 2, \dots, L$. Furthermore, we consider that the sub-channels SD and RD have equal SNRs, i.e., $SNR = SNR_{SD} = SNR_{RD}$.

In this work, we carry out a comparative study between the random committee MLT with the AdaBoost MLT using a 10-fold cross-validation [43] on the training set described above. Here, the number of training subsets is set to ten for both random committee and AdaBoost MLTs (i.e., $T = 10$). For all the results, we consider a cooperative MIMO system with $N_A = 4$, $L = 3$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB and $e^2 = 0.1$ where $|\rho| = 0.5$. Here, we choose $SNR = -5$ dB to evaluate the MLTs performance at low SNR as a good detection performance, i.e., 100%, can be achieved at acceptable SNR values as shown in our proposal presented in [21].

Tables 3 and 4 display the detailed accuracy by modulation type for the random committee and the Adaboost MLTs, respectively, in the case of cooperative MIMO system with $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB, $e^2 = 0.1$ and $|\rho| = 0.5$. By comparing the average of the presented metrics, i.e., TP rate, FP rate, precision, recall, and F-Measure, it is clearly shown that the random committee MLT offers a gain compared to the Adaboost MLT in terms of modulation detection as the values of TP rate, precision, recall, and F-Measure of the random committee MLT are higher than the ones of the Adaboost MLT. However, the value of the FP rate for the random committee is lower than the one of the Adaboost MLT. Therefore, the random committee MLT can be adopted for modulation detection.

Tables 5 and 6 confirm the obtained results. In fact, the percentages of correctly detected modulations are 84.86% and 84.285% for random committee and the Adaboost MLTs, respectively.

Table 3. Detailed accuracy by modulation type for the random committee using random tree as a base MLT with $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$.

\mathcal{M}	TP Rate	FP Rate	Precision	Recall	F-Measure
BPSK	0.952	0.014	0.945	0.952	0.949
QPSK	0.703	0.081	0.686	0.703	0.694
8PSK	0.652	0.071	0.696	0.652	0.674
16QAM	0.960	0.014	0.945	0.960	0.952
64QAM	0.976	0.010	0.962	0.976	0.969
Avg.	0.849	0.038	0.847	0.849	0.847

Table 4. Detailed accuracy by modulation type for Adaboost using J48 as a base MLT with $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$.

\mathcal{M}	TP Rate	FP Rate	Precision	Recall	F-Measure
BPSK	0.951	0.013	0.949	0.951	0.950
QPSK	0.659	0.079	0.677	0.659	0.668
8PSK	0.663	0.083	0.667	0.663	0.665
16QAM	0.965	0.014	0.945	0.965	0.954
64QAM	0.978	0.008	0.967	0.978	0.972
Avg.	0.843	0.039	0.841	0.843	0.842

Table 5. Confusion matrix for random committee using random tree as a base MLT with $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$.

a	b	c	d	e	Detected as
3807	55	78	40	20	a = BPSK
75	2813	1026	54	32	b = QPSK
109	1211	2609	49	22	c = 8PSK
30	19	31	3838	82	d = 16QAM
6	5	2	82	3905	e = 64QAM
Correctly detected modulations				16,972	84.86%
Incorrectly detected modulations				3028	15.14%

Table 6. Confusion matrix for Adaboost using J48 as a base MLT with $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $SNR = -5$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$.

a	b	c	d	e	detected as
3803	62	82	39	14	a = BPSK
71	2635	1214	55	25	b = QPSK
107	1168	2650	56	19	c = 8PSK
21	24	21	3858	76	d = 16QAM
6	3	4	76	3911	e = 64QAM
Correctly detected modulations				16,857	84.285%
Incorrectly detected modulations				3143	15.715%

To confirm these results, we show in Figure 4 the probability of correct detection ($P_{\text{Correct detection}}$) of the proposed algorithm using random committee MLT as a function of the SNR compared to the algorithm proposed in [21] using J48 MLT alone and AdaBoost MLT operating with J48, where $L = 3$, $N_A = 4$, $SNR_{SR} = 20$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$. Here, we consider that the test set consists of 1000 Monte Carlo trials for each modulation scheme (i.e., 5000 Monte Carlo trials in total $N_{\text{total}} = 5000$). For each trial, N_A test signals are considered where each signal consists of 512 i.i.d. symbols. It clearly shown that the

proposed algorithm provides a good performance, as $P_{\text{Correct detection}}$ reaches 100% (i.e., $P_{\text{Correct detection}} \simeq 100\%$) at acceptable SNR. Furthermore, one can see that the modulation detection of our proposal is enhanced compared to the algorithm proposed in [21] for the two cases: using both the AdaBoost and J48 MLTs. For example, $P_{\text{Correct detection}}$ achieves about 100% at SNR equal to 5 dB and 10 dB for our proposal and the algorithm proposed in [21] using AdaBoost MLT, respectively. It is also shown that the MLP has the worst performance.

