Automatic Modulation Classification for MIMO Systems Using Fourth-Order Cumulants

Michael S. Mühlhaus*, Mengüç Öner†, Octavia A. Dobre‡, Holger U. Jäkel*, Friedrich K. Jondral*

* Karlsruhe Institute of Technology, Germany, Email: {michael.muehlhaus, holger.jaekel, friedrich.jondral}@kit.edu
†Isik University, Sile, Turkey, Email: oner@isikun.edu.tr
‡ Memorial University, St. John's, Canada, Email: odobre@mun.ca

Abstract—**Automatic classification of the modulation type of an unknown communication signal is a challenging task, with applications in both commercial and military contexts, such as spectrum surveillance, cognitive radio, and electronic warfare systems. Most of the automatic modulation classification (AMC) algorithms found in the literature assume that the signal of interest has been transmitted using a single antenna. In this paper, a novel AMC algorithm for multiple input multiple output (MIMO) signals is proposed, which employs fourth-order cumulants as features for classification. First, perfect channel state information (CSI) is assumed. Subsequently, a case of more practical relevance is considered, where the channel matrix is unknown and has to be estimated blindly by employing independent component analysis (ICA). The performance of the proposed classification algorithm is investigated through simulations and compared with an average likelihood ratio test (ALRT) which can be considered as optimum in the Bayesian sense, but has a very high computational complexity.**

Index Terms—**Automatic modulation classification, multiple input multiple output, independent component analysis, blind source separation, fourth-order cumulant.**

I. Introduction

Automatic modulation classification (AMC) techniques, which are used for identifying the modulation type of an unknown modulated signal corrupted by noise and fading, play an important role in spectrum surveillance and cognitive radio applications. AMC is essentially a multiple hypothesis testing problem, where, by using the received signal of interest in a given observation interval, the classifier has to infer which modulation type M_i from the set $\mathcal{M}$ of possible modulation schemes has been employed by the transmitter.

Research in the field of AMC has been intensively carried out for single input single output (SISO) systems in the last decades. In the literature, two fundamentally different approaches are observed for tackling the task of SISO AMC: Likelihood-based tests, which provide optimal performance in the Bayesian sense, but require a large amount of a-priori information and a high computational power, and feature-based classifiers that are sub-optimal but require less a priori information and have a lower computational complexity [1].

The feature-based classifiers extract modulation-type specific features out of the received signal and use them to discriminate among the elements of the set of candidate modulation types, $\mathcal{M}$. The instantaneous frequency, phase, and amplitude [2], and higher-order statistics, such as cumulants of orders four or higher, are examples of features exploited for classification [3], [4], [5], [6].

Recently, multiple input multiple output (MIMO) systems have been drawing considerable interest, where multiple transmit and receive antennas are used to increase the data rates and/or the robustness of the communication. To the best of our knowledge, only two authors have investigated the problem of AMC in MIMO systems in the literature. Choqueuse et al. [7] have formulated an average likelihood ratio test (ALRT) for MIMO spatial multiplexing systems, while Hassan et al. [8] have proposed a neural network-based classifier. In this work, we propose a novel feature-based classifier for MIMO systems by exploiting the fourth-order cumulants of the received signal. The performance of the proposed classifier is investigated via simulations and compared with the ALRT test provided in [7], which can be considered as optimal in the Bayesian sense. First, the ideal case is considered, where the receiver has perfect knowledge of the channel matrix (as in the case of ALRT). Subsequently, a more practical case is investigated, where the receiver has no knowledge of the channel matrix and estimates it blindly using the independent component analysis (ICA).

The rest of the paper is organized as follows. In Section II the MIMO system model is described. Section III presents the ALRT, which is used as a benchmark for the performance of the proposed algorithm. In Section IV the proposed fourth-order cumulant-based classifier is introduced and both the cases with and without the knowledge of the channel matrix are considered. Subsequently, the performance of the proposed classifier is investigated using simulations in Section V. The paper concludes with Section VI.

