

Automatic Modulation Classification of Real Signals in AWGN Channel Based on Sixth-Order Cumulants

Mirko SIMIC¹, Milos STANKOVIC², Vladimir D. ORLIC¹

¹ Institute Vlatacom, Blvd M Milankovic 5, 11000 Belgrade, Serbia

² Singidunum University, Danijelova 32, 11000 Belgrade, Serbia

{mirko.simic, vladimir.orlic}@vlatacom.com, milos.stankovic@singidunum.ac.rs

Submitted October 16, 2020 / Accepted February 24, 2021

Abstract. *Automatic modulation classification (AMC) represents an important integral part of modern communication systems. While novel AMC algorithms based on complex neural network structures showed significant performance improvements, in practical applications low algorithm complexity of AMC algorithms based on higher-order cumulants still makes them very attractive. AMC algorithm based on sixth-order cumulants showed very good performance in this context, especially when it comes to distinguishing Binary Phase Shift Keying (BPSK) signals from complex constellations. Still, no further analysis of expected performance with other real constellations was presented for this algorithm so far. In this paper, the performance was explored in a wider context of real signals classification, by observing various Pulse Amplitude Modulation (PAM) constellations, whose statistical features are presented for the first time. Their classification performance was tested via Monte Carlo simulations, and explained through the presence of bias under conditions of strong additive white Gaussian noise channel, reported in this paper for real signals for the first time. One new approach in AMC is proposed, which ensures improvement in the classification of real signal constellations. Achieved improvement is confirmed in many Monte Carlo experiments, where the proposed new AMC scheme is tested versus the most popular standard higher-order cumulants-based algorithms.*

Keywords

Automatic modulation classification, feature-based, cumulants, pulse amplitude modulation, real constellations, AWGN

1. Introduction

Automatic modulation classification (AMC) is a technique commonly connected with wireless systems and applications, standing for modulation format recognition process and further demodulation of a priori unknown signals at the receiver side. It has found importance in

electronic warfare, surveillance, and countermeasures, but also in many civilian communication applications, like software-defined radio, spectrum management, cognitive radio, smart reconfigurable transceivers, intelligent modems, and Internet of Things (IoT), [1–5].

There are two approaches in AMC algorithms development widely used by many authors: likelihood-based (LB) methods, which are leading to optimal solutions at the price of high computations, and feature-based (FB) methods, characterized with lower computational complexity, but still able to show performance close to optimal when properly designed [6]. FB algorithms are based on pattern recognition approach, extracting various instantaneous features of the received communication signal, and very frequently using complex classifiers for extended performance, such as neural networks [7], or deep-learning methods [3], [8]. Many modern algorithms perform extraction of several different features simultaneously and combine their classification properties.

Higher-order statistics – cumulants and moments, are especially popular as features of interest in FB AMC schemes. Simple structures of fourth-order cumulants [9] or sixth-order cumulants [10] are still considered as the state-of-the-art of AMC [11]. Sixth-order cumulants showed significantly better performance than fourth-order cumulants [12]. They also show to be superior in low complexity, memory requirements, and inference time when compared with other up-to-date AMC algorithms, like neural networks (with difference measured in several orders of magnitude) – but requiring additional performance improvements to remain competitive with those algorithms [5]. Published research of cumulant-based AMC algorithms are mostly focused on complex signal constellations, like various Quadrature Amplitude Modulation (QAM) schemes. When it comes to real signal constellations, those are quite rarely considered in research, with some exceptions: Binary Phase Shift Keying (BPSK) signals are included in many research works; Pulse Amplitude Modulation (PAM) formats from PAM-4 to PAM-64 were considered in classical work of Swami [13] under fourth-order cumulants AMC, and also in [2] under complex network classifier; PAM-4 signals were considered in [14]

under deep learning aided classification method, and [15] under fusion neural network classifier; BPSK and PAM-4 in [16] under LB classifier, and [17] under Convolution Neural Network (CNN) classifier; PAM-4 and PAM-8 in [18] under CNN classifier, and [19] under multi-layer support vector machine classifier.

While none of the research published so far considers a wider set of PAM signals classification under sixth-order cumulants (or other, higher-order cumulant structures), these constellations are still being of interest in modern communication systems, like IEEE 802.3 gigabit Ethernet [20–22], digital television [23], gigabit optical fiber links [24], [25], or free-space optical systems [26]. PAM signal constellations are presented in Fig. 1.

In this paper we discuss the classification performance of classical sixth-order cumulants-based AMC algorithm for a wide set of PAM constellations, present their error variances and parameters for efficiency in distinguishing particular constellations, for the first time. These parameters are compared with those corresponding with fourth-order cumulants-based AMC algorithm, under the same set of PAM constellations, and tested in Monte Carlo simulations. Very good performance of sixth-order cumulants in distinguishing PAM constellations from complex signal constellations is further explained with the presence of bias under strong AWGN conditions, which is also reported for the first time. A novel AMC scheme is then proposed, which provides both: good performance in distinguishing PAM constellations from complex signal constellations, and mutual distinguishing of various PAM constellations, simultaneously. This novel scheme is based on the convenient manipulation over the basic formula for the real signal cumulant value estimation (unlike the one based on the modulation order reduction [5], which is not applicable to real signals due to their properties presented in this paper). In comparison with complex state-of-the-art AMC algorithms, our resulting solution remains superior in low complexity, memory requirements, and inference time.