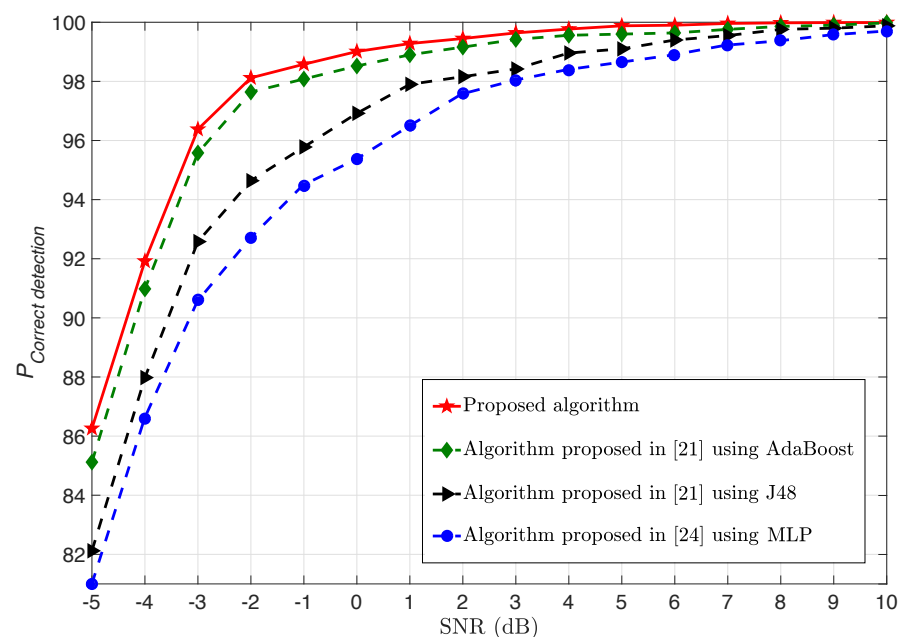


Figure 4. $P_{\text{Correct detection}}$ as a function of SNR for the proposed algorithm using random committee compared to MLP [24], algorithm proposed in [21] using J48 MLT alone and AdaBoost operating with J48, where $L = 3$, $N_A = 4$, $\text{SNR}_{\text{SR}} = 20$ dB, $e^2 = 0.1$, and $|\rho| = 0.5$.

In addition to the provided modulation detection gain, the random committee MLT has a low complexity compared to the Adaboost MLT. Indeed, the required time taken to build the model for the random committee is more than seven times faster than the Adaboost MLT as shown in Figure 5. One can also show that the training time of the MLP is long. Consequently, the proposed algorithm provides a good tradeoff between modulation detection performance and complexity.

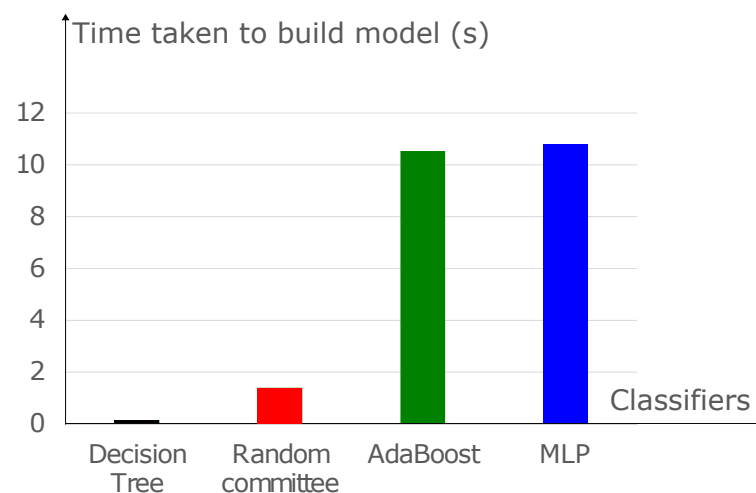


Figure 5. Time taken to build model: random committee Vs. AdaBoost Vs. J48 Vs. MLP with $L = 3$, $N_A = 4$, $\text{SNR}_{\text{SR}} = 20$ dB, $\text{SNR} = -5$ dB, $e^2 = 0.1$ and $|\rho| = 0.5$.