II. MIMO System Model

Fig. 1 shows the general setup of a MIMO system with N_t transmit and N_r receive antennas. The transmitter encodes a data block of length N using a specific space-time code. In this work, as in [7], the case of a vertical Bell Labs layered (V-BLAST) transmitter architecture is considered, which is also referred to as spatial multiplexing. In such a case, the data symbols are multiplexed into N_t independent data streams; hence, the transmitted symbols are independent both in time and space, and the overall data rate is N_t times higher compared to that of a SISO system. The received signal vector at time instant k, affected by fading and complex circular
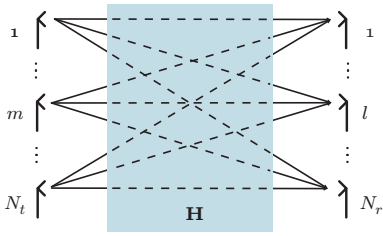
Fig. 1. General setup of a MIMO system with N_t transmit and N_r receive antennas.

additive white Gaussian noise (AWGN), is expressed as

$$\mathbf{r}[k] = \mathbf{H}\mathbf{s}[k] + \mathbf{w}[k], \qquad (1)$$

where $\mathbf{r}[k] = [r_1[k], \ldots, r_{N_r}[k]]^T$ is the vector consisting of the received signals at each of the N_r receive antennas, $\mathbf{s}[k] = [s_1[k], \ldots, s_{N_t}[k]]^T$ is the vector consisting of the transmitted symbol streams at each of the N_t transmit antenna, and $\mathbf{w}[k]$ is the noise vector of length N_r. The MIMO channel is described by the channel matrix

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r,1} & \cdots & h_{N_r,N_t} \end{bmatrix}, \qquad (2)$$

where $h_{m,l}$, $m = 1, \ldots, N_r$, $l = 1, \ldots, N_t$, represents the channel coefficient between the m-th receive antenna and l-th transmit antenna. The channel coefficients are modeled as statistically independent Gaussian distributed complex circular random variables with zero mean and unit variance. Throughout the paper we assume that the channel is a flat block fading channel, i.e., the channel matrix remains constant for a given observation interval.

Without loss of generality, we assume that the variance σ_S^2 of the transmitted symbols per antenna equals one. The average signal-to-noise ratio (SNR) is defined as the total received mean signal power, $N_t N_r \sigma_S^2 = N_t N_r$, divided by the noise power $N_r \sigma_w^2$

$$\text{SNR} = \frac{N_t \sigma_S^2}{\sigma_w^2} = \frac{N_t}{\sigma_w^2}. \qquad (3)$$

III. AVERAGE LIKELIHOOD RATIO TEST CLASSIFIER

In [7], the ALRT-based classifier has been proposed for MIMO spatial multiplexing systems, under the assumption of a known channel matrix and noise variance. The classifier decides for the modulation type $M_i \in \mathcal{M}$ which maximizes the log-likelihood function, i.e.,

$$M_i = \arg \max_{M_q \in \mathcal{M}} \left(\lambda \left(\mathbf{R} | M_q, \mathbf{H}, \sigma_w^2 \right) \right), \qquad (4)$$

with the log likelihood function $\lambda(\cdot)$ given as

$$\lambda(\mathbf{R}|M_q, \mathbf{H}, \sigma_w^2) = -TN_t \log(m_q) - N_r \log(2\pi\sigma_w^2)$$
$$+ \sum_{k=1}^{T} \log \left(\sum_{p=1}^{(m_q)^{N_t}} \exp \left(-\frac{||\mathbf{r}[k] - \mathbf{H}\tilde{\mathbf{s}}_p[k]||^2}{2\sigma_w^2} \right) \right), \qquad (5)$$

where $\mathbf{R} = [\mathbf{r}[1], \mathbf{r}[2], \ldots, \mathbf{r}[T]]$ is the received signal matrix for an observation interval of length T, m_q represents the number of constellation points of the modulation type M_q, $|| \cdot ||$ is the Euclidian norm, and the inner sum is taken over all $(m_q)^{N_t}$ possible transmitted signal vectors $\tilde{\mathbf{s}}_p[k]$ for the modulation type M_q (see [7] for further details).

Note that the ALRT is optimal in the Bayesian sense, provided that the channel matrix and the noise variance are known. However, it is very computationally intensive, and when the number of transmit antennas N_t and/or the order of the modulation scheme m_q are large, it basically becomes impractical. In this work, the ALRT is used as a performance benchmark for the proposed feature-based AMC.