The rest of the paper is organized as follows: in Sec. 2 we present standard sixth-order cumulant-based AMC algo-

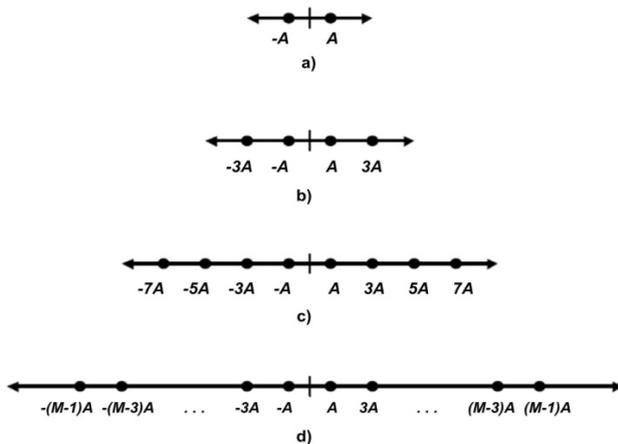


Fig. 1. PAM signal constellations: a) PAM-2, b) PAM-4, c) PAM-8, d) general PAM-M.

rithm and its statistics calculated for PAM constellations. Performance tests of standard cumulant-based algorithm in distinguishing real from complex constellations is presented in Sec. 3, while in Sec. 4 we propose our novel two-stage AMC scheme. This novel scheme was tested via simulations presented in Sec. 5, along with analysis of achieved results. Corresponding conclusions are given in Sec. 6.

2. AMC Algorithm Based on Sixth-Order Cumulants

In a standard communication system model where the received signal is corrupted by AWGN only during the propagation (as explained in [5], for instance), the received signal sequence $y(n)$ can be represented by:

$$y(n) = x(n) + g(n) \tag{1}$$

where $x(n)$ stands for transmitted symbols of an unknown modulation, and $g(n)$ represents AWGN with a zero mean and variance of σ_g^2 . For zero-mean random variable x , associated with transmitted data sequence $x(n)$, the second-order cumulant $C_{21,x} = \text{cum}(x, x^*)$ is given by:

$$C_{21,x} = E(|x|^2). \tag{2}$$

The sixth-order cumulant of the same random variable x $C_{63,x} = \text{cum}(x, x, x, x^*, x^*, x^*)$ is given by [10]:

$$C_{63,x} = E(|x|^6) - 9E(|x|^4)E(|x|^2) + 12|E(x^2)|^2 E(|x|^2) + 12E^3(|x|^2) \tag{3}$$

while the self-normalized sixth-order cumulant is derived from the previous two equations as:

$$\hat{C}_{63,x} = C_{63,x} / (C_{21,x})^3. \tag{4}$$

The cumulants of random variable y , associated with received sequence $y(n)$, can further be expressed in the following manner:

$$C_{63,y} = C_{63,x} \tag{5}$$

$$C_{21,y} = C_{21,x} + \sigma_g^2. \tag{6}$$

Consequently, the following relation is established between the cumulants' values:

$$\hat{C}_{63,x} = \frac{C_{63,y}}{(C_{21,y} - \sigma_g^2)^3}. \tag{7}$$

While the noise power σ_g^2 can be measured at the receiving point, the calculation of cumulants of a received signal in practice is executed via the calculation of mean-values over an ensemble of collected signal samples, which is implementable quite easily. If the number of samples is represented with N , equation (7) in its practical realization takes the form of:

$$\begin{aligned} \tilde{C}_{63,x} = & \left[\frac{1}{N} \sum_{n=1}^N |y(n)|^6 - 9 \left(\frac{1}{N} \sum_{n=1}^N |y(n)|^4 \cdot \frac{1}{N} \sum_{n=1}^N |y(n)|^2 \right) \right. \\ & + 12 \left(\frac{1}{N} \sum_{n=1}^N |y(n)|^2 \right)^2 \cdot \frac{1}{N} \sum_{n=1}^N |y(n)|^2 \left. \right] \quad (8) \\ & + 12 \left(\frac{1}{N} \sum_{n=1}^N |y(n)|^2 \right)^3 \cdot \frac{1}{\left(\frac{1}{N} \sum_{n=1}^N |y(n)|^2 - \sigma_g^2 \right)^3}. \end{aligned}$$

In practice, estimated values of normalized cumulants are always shaped with some portion of dispersion around expected (theoretical) values. This phenomenon was explored and described originally in [13], for fourth-order cumulants, which are calculated as:

$$C_{42,x} = E(|x^4|) - |E(x^2)|^2 - 2E^2(|x^2|), \quad (9)$$

$$\hat{C}_{42,x} = C_{42,x} / (C_{21,x})^2 \quad (10)$$

where the variance of the sample estimates of $C_{42,x}$ for complex constellations, with N samples, is given by:

$$\begin{aligned} N \text{ var}(C_{42,x}) = & \\ [m_{8,4} - m_{4,2}^2] + 4m_{2,1}[3m_{4,2}m_{2,1} - 2m_{6,3} + 2m_{2,1}^3], & \quad (11) \end{aligned}$$

with $m_{k,m} = E[y^{k-m}(y^*)^m]$. For real constellations, the variance of the sample estimates of $C_{42,x}$ in [13] is given by:

$$\begin{aligned} N \text{ var}(C_{42,x}) = & \\ [m_{8,4} - m_{4,2}^2] + 6m_{2,1}[5m_{4,2}m_{2,1} - 2m_{6,3} + 3m_{2,1}^3]. & \quad (12) \end{aligned}$$

The similar conclusion was reported for sixth-order cumulants in [27], where the variance of $C_{63,x}$ for complex signals with N samples, is given with:

$$\begin{aligned} N \text{ var}(C_{63,x}) = & [m_{12,6} - m_{6,3}^2] + 9[m_{2,1}^2(48m_{4,2}m_{2,1}^2 - 54m_{4,2}^2 \\ & + 96m_{2,1}^4 - 64m_{6,3}m_{2,1}) + m_{4,2}(9m_{4,2}^2 + 16m_{6,3}m_{2,1} \\ & - 2m_{8,4}) + m_{2,1}(17m_{8,4}m_{2,1} - 2m_{10,5})] \quad (13) \end{aligned}$$

while for real constellations, this variance is given by:

$$\begin{aligned} N \text{ var}(C_{63,x}) = & [m_{12,6} - m_{6,3}^2] + 9[m_{2,1}^2(384m_{4,2}m_{2,1}^2 - 126m_{4,2}^2 \\ & + 384m_{2,1}^4 - 128m_{6,3}m_{2,1}) + m_{4,2}(9m_{4,2}^2 + 16m_{6,3}m_{2,1} \\ & - 2m_{8,4}) + m_{2,1}(25m_{8,4}m_{2,1} - 2m_{10,5})]. \quad (14) \end{aligned}$$

Reported error variances are directly proportional with sample size N and take different values for different modulation formats. While variances of the sample estimates of fourth-order cumulants for various PAM constellations were published in [13], we calculate variances of the sample estimates $C_{63,x}$ for the first time in this paper. In Tab. 1 the theoretic cumulant values and corresponding variances for various modulation constellations are shown.