5. Conclusions

In this paper, we studied the detection of modulations for the multi-relay cooperative MIMO systems in the presence of spatially correlated channels. At destination node, we extracted the higher-order statistics (HOSs) as discriminating features of the received signals. After applying the principal component analysis (PCA) technique, we carried out a comparative study between the random committee and the AdaBoost MLTs at low SNR. The efficiency metrics, including the true positive rate, false positive rate, precision, recall, F-Measure, and the time taken to build the model, are used for the performance comparison. Simulation results demonstrated that the use of the random committee MLT, as compared to the AdaBoost MLT, offers gain in terms of the complexity and modulation detection.

Author Contributions: Conceptualization: H.B.C.; methodology: H.B.C. and A.A.; validation: H.B.C. and W.K.; formal analysis: H.B.C. and A.A.; writing—original draft preparation: H.B.C. and A.A.; writing—review and editing: H.B.C. and W.K.; supervision: H.B.C. and W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Jounf university.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are not publicly available due to project restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Based on (17), the second-order HOM can be expressed as

$$M_{20} = \mathbb{E} \left\{ \left(\mathbf{y}_D^{(a)} \right)^2 \right\}, \quad \text{and} \quad M_{21} = \mathbb{E} \left\{ \left| \mathbf{y}_D^{(a)} \right|^2 \right\}. \quad (\text{A1})$$

C_{20} and C_{21} are written using (19) as

$$C_{20} \left(\mathbf{y}_D^{(a)} \right) = \text{Cum} \left[\mathbf{y}_D^{(a)}, \mathbf{y}_D^{(a)} \right], \quad C_{21} \left(\mathbf{y}_D^{(a)} \right) = \text{Cum} \left[\mathbf{y}_D^{(a)}, \overline{\mathbf{y}_D^{(a)}} \right]. \quad (\text{A2})$$

Assuming $m = 3$, the available set of indices is $(1, 2, 3)$ and four distinct types of partitioning can be obtained for that set: $\{(1), (2), (3)\}$ leading to $\beta = 3$, $\{(1), (2, 3)\}$ leading to $\beta = 2$, $\{(2), (1, 3)\}$ leading to $\beta = 2$, $\{(3), (1, 2)\}$ leading to $\beta = 2$, $\{(1, 2, 3)\}$ leading to $\beta = 1$. Therefore,

$$\begin{aligned} \text{Cum} \left[\mathbf{y}_{D1}^{(a)}, \mathbf{y}_{D1}^{(a)}, \mathbf{y}_{D3}^{(a)} \right] &= (-1)^{1-1} (1-1)! \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D2}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} \\ &\quad + (-1)^{2-1} (2-1)! \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} \\ &\quad + (-1)^{2-1} (2-1)! \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} \\ &\quad + (-1)^{2-1} (2-1)! \mathbb{E} \left\{ \mathbf{y}_{D3}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D2}^{(a)} \right\} \\ &\quad + (-1)^{3-1} (3-1)! \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D3}^{(a)} \right\} \\ &= \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D2}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} - \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} \\ &\quad - \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D3}^{(a)} \right\} - \mathbb{E} \left\{ \mathbf{y}_{D3}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \mathbf{y}_{D2}^{(a)} \right\} \\ &\quad + 2 \mathbb{E} \left\{ \mathbf{y}_{D1}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D2}^{(a)} \right\} \mathbb{E} \left\{ \mathbf{y}_{D3}^{(a)} \right\} \end{aligned} \quad (\text{A3})$$

By following the same manner, one can express the HOC up to eighth order in terms of HOMs. For example, C_{40} , C_{60} , and C_{80} are defined, respectively, as

$$C_{40} = M_{40} - 3(M_{20})^2, \quad (\text{A4})$$

$$C_{60} = M_{60} - 15M_{20}M_{40} + 30(M_{20})^3, \quad (\text{A5})$$

$$C_{80} = M_{80} - 35(M_{40})^2 - 630(M_{20})^4 + 420(M_{20})^2M_{40}. \quad (\text{A6})$$