IV. PROPOSED CUMULANT-BASED CLASSIFIER

Due to the high computational complexity of the ALRT, alternative classification methods should be considered for practical applications. In the literature, higher-order cumulants are widely used as features for AMC in SISO systems [3], [4], [6]. The fourth-order non-conjugate cumulant for a complex-valued zero-mean stationary process $x[k]$ is defined as [9]

$$\kappa = E\{x[k]^4\} - 3E\{x[k]^2\}^2. \qquad (6)$$

The fourth-order cumulants for different modulation schemes normalized to unit variance are listed in Table I.

	BPSK	QPSK	8-PSK	16-QAM
κ_{M_q}	-2	1	0	-0.68

TABLE I
THEORETICAL VALUES OF CUMULANTS FOR DIFFERENT MODULATION SCHEMES WITH UNIT ENERGY.

The cumulant features cannot be directly extracted from the components of the received signal vector $\mathbf{r}[k]$ at the receive antennas, since they consist of linear mixtures of the components of the transmitted signal vector $\mathbf{s}[k]$ and additive noise. Thus, we estimate the channel matrix and compensate the channel effect prior to the extraction of the cumulant features and classification. The proposed classification algorithm is illustrated in Fig. 2. First, channel estimation and compensation are carried out, and estimates of the N_t individual transmitted signals $\hat{s}_m[k], m = 1, ..., N_t$ are obtained. Subsequently, the cumulants are estimated for each of the N_t branches, and finally, these are combined and used for classification. In this work, we consider both the ideal case of perfect knowledge of the channel matrix $\mathbf{H}$ and the case where $\mathbf{H}$ and $\hat{\mathbf{s}}[k]$ are estimated blindly using ICA.
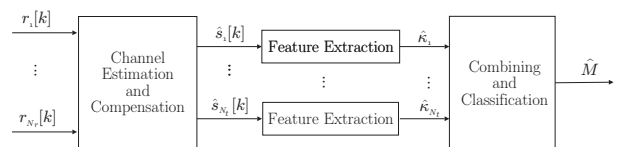

Fig. 2. Proposed classifier.

A. Perfect Channel State Information (CSI)

If the channel matrix $\mathbf{H}$ is known, the estimation of $\hat{\mathbf{s}}[k] = [\hat{s}_1[k], ..., \hat{s}_{N_t}[k]]^T$ can be carried out using zero forcing (ZF) and minimum mean square error (MMSE) methods.

The ZF method: It is assumed that $\mathbf{H}$ has the full column rank N_t. The received signal vector $\mathbf{r}[k]$ is multiplied by the pseudo inverse of the channel matrix $\mathbf{H}^+ := (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$, where $(\cdot)^H$ denotes the Hermitian operation. The estimated signal $\hat{\mathbf{s}}[k]$ becomes:

$$
\begin{aligned}
\hat{\mathbf{s}}[k] &= \mathbf{H}^+\mathbf{r}[k] \\
&= (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H(\mathbf{H}\mathbf{s}[k] + \mathbf{w}[k]) \\
&= \mathbf{s}[k] + \mathbf{H}^+\mathbf{w}[k] = \mathbf{s}[k] + \tilde{\mathbf{w}}[k].
\end{aligned}
\tag{7}
$$

In a noise free scenario, the ZF reconstructs $\mathbf{s}[k]$ perfectly. However, if the channel is badly conditioned (having small singular values), the noise is severely amplified. The noise variance equals the diagonal entries of the covariance matrix of the filtered noise $\tilde{\mathbf{w}}[k]$, which is no longer white. The noise power at antenna m is $P_m = (\text{cov}(\tilde{\mathbf{w}}[k]))_{mm}$ with

$$
\begin{aligned}
\text{cov}(\tilde{\mathbf{w}}[k]) &= (\mathbf{H}^+)\text{cov}(\mathbf{w}[k])(\mathbf{H}^+)^H \\
&= (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\sigma_w^2\mathbf{I}_{N_r}\mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1} \\
&= \sigma_w^2(\mathbf{H}^H\mathbf{H})^{-1},
\end{aligned}
\tag{8}
$$

where $\mathbf{I}_{N_r}$ is an $N_r \times N_r$ identity matrix and σ_w^2 is the AWGN noise variance.