The decision-making process for modulation recognition is based on a comparison of obtained values of normalized cumulant estimates with predefined thresholds. It was shown, on the basis of intensive computer simulations,

Constellation	\hat{C}_{63}	$N \text{ var}(C_{63,x})$	\hat{C}_{42}	$N \text{ var}(C_{42,x})$
BPSK	16.0000	5040.00	-2.0000	36.00
PAM-4	12.1600	4612.00	-1.3600	34.72
PAM-8	11.7600	4455.00	-1.2381	32.27
PAM-16	11.6817	4419.70	-1.2094	31.67
PAM-32	11.6632	4411.00	-1.2023	31.52
PAM-64	11.6587	4409.00	-1.2006	31.49
QPSK	4.0000	576.00	-1.0000	12.00
QAM-16	2.0800	332.31	-0.6800	9.54
QAM-64	1.7970	289.38	-0.6190	8.82

Tab. 1. Theoretical cumulant statistics for various constellation types, and variances of their sample estimates.

that optimal comparison threshold values are positioned at the middle of intervals between expected (theoretical) values corresponding with particular modulation formats [5].

3. Performance in Distinguishing Real Constellations from Complex Constellations

In order to compare the expected classification performance of fourth and sixth-order cumulants-based algorithms, in [27] a ratio of standard deviation and mutual distance of nearby values was used, for both $C_{42,x}$ and $C_{63,x}$, as a measure of the algorithm's selectivity. Based on this parameter's value, distinguishing BPSK from QPSK was reported to be better with sixth-order cumulants criteria, and confirmed with the simulations through 2000 Monte Carlo trials and $N=250$ received data samples collected for AMC in each trial. An algorithm with $C_{63,x}$ features was simulated along with an algorithm based on $C_{42,x}$ features in non-dispersive channel conditions (AWGN only), and with noise power σ_g^2 considered to be known. Correct classification probability P_{CC} was calculated versus SNR , with two scenarios of modulation candidates considered from the set {BPSK, QPSK}. These results are presented in Fig. 2.

From Tab. 1 it can be clearly noted that cumulant values of QPSK signals, among all complex constellations, are the closest to cumulant values of all considered real constellations. Thus, the ability of both fourth- and sixth- order cumulants-based algorithm to distinguish a particular real constellation from complex signals can be evaluated through

No. of simulation trials	2000
Data samples within each trial	$N=250$
Channel model	AWGN
Modulation candidates	QPSK, PAM-4, PAM-8, PAM-16, PAM-64
Simulation software	MATLAB R2015a
Hardware platform	i3-5005 CPU, 2.00 GHz

Tab. 2. The main simulation parameters.

the ability of distinction from QPSK, as the border case. By calculating the values of the same comparison parameter (a ratio of standard deviation and mutual distance of cumulant's theoretical values) for PAM signals, the same conclusion of better classification with sixth-order cumulants criteria can be easily established. To test this expectation, we carry out the simulations through 2000 Monte Carlo trials and $N=250$ received data samples collected for AMC in each trial. All simulations are executed in the Matlab software package, with the main simulation parameters presented in Tab. 2.

Again, the algorithm with $C_{63,x}$ features is simulated along with an algorithm based on features $C_{42,x}$ in AWGN channel conditions, with noise power considered to be known. Correct classification probability was calculated versus SNR, under five scenarios of different sets of modulation candidates considered: (i) {PAM-4, QPSK}, (ii) {PAM-8, QPSK}, (iii) {PAM-16, QPSK}, (iv) {PAM-32, QPSK} and (v) {PAM-64, QPSK}. Simulation results for scenario (i) are presented in Fig. 3, results for scenario (ii) are shown in Fig. 4, for scenario (iii) in Fig. 5, scenario (iv) in Fig. 6, while results for scenario (v) are presented in Fig. 7. The simulation results are essentially in agreement with theoretical arguments: the algorithm based on $C_{63,x}$ features shows better performance than the one based on $C_{42,x}$ features, in all cases considered.

Furthermore, results could be even commented as superior in the case of a sixth-order cumulants-based algorithm. Careful inspection of estimated $\hat{C}_{63,x}$ values achieved under low SNR conditions leads to one very interesting conclusion: a presence of bias in estimated $\hat{C}_{63,x}$ values for PAM constellations can be noted, which increases these estimates' values, leading to more accurate classification. While a similar effect was reported for BPSK signals in multipath channel conditions in [10], for PAM signals in AWGN it is reported here for the first time. This effect is illustrated in Fig. 8, where resulting histograms of sixth-order cumulant values estimated in Monte Carlo trials are presented at SNR = 20 dB and SNR = 10 dB, for scenario (v).