References

- Li, Y.N.R.; Gao, B.; Zhang, X.; Huang, K. Beam Management in Millimeter-Wave Communications for 5G and Beyond. *IEEE Access* **2020**, *8*, 13282–13293. [\[CrossRef\]](#)
- Nadeem, Q.-U.-A.; Kammoun, A.; Alouini, M.S. Elevation beamforming with full dimension MIMO architectures in 5G systems: A tutorial. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3238–3273. [\[CrossRef\]](#)
- Moualeu, J.M.; Hamouda, W.; Xu, H.; Takawira, F. Multi-Relay Turbo-Coded Cooperative Diversity Networks Over Nakagami-*m* Fading Channels. *IEEE Trans. Veh. Technol.* **2013**, *62*, 4458–4470. [\[CrossRef\]](#)
- Fang, Y.; Chen, P.; Cai, G.; Lau, F.C.; Liew, S.C.; Han, G. Outage-limit-approaching channel coding for future wireless communications: Root-protograph low-density parity-check codes. *IEEE Veh. Technol. Mag.* **2019**, *14*, 85–93. [\[CrossRef\]](#)
- Arti, M.; Mallik, R.K.; Schober, R. Beamforming and combining in two-way AF MIMO relay networks. *IEEE Commun. Lett.* **2013**, *17*, 1400–1403. [\[CrossRef\]](#)
- Arti, M.; Bhatnagar, M.R. Maximal Ratio Transmission in AF MIMO Relay Systems Over Nakagami-Fading Channels. *IEEE Trans. Veh. Technol.* **2015**, *64*, 1895–1903.
- Sugiura, S.; Chen, S.; Hanzo, L. MIMO-aided near-capacity turbo transceivers: Taxonomy and performance versus complexity. *IEEE Commun. Surv. Tutor.* **2011**, *14*, 421–442. [\[CrossRef\]](#)
- Yang, Y.; Hu, H.; Xu, J.; Mao, G. Relay technologies for WiMAX and LTE-advanced mobile systems. *IEEE Commun. Mag.* **2009**, *47*, 100–105. [\[CrossRef\]](#)
- Mostafa, H.; Marey, M. Soft-Information Assisted Modulation Recognition for Reconfigurable Radios. *IEEE Wirel. Commun. Lett.* **2020**. [\[CrossRef\]](#)
- Eldemerdash, Y.A.; Dobre, O.A.; Üreten, O.; Yensen, T. A robust modulation classification method for PSK signals using random graphs. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 642–644. [\[CrossRef\]](#)
- Dabbagh, A.D.; Love, D.J. Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch. *IEEE Trans. Commun.* **2008**, *56*, 1859–1868. [\[CrossRef\]](#)
- Nguyen, D.D.; Liu, Y.; Chen, Q. On the energy efficient multi-pair two-way massive MIMO AF relaying with imperfect CSI and optimal power allocation. *IEEE Access* **2017**, *6*, 2589–2603. [\[CrossRef\]](#)
- Turan, M.; Öner, M.; Çırpan, H.A. Joint modulation classification and antenna number detection for MIMO systems. *IEEE Commun. Lett.* **2015**, *20*, 193–196. [\[CrossRef\]](#)
- Mohammadkarimi, M.; Karami, E.; Dobre, O.A. A novel algorithm for blind detection of the number of transmit antenna. In Proceedings of the International Conference on Cognitive Radio Oriented Wireless Networks, Doha, Qatar, 21–23 April 2015; pp. 441–450.
- Oularbi, M.R.; Gazor, S.; Aissa-El-Bey, A.; Houcke, S. Enumeration of base station antennas in a cognitive receiver by exploiting pilot patterns. *IEEE Commun. Lett.* **2012**, *17*, 8–11. [\[CrossRef\]](#)
- Marey, M.; Dobre, O.A.; Liao, B. Classification of STBC systems over frequency-selective channels. *IEEE Trans. Veh. Technol.* **2014**, *64*, 2159–2164. [\[CrossRef\]](#)
- Karami, E.; Dobre, O.A. Identification of SM-OFDM and AL-OFDM signals based on their second-order cyclostationarity. *IEEE Trans. Veh. Technol.* **2014**, *64*, 942–953. [\[CrossRef\]](#)
- Liu, K.; Da Costa, J.P.C.; So, H.C.; De Almeida, A.L. Semi-blind receivers for joint symbol and channel estimation in space-time-frequency MIMO-OFDM systems. *IEEE Trans. Signal Process.* **2013**, *61*, 5444–5457. [\[CrossRef\]](#)
- Cai, S.; Matsumoto, T.; Yang, K. SIMO channel estimation using space-time signal subspace projection and soft information. *IEEE Trans. Signal Process.* **2012**, *60*, 4219–4235. [\[CrossRef\]](#)
- Ullah, R.; Marwat, S.N.K.; Ahmad, A.M.; Ahmed, S.; Hafeez, A.; Kamal, T.; Tufail, M. A Machine Learning Approach for 5G SINR Prediction. *Electronics* **2020**, *9*, 1660. [\[CrossRef\]](#)
- Ben Chikha, H. Performance of AdaBoost in modulation recognition for spatially-correlated cooperative MIMO systems. In Proceedings of the 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 15–18 April 2018; pp. 1–6.
- Ben Chikha, W.; Chaoui, S.; Attia, R. Identification of superposed modulations for two-way relaying MIMO systems with physical-layer network coding. *IET Commun.* **2017**, *11*, 225–231. [\[CrossRef\]](#)
- Lau, K.; Salibian-Barrera, M.; Lampe, L. Modulation recognition in the 868 MHz band using classification trees and random forests. *AEU-Int. J. Electron. Commun.* **2016**, *70*, 1321–1328. [\[CrossRef\]](#)
- Ben Chikha, W.; Dayoub, I.; Hamouda, W.; Attia, R. Modulation Recognition for MIMO Relaying Broadcast Channels with Direct Link. *IEEE Wirel. Commun. Lett.* **2014**, *3*, 50–53. [\[CrossRef\]](#)
- Ben Chikha, W.; Dayoub, I.; Attia, R. Modulation detection in intelligent transportation systems based on cooperative MIMO networks over nakagami-*m* fading. In Proceedings of the 2014 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), Tunis, Tunisia, 3–6 November 2014; pp. 1–6.