The MMSE method: Instead of $\mathbf{H}^+$ as for ZF, $\mathbf{H}_{MMSE}$ is given by

$$
\mathbf{H}_{MMSE} := (\mathbf{H}^H\mathbf{H} + \sigma_w^2\mathbf{I}_{Nt})^{-1}\mathbf{H}^H.
\tag{9}
$$

Comparable to the ZF detector, the noise variance can be calculated as

$$
\begin{aligned}
\text{cov}(\tilde{\mathbf{w}}[k]) &= (\mathbf{H}_{MMSE})\text{cov}(\mathbf{w}[k])(\mathbf{H}_{MMSE})^H \\
&= (\mathbf{H}^H\mathbf{H} + \sigma_w^2\mathbf{I}_{N_t})^{-1}\mathbf{H}^H\sigma_w^2\mathbf{I}_{N_r}\mathbf{H} \\
&\quad \times (\mathbf{H}^H\mathbf{H} + \sigma_w^2\mathbf{I}_{N_t})^{-1} \\
&= \sigma_w^2\left(\mathbf{H}^H\mathbf{H} + \sigma_w^4(\mathbf{H}^H\mathbf{H})^{-1} + 2\sigma_w^2\mathbf{I}_{N_t}\right)^{-1}.
\end{aligned}
\tag{10}
$$

Note that in both cases, the noise vector after compensation $\tilde{\mathbf{w}}[k]$ is non-white, and, thus, the noise terms at each of the N_t estimated signal streams are correlated and have different powers. Furthermore, for the feature extraction, the contribution of the signal power in the total power of the reconstructed symbol streams $\hat{s}_m[k]$ has to be normalized to one. This is done in case of ZF by definition, since the signal part $\mathbf{s}[\mathbf{k}]$ is perfectly reconstructed and the power of the transmit symbols is normalized to one. For MMSE, however, the reconstructed signals must be normalized according to:

$$
\bar{s}_m[k] = \frac{\hat{s}_m[k]\sqrt{1 + P_m}}{\sigma_{\hat{s}_m}},
\tag{11}
$$

where $\bar{s}_m[k]$ is the m-th normalized estimated signal stream, $\sigma_{\hat{s}_m}^2$ is the variance of the estimated signal at the m-th signal stream, P_m is the noise, amplified according to equation (8)

and can be interpreted as the noise power at the m-th estimated symbol stream.

B. Independent Component Analysis

In most of the scenarios where AMC is employed, the perfect knowledge of the channel matrix by the classifier is an unrealistic assumption. Why should a receiver have complete knowledge about the channel, but no information about the used modulation scheme? In such cases, the channel matrix needs to be estimated blindly and compensated for obtaining the estimates of the transmitted symbol streams. Since the transmitted symbol streams of a MIMO V-BLAST system are statistically independent, and the received signal consists of a linear mixture of these independent symbol streams plus noise, the ICA framework is a useful tool for this task.

The ICA is an umbrella term for computational methods that are used to blindly separate linear mixtures of random processes into their individual components. Most of the ICA algorithms are based on three fundamental assumptions: the random processes to be separated have to be statistically independent, at most one of the components may have a Gaussian distribution, and the number of sensors have to be equal to or larger than the number of independent components [10]. Obviously, the MIMO spatial multiplexing system investigated in this work fulfills all three requirements, with $\mathbf{H}$ as the mixing matrix and the transmitted signal streams $s_m[k]$ as the independent components to be separated prior to the feature extraction stage in Fig. 2. The ICA algorithms found in the literature are basically iterative optimization algorithms with the task of maximizing an objective function that characterizes the statistical independence of the separated components. The various ICA algorithms found in the literature differ in the metric that measures the independence and in the methodology of solving the optimization problem. For a thorough account on ICA, the interested reader is referred to [10] and the references therein.