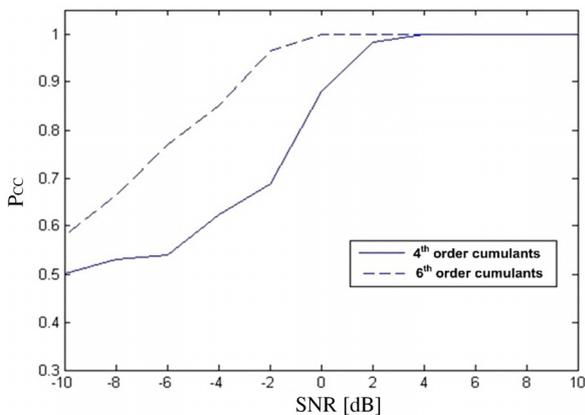


Fig. 2. Correct classification probability in {BPSK, QPSK} scenario, $N=250$, [27].

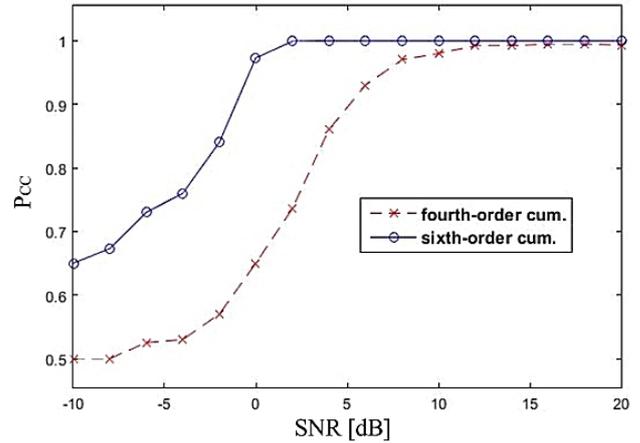


Fig. 3. Correct classification probability in {PAM-4, QPSK} scenario, $N=250$.

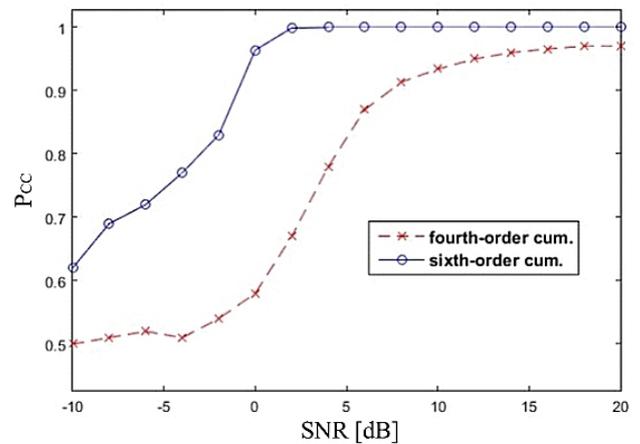


Fig. 4. Correct classification probability in {PAM-8, QPSK} scenario, $N=250$.

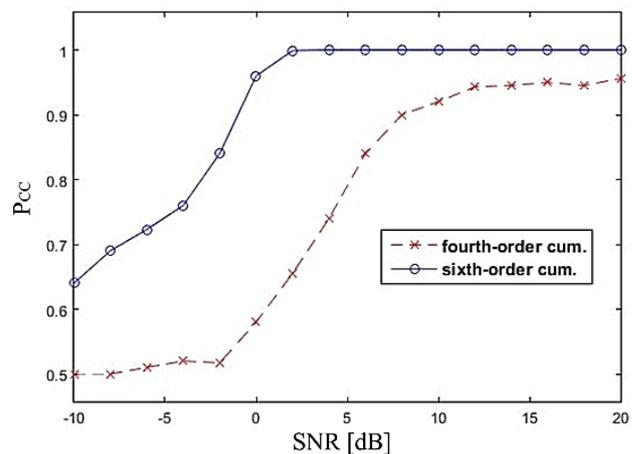


Fig. 5. Correct classification probability in {PAM-16, QPSK} scenario, $N=250$.

As it can be noticed from Fig. 8, while cumulant estimates of complex (QPSK) constellation are strictly unbiased, lower SNR values introduce stronger bias for PAM-64 signal's cumulant estimates. This effect is present for all PAM constellations simulated and represents the reason for the excellent classification performances presented above. Since both theoretical cumulant values and

their error variances for all considered PAM constellations are mutually very similar (Tab. 1), distinguishing from complex (QPSK) signals in all particular simulations comes with approximately the same superior performance, being errorless at all $SNR > 2$ dB values, and providing improved P_{CC} values even at very low SNR . Approximately the same error variances of PAM signals are being compensated with an approximately same bias of mutually close theoretical values, thus resulting in approximately the same AMC performance.

An explanation for this effect can be found in the structure of the used $C_{63,x}$ formula (3). While that formula, originally proposed in [10], is being used by some authors, other authors also report and use different structures for $C_{63,x}$ in their work: for example, in [3, 5, 18, 28] all mutually different formulas were used. In fact, these formulas mostly represent an approximation, simplified forms of the theoretical cumulant formula, which are correct for some constellations but not in the general case.

The exact sixth-order cumulant formula of random variable x can be expressed from the joint cumulant formula [29]:

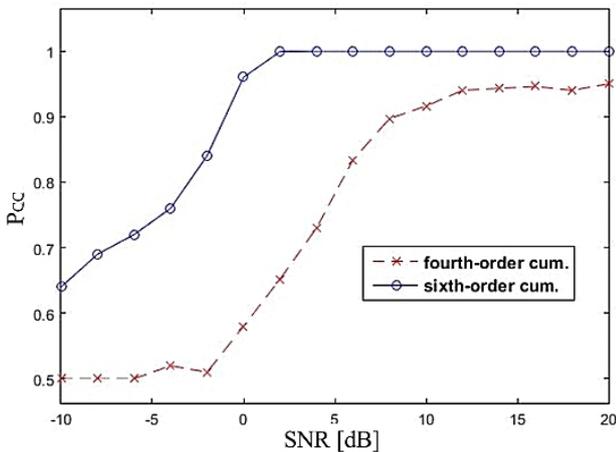


Fig. 6. Correct classification probability in {PAM-32, QPSK} scenario, $N = 250$.