26. Ha, C.B.; You, Y.H.; Song, H.K. Machine learning model for adaptive modulation of multi-stream in MIMO-OFDM system. *IEEE Access* **2018**, *7*, 5141–5152. [[CrossRef](#)]
27. Laneman, J.N.; Tse, D.N.; Wornell, G.W. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Trans. Inf. Theory* **2004**, *50*, 3062–3080. [[CrossRef](#)]
28. Peel, C.B.; Hochwald, B.M.; Swindlehurst, A.L. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: Channel inversion and regularization. *IEEE Trans. Commun.* **2005**, *53*, 195–202. [[CrossRef](#)]
29. Wan, H.; Chen, W.; Ji, J. Efficient linear transmission strategy for MIMO relaying broadcast channels with direct links. *IEEE Wirel. Commun. Lett.* **2012**, *1*, 14–17. [[CrossRef](#)]
30. Shiu, D.S.; Foschini, G.J.; Gans, M.J.; Kahn, J.M. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.* **2000**, *48*, 502–513. [[CrossRef](#)]
31. Loyka, S.L. Channel capacity of MIMO architecture using the exponential correlation matrix. *IEEE Commun. Lett.* **2001**, *5*, 369–371. [[CrossRef](#)]
32. Van Zelst, A.; Hammerschmidt, J. A single coefficient spatial correlation model for multiple-input multiple-output (MIMO) radio channels. In Proceedings of the URSI General Assembly, Maastricht, The Netherlands, 17–24 August 2002; pp. 17–24.
33. Zhang, B.; He, Z.; Niu, K.; Zhang, L. Robust linear beamforming for MIMO relay broadcast channel with limited feedback. *IEEE Signal Process. Lett.* **2010**, *17*, 209–212. [[CrossRef](#)]
34. Thornton, C.; Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 847–855.
35. Swami, A.; Sadler, B. Hierarchical digital modulation classification using cumulants. *IEEE Trans. Commun.* **2000**, *48*, 416–429. [[CrossRef](#)]
36. Dobre, O.A.; Abdi, A.; Bar-Ness, Y.; Su, W. Survey of automatic modulation classification techniques: Classical approaches and new trends. *IET Commun.* **2007**, *1*, 137–156. [[CrossRef](#)]
37. McCullagh, P. *Tensor Methods in Statistics*; Chapman Hall: London, UK, 1987.
38. Spooner, C.M. On the Utility of Sixth-Order Cyclic Cumulants for RF Signal Classification. In Proceedings of the Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 4–7 November 2001; Volume 1, pp. 890–897.
39. Kwak, N. Principal component analysis based on L1-norm maximization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 1672–1680. [[CrossRef](#)] [[PubMed](#)]
40. Cutler, A.; Zhao, G. *Fast Classification Using Perfect Random Trees*; Utah State University: Logan, UT, USA, 1999.
41. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. In Proceedings of the European Conference on Computational Learning Theory, Barcelona, Spain, 13–15 March 1995; pp. 23–37.
42. Riedmiller, M.; Braun, H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In Proceedings of the IEEE international conference on neural networks, San Francisco, CA, USA, 28 March–1 April 1993; pp. 586–591.
43. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-Validation. *Encycl. Database Syst.* **2009**, *5*. [[CrossRef](#)]