In this work, we use the well known joint approximate diagonalization of eigen-matrices (JADE) algorithm proposed in [11] for the estimation of the channel matrix $\mathbf{H}$ and the separation of the independent symbol streams $s_m[k]$ from the noisy linear mixtures $r_l[k]$. Note that ICA algorithms are able to separate independent components up to a multiplicative factor and a phase rotation. Thus, the estimated transmit signals have to be normalized, as in the case of MMSE (see eq. (11)). For both cases, this normalization is carried out in the feature extraction blocks.

C. Feature Extraction and Classification

Following the compensation of the channel, the estimated transmitted signals are normalized if necessary and the cumulant features $\hat{\kappa}_1, \cdots, \hat{\kappa}_{N_t}$ are extracted for each of the N_t branches. Subsequently, the individual estimated features are combined and used in the classification unit. In this work, we propose the following linear combination of the features to be

used for classification

$$\hat{K} = \frac{\sum_{m=1}^{N_t} \frac{|\hat{\kappa}_m|}{P_m^2}}{\sum_{m=1}^{N_t} \frac{1}{P_m^2}}, \qquad (12)$$

i.e., the individual cumulant estimates are weighted with the square of the SNR at their respective signal path ($\sim 1/P_m$) and averaged. Thus, the cumulant estimates which are affected by more noise carry only a smaller weight, while those which are less affected by noise carry a larger weight.[1]

The decision on a specific modulation scheme $\hat{M} = M_i$ is performed by employing the criterion of minimum Euclidean distance between the combined feature estimate $\hat{K}$ and the absolute value of the theoretical values

$$\hat{M} = \arg \min_{M_q \in \mathcal{M}} \sqrt{\left(\left| \kappa_{M_q} \right| - \hat{K} \right)^2}. \qquad (13)$$

V. SIMULATION RESULTS

The classification performance is analyzed by simulations and measured by the average probability of correct classification defined as

$$P_{CC} = \sum_{i=1}^{|\mathcal{M}|} P(\hat{M} = M_i | M_i) P(M_i), \qquad (14)$$

where $|\mathcal{M}|$ is the cardinality of the set of possible modulation schemes. $P(M_i)$ is the transmission probability of the modulation scheme M_i. It is assumed that all modulations are equally probable. $P(\hat{M} = M_i | M_i)$ is the probability that the modulation scheme M_i is classified given that M_i is transmitted.

The following simulation parameters are used unless mentioned otherwise. The set of modulation schemes is $\mathcal{M} = \{$BPSK, QPSK, 8-PSK, 16-QAM$\}$, the number of transmit antennas is $N_t = 2$, the number of receive antennas is $N_r \in \{2, 4, 6\}$ and the observation length is $T = 1000$. For each SNR value, 5000 iterations were performed.

A. Performance with perfect channel knowledge

Fig. 3 shows the classification performance of ALRT. It can be seen that the probability of correct classification of 0.9 is achieved at an SNR of about 0 dB if two receive antennas are used. The use of 4 instead of 2 receive antennas gains 4 dB SNR. Two more receive antennas bring a further performance improvement of 2 dB.

Fig. 4 shows the performance of the cumulant-based classifier with ZF and MMSE equalizers in comparison to ALRT. Note that for all three cases, perfect knowledge of the CSI is assumed. The performance of the proposed classifier with MMSE is a marginally better than with ZF, and is approximately 5 dB worse compared with ALRT at a P_{CC} of 0.9. While this loss in performance is considerable, the computational complexity of the AMC is reduced to only a fraction of that of ALRT. This makes the cumulant-based classifier a good candidate for practical applications.

---

[1]The idea is similar to maximal ratio combining. A detailed statistical analysis of the combined feature $\hat{K}$ will be provided in future work.
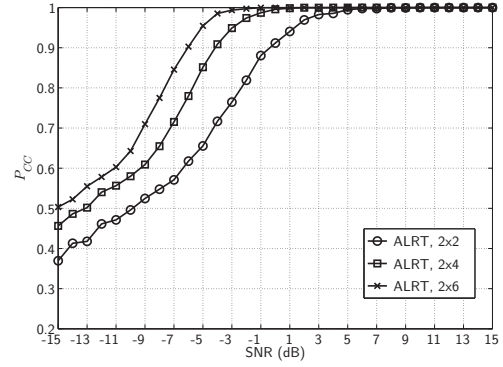