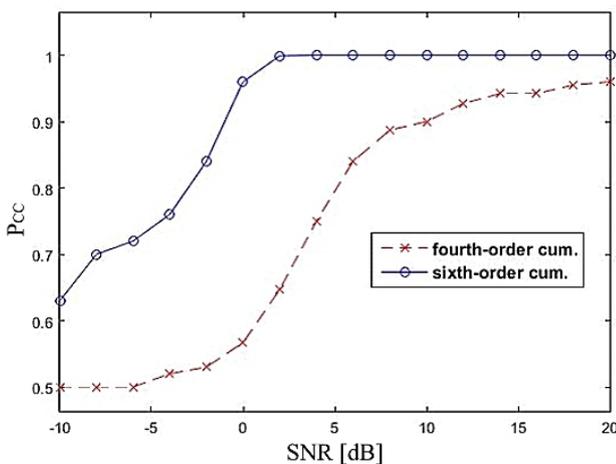


Fig. 7. Correct classification probability in {PAM-64, QPSK} scenario, $N = 250$.

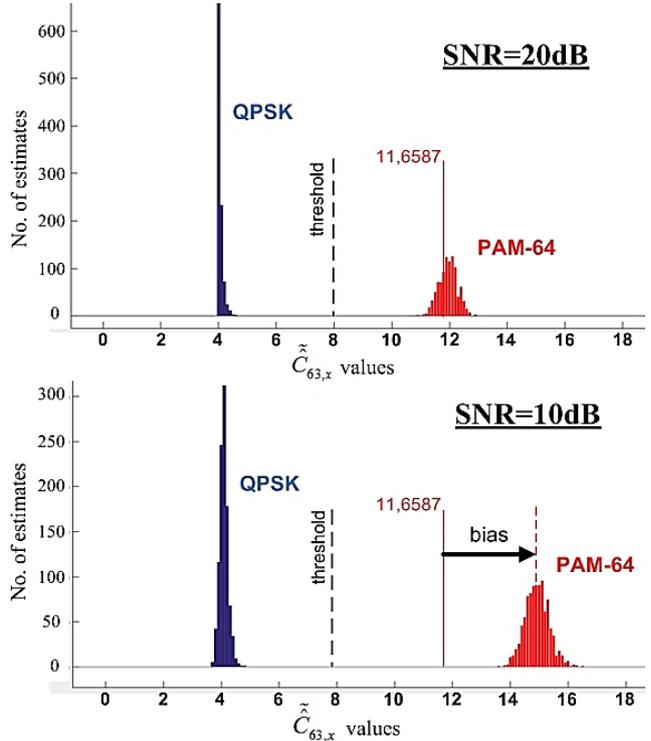


Fig. 8. Resulting histogram of sixth-order cumulant values estimated in Monte Carlo trials at $SNR = 20$ dB (up) and $SNR = 10$ dB (down), corresponding with PAM-64 (red) and QPSK (blue) transmitted signal.

$$\text{cum}(x_1, \dots, x_n) = \sum_{\pi} (|\pi| - 1)! (-1)^{|\pi| - 1} \prod_{B \in \pi} E(\prod_{i \in B} x_i) \quad (15)$$

where π runs through the list of all partitions of $\{1, \dots, n\}$, and B runs through the list of all blocks of the partition π . For zero-mean random variable x , strict developing of all members in the sum of (15) for the $\text{cum}(x, x, x, x^*, x^*, x^*)$ case, leads to the following formula derived here for general and unbiased cumulant:

$$\begin{aligned} C_{63,x_UNB} &= E(|x|^6) - 9E(|x|^4)E(|x|^2) \\ &\quad + 18|E(x^2)|^2 E(|x|^2) - 6|E(x^2)| E(x^2|x^2) \quad (16) \\ &\quad + 12E^3(|x|^2). \end{aligned}$$

For complex signals, $E(x^2) = 0$ stands and equation (16) gets reduced to the form of (3). Similarly, other forms used by authors who focused on applications with complex signals only are derived. However, it is of interest to cover applications of real signals as well. When equation (16) is used instead of (3) in the AMC algorithm given by equations (4)–(8), the appearance of bias in estimated \hat{C}_{63,x_UNB} values for PAM constellations under low SNR conditions are avoided, but error estimates and mean values of cumulants are much less convenient for AMC application. This is illustrated in Fig. 9, where the situation equivalent to the one presented in Fig. 8 is shown, now corresponding with the application of C_{63,x_UNB} instead of $C_{63,x}$.

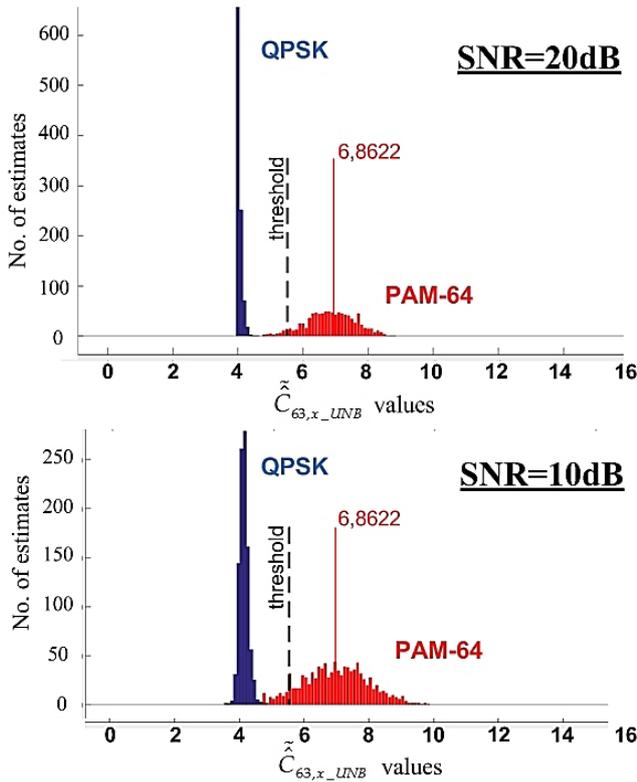


Fig. 9. Resulting histogram of sixth-order cumulant values estimated in Monte Carlo trials at $SNR = 20$ dB (up) and $SNR = 10$ dB (down), corresponding with PAM-64 (red) and QPSK (blue) transmitted signal.