Fig. 3. Classification performance of ALRT with $T = 1000$, $N_t = 2$, $N_r \in \{2, 4, 6\}$, and $\mathcal{M} = \{$BPSK, QPSK, 8-PSK, 16-QAM$\}$.
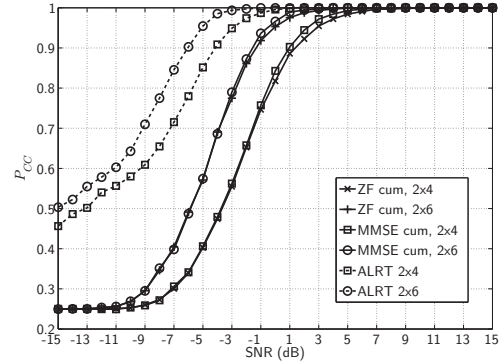

Fig. 4. Performance of the proposed classifier with perfect CSI using MMSE and ZF with $T = 1000$, $N_t = 2$, $N_r \in \{4, 6\}$, and $\mathcal{M} = \{$ BPSK, QPSK, 8-PSK, 16-QAM $\}$.

B. Performance with estimated channel

Fig. 5 compares the performance of the cumulant-based classifier with blind channel estimation using JADE ICA to the performance with perfect CSI knowledge. It can be seen that the performance of the cumulant-based classifier with ICA is very close to the MMSE classifier with perfect knowledge of the CSI.
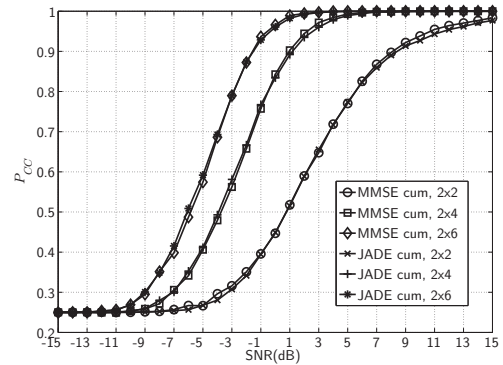

Fig. 5. Classification performance of the proposed classifier with perfect CSI using MMSE and without CSI using JADE ICA with $T = 1000$, $N_t = 2$, $N_r = 4$, and $\mathcal{M} = \{$BPSK, QPSK, 8-PSK, 16-QAM $\}$.

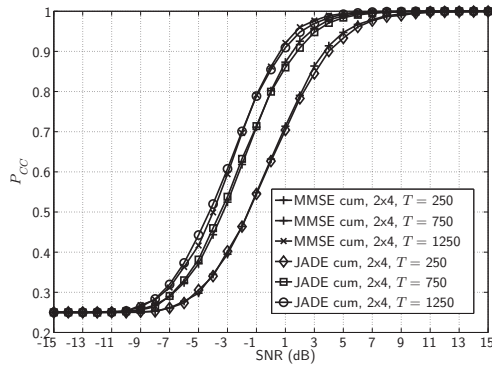Fig. 6 shows the performance of the cumulant-based clas-

Fig. 6. Effect of the length of the observation window T on the classification performance of the proposed classifier with perfect CSI using MMSE and without CSI using JADE ICA with $N_t = 2$, $N_r = 4$, $\mathcal{M}$ ={BPSK, QPSK, 8-PSK, 16-QAM}.

SNR= −3 dB	M_k \ M_i	BPSK	QPSK	8-PSK	16-QAM
	BPSK	0.9155	0.0760	0.0005	0.0080
	QPSK	0.1345	0.6235	0.0135	0.2285
	8-PSK	0.0790	0.1800	0.2680	0.4730
	16-QAM	0.1010	0.3415	0.0400	0.5175

SNR= 0 dB	M_k \ M_i	BPSK	QPSK	8-PSK	16-QAM
	BPSK	0.9800	0.0180	0	0.0020
	QPSK	0.0260	0.8460	0.0035	0.1245
	8-PSK	0.0135	0.0325	0.7245	0.2295
	16-QAM	0.0135	0.1735	0.0270	0.7860