Feature	BPSK	PAM-4	PAM-8	PAM-16	PAM-32	PAM-64
\hat{C}_{63_UNB}	16.0000	8.3200	7.1899	6.9381	6.8733	6.8622

Tab. 3. Theoretical unbiased sixth-order cumulants of real constellations.

While for complex constellations data presented in Tab. 1 are also standing with C_{63,x_UNB} , in Tab. 3 the theoretic C_{63,x_UNB} values of considered real constellations are shown, since they're different.

Although the performance in distinguishing real constellations from QPSK (and other complex constellations) is clearly better with $C_{63,x}$ formula, the presence of reported bias makes mutual distinguishing among various PAM constellations practically impossible with this feature: all PAM constellations are affected with bias, so using fixed threshold values based on theoretical cumulants for decision - making in AMC is pointless. On the other side, the structure C_{63,x_UNB} is relaxed from this issue. Thus, it would be of interest to inspect possibilities for a solution which will benefit from both: 1) bias of $C_{63,x}$ which provides superior separation of real signals from complex signals, and 2) unbiased C_{63,x_UNB} structure for the mutual separation of real constellations.

4. Novel Two-Stage AMC Scheme

The formula given in (16) can be rewritten in the following manner:

$$C_{63,x_UNB} = C_{63,x} + \underbrace{6|E(x^2)|^2 E(|x|^2) - 6|E(x^2)| E(x^2|x^2)}_{O_{63,x}} \quad (17)$$

where $C_{63,x}$ is given by (3), whose simple expansion with the values of offset $O_{63,x}$ results in the conversion of biased into unbiased sixth-order cumulants. In practical implementation normalized unbiased sixth-order cumulant can be expressed as:

$$\tilde{C}_{63,x_UNB} = \tilde{C}_{63,x} \left\{ \frac{+ \left[6 \left(\frac{1}{N} \sum_{n=1}^N y^2(n) \right)^2 \cdot \frac{1}{N} \sum_{n=1}^N |y(n)|^2 \right] - 6 \left(\frac{1}{N} \sum_{n=1}^N y^2(n) \right) \cdot \frac{1}{N} \sum_{n=1}^N y^2(n) |y(n)|^2 \right]}{\left(\frac{1}{N} \sum_{n=1}^N |y(n)|^2 - \sigma_g^2 \right)^3} \right\} \tilde{O}_{63,x} \quad (18)$$

where $\tilde{C}_{63,x}$ is given by (8). From (18) novel two-stage scheme AMC can be derived, as presented in Fig. 10, which is based on exploiting the benefits of both classical (biased) and general (unbiased) sixth-order cumulant structures.

As it can be noticed from Fig. 10, in the proposed AMC algorithm, within Stage 1, classical sixth-order cumulants of the received signal are estimated and used for distinguishing between real and complex constellations. In this manner, all the excellent properties reported in Sec. 3 of this paper are exploited, and very good classification performance in this step is provided. As a comparison threshold, the “middle of the interval” value between QPSK and PAM-64 (Tab. 1) was used, since these two constellations represent the border cases of two constellation clusters.

If the constellation is initially recognized as belonging to a cluster of real signals, Stage 2 follows where estimates of sixth-order cumulants are first converted into unbiased estimates form by adding the value of estimated offset $\tilde{O}_{63,x}$, and then decision-tree-based recognition of exact real constellation is performed. Comparison thresholds within Stage 2 are calculated as the “middle of the interval” values between neighboring theoretical unbiased estimates of PAM (and BPSK) constellations, presented in Tab. 3. In this manner, AMC of particular real constellations from the wider constellation set is achieved, which is feasible only with unbiased cumulants.

When compared with the structure of standard sixth-order cumulants AMC algorithm, the one presented in Fig. 10 differs numerically only in the estimation of $\tilde{O}_{63,x}$ i.e. calculation of \tilde{C}_{63,x_UNB} (and correspondingly changed values for comparison when deciding real signal constella-

tion). Thus, the only added complexity comes from the $\tilde{\hat{O}}_{63,x}$ estimation. The scheme in Fig. 10 and its corresponding equations (17)–(18) are proposed here for the first time.

At this point it is of interest to highlight the difference between the novel AMC algorithm proposed here and the one described in [5]: although both solutions provide im-

provements in cumulant-based recognition via two-step procedure, their nature and applicability is completely different. Algorithm described in [5] represents an ad-hoc solution for QAM constellations recognition, based on modulation order reduction. Important precondition for its application lays in absence of any bias in cumulants' estimates. On the other hand, the presence of strong bias reported for PAM signals in this work makes the solution [5] being useless in classification of PAM constellations. Instead of any manipulations over the signal structure, here we propose the method based on smart switching between the fundamental formulas for sixth-order cumulants calculation, dedicated for PAM signals exactly, as described previously in this section. It should be also noted that neither the algorithm proposed in this paper makes any improvement in the classification of QAM signals, meaning that two methods remain contributing to completely different classes of signal constellations.