SNR= 3 dB	M_k \ M_i	BPSK	QPSK	8-PSK	16-QAM
	BPSK	0.9985	0.0015	0	0
	QPSK	0.0040	0.9580	0.0015	0.0365
	8-PSK	0.0015	0.0045	0.9455	0.0485
	16-QAM	0.0005	0.0510	0.0095	0.9390

TABLE II
CONFUSION MATRICES OF THE PROPOSED CUMULANT-BASED CLASSIFIER USING BLIND CHANNEL ESTIMATION WITH JADE.

sifier with (MMSE) and without CSI knowledge using JADE ICA for different observation lengths. While the increase of the observation length from 250 to 750 symbols gains about 2 dB, the further increasing from 750 to 1250 symbols gains just one more dB.

C. Confusion matrices

To give a better idea about the internal behavior of the cumulant based classifier, Table II shows the confusion matrices for SNR = -3dB, SNR = 0 dB and SNR = 3dB. The confusion matrix displays the probability of classification of a modulation scheme $\hat{M} = M_i$ given that the transmitter employs the modulation scheme M_k, i.e. $P(\hat{M} = M_i|M_k)$.

If the received signal is very noisy, the combined cumulant feature of the signal becomes large. Hence, the classifier usually decides for BPSK. The better the SNR becomes, the smaller are the values for $\hat{K}$ and hence the probability first of QPSK detection and then of 16-QAM detection increases. The classification of 8-PSK is the most difficult case for the proposed cumulant-based classifier since its cumulant feature is ideally equal to zero.

VI. CONCLUSION

A novel cumulant-based classifier for automatic modulation classification for MIMO systems is proposed. Both the cases with and without CSI are considered. Simulation results show that, in the case of perfect CSI, the proposed classifier performs marginally better with MMSE equalization than with ZF. In the case without CSI the equalization is performed blindly using JADE, which is a well known ICA algorithm. Interestingly, the performance of the classifier with blind equalization using ICA is almost the same as in the case with perfect CSI using MMSE.

The classification performance of the cumulant-based classifier is about 5 dB worse compared to the optimal ALRT solution at a $P_{CC} = 0.9$. However, the computational complexity is reduced to only a fraction of that of the ALRT classifier.

A possible extension of the described cumulant-based classifier is to use a feature vector containing additional higher-

order cumulant features, such as sixth or eighth-order cumulants. This should improve the classification performance, but increase the computational complexity, which will, however, still remain lower than the complexity of the ALRT.

If other space-time codes are used and the receiver has knowledge of the employed space time codes, this space-time redundancy can also be exploited to further increase the probability of correct classification.

REFERENCES

[1] O. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137 –156, April 2007.
[2] F. Jondral, "Automatic classification of high frequency signals," *Signal Processing*, vol. 9, pp. 177–190, 1985.
[3] A. Swami and B. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 416 –429, Mar. 2000.
[4] C. Spooner, "Classification of co-channel communication signals using cyclic cumulants," in *Proc. IEEE Asilomar*, 1995, pp. 531 –536.
[5] ——, "On the utility of sixth-order cyclic cumulants for rf signal classification," in *Proc. IEEE Asilomar*, 2001, pp. 890 –897.
[6] O. Dobre, Y. Bar-Ness, and W. Su, "Higher-order cyclic cumulants for high order modulation classification," in *IEEE Milcom*, 2003, pp. 112 – 117.
[7] V. Choqueuse, S. Azou, K. Yao, L. Collin, and G. Burel, "Blind modulation recognition for mimo systems," *MTA Review*, vol. 19, pp. 183–196, 2009.
[8] K. Hassan, C. Nzandza, M. Berbineau, W. Hamouda, and I. Dayoub, "Blind modulation identification for mimo systems," in *Proc. IEEE GLOBECOM*, 2010, pp. 1 –5.
[9] W. Gardner and C. Spooner, "The cumulant theory of cyclostationary time-series, part 1: Foundation," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3387 –3408, Dec. 1994.
[10] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.
[11] J. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," *IEE Proceedings on Radar and Signal Processing,*, vol. 140, no. 6, pp. 362 –370, Dec. 1993.