5. Simulations and Performance Analysis

We carried out the simulations through 2000 Monte Carlo trials and $N = 250$ received data samples were collected for AMC in each trial, where the algorithm proposed in Sec. 4 was tested in AWGN channel conditions, in scenarios with modulation candidates considered from the set: (i) {QPSK, PAM-4, BPSK}, (ii) {QPSK, PAM-8, PAM-4, BPSK}, and (iii) {QPSK, PAM-64, PAM-32, PAM-16, PAM-8, PAM-4, BPSK}. Modulation candidates are selected in order to demonstrate performance in both distinguishing real from complex constellations (i.e. from QPSK as the border case of complex constellations) and in mutual distinguishing among different real constellations within the same scenario, with a gradual increase in scenarios' complexity. Similarly as explained previously in Sec. 3, all simulations are executed in the Matlab software package, using the package's built-in functions for modeling the AWGN channel, with the main simulations' hardware and software parameters presented in Tab. 2. Simulations code is open for considerations and use of other researchers, and can be found at [30].

In order to provide a fair comparison of achieved performance with other comparable algorithms, the AMC algorithm based on standard fourth-order cumulants [9] was simulated as well, under the same set of modulation candidates and sample size $N = 250$, along with the AMC algorithm based on standard (biased) sixth-order cumulants [10]. The value of N was selected to match directly with the one used in simulations described in [9] and represents the main controlling parameter from the aspect of AMC performance. It should be also noted that the algorithm [9] is unbiased in its nature (i.e. for all simulated signal constellations). The AWGN channel was simulated with noise power σ_g^2 considered to be known. For each particular SNR value, and within every particular Monte Carlo trial, the same set of N samples was processed by all simulated algo-

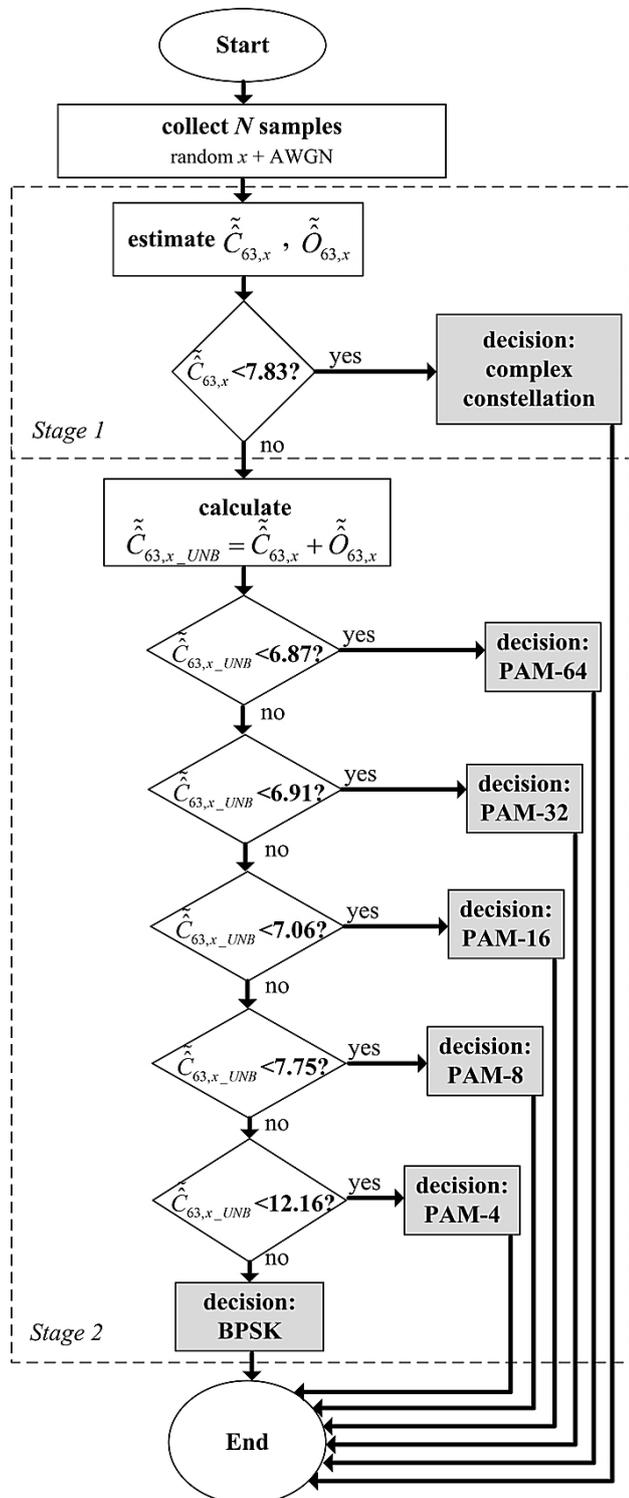


Fig. 10. Diagram of novel two-stage AMC algorithm based on sixth-order cumulants.

rithms, thus providing the fair comparison under the exactly same channel conditions and over the exactly same input data.

Comparison thresholds for sixth-order cumulants AMC algorithm are having the values as previously described in this paper, while comparison thresholds for the algorithm [9] were selected as the “middle of interval” values between the theoretical fourth-order cumulants of considered modulation formats (Tab. 1).

Correct classification probability P_{CC} was calculated versus SNR , and Figure 11 illustrates the results of simulation in scenario (i); results in scenario (ii) are presented in Fig. 12, while Figure 13 illustrates the results of simulation in scenario (iii).

As can be confirmed from Fig. 11 and Fig. 12, the proposed novel two-stage AMC algorithm based on sixth-order cumulants shows better performance than the algorithm based on standard fourth-order cumulants, and this difference in performance is more significant as the SNR value is getting lower. Performance of standard sixth-order cumulants is generally poor, due to the presence of bias, as described in this paper, but its robustness in the classification

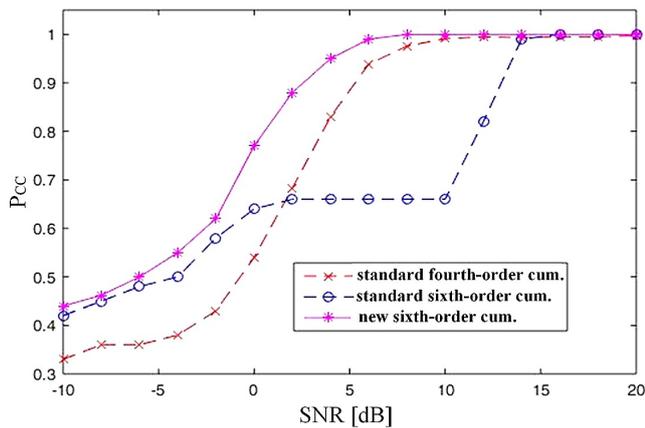


Fig. 11. Correct classification probability in {QPSK, PAM-4, BPSK} scenario, $N = 250$.

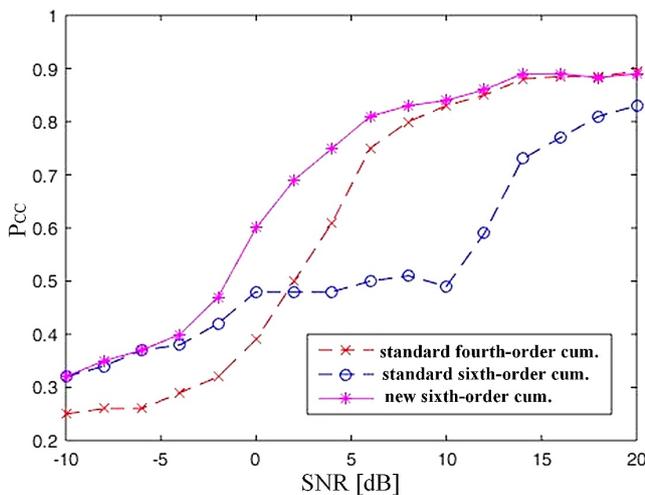


Fig. 12. Correct classification probability in {QPSK, PAM-8, PAM-4, BPSK} scenario, $N = 250$.

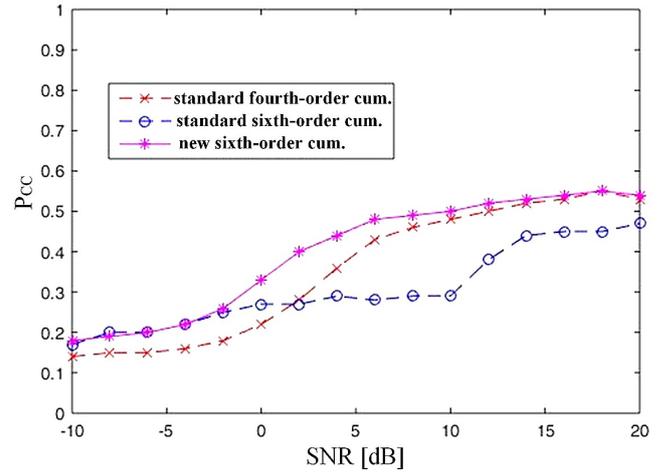


Fig. 13. Correct classification probability in {QPSK, PAM-64, PAM-32, PAM-16, PAM-8, PAM-4, BPSK} scenario, $N = 250$.

of complex (QPSK) signal remains obvious even at low SNR . Achieved classification performance of the proposed algorithm comes mainly from this, well-exploited, excellent classification of complex (QPSK) signals, while moderate performance in recognition of real constellations is provided at the same time - very similar to the one given by the standard fourth-order cumulants. Also, it should be noted that as the number of constellations under test rises (four constellations in scenario (ii), in comparison with three constellations in scenario (i)) general performance of all simulated algorithms gets lower in terms of P_{CC} . This is especially illustrated in Fig. 13, where a wide set of seven constellations was considered in simulation, i.e. all real constellations discussed in this paper, along with QPSK.

From Fig. 13 it can be noted that at $SNR = 20$ dB the most competitive algorithms under the test of seven constellations show the performance of around $P_{CC} = 0.55$, while for the same SNR value in a scenario with four constellations (Fig. 12) their performance was approximately $P_{CC} = 0.85$. Nevertheless, Figure 13 confirms again that the proposed novel AMC algorithm based on sixth-order cumulants outperforms the classical algorithms based on higher-order cumulants, even in the problem with the widest set of modulation candidates.

In order to explore the impact of sample size N on reported classification performance, we repeat simulations in scenario (iii), now with $N = 2000$ received data samples for AMC. The results are presented in Fig. 14.

As can be concluded from Fig. 14, bigger sample size N directly contributes to the classification performance of algorithms under test: at $SNR = 20$ dB the most competitive algorithms reach the value of $P_{CC} = 0.65$ with $N = 2000$ samples, which represents an obvious improvement in comparison with the same scenario and $N = 250$ samples from Fig. 13. Again, the achieved classification performance of the proposed new AMC algorithm based on sixth-order cumulants is confirmed as better than of the classical algorithm based on fourth-order cumulants. The

algorithm based on standard sixth-order cumulant structure is outperformed by the others in all considered scenarios.

Reported performance stands in direct correspondence with the ability of a particular algorithm to distinguish complex from real signals effectively. Simulations showed that distinguishing complex (QPSK) signals from real constellations is errorless with fourth-order cumulants and $N=250$ samples at $SNR > 7$ dB, while with sixth-order cumulants and the same N value distinguishing complex from real constellations is errorless even around $SNR = 2$ dB. With $N=2000$ samples, fourth-order cumulants achieve QPSK signal classification without any errors at $SNR > 4$ dB, while the proposed sixth-order cumulants algorithm provides an errorless classification of QPSK signals even around the values low as $SNR = -1$ dB. Reported numerical values are stated directly in an observed performance presented in Fig. 15.

Although the only channel disturbance source considered (and simulated) in these experiments is the noise, the nature of cumulants provides that all presented considerations will also stand in the case of the flat-fading channel (this property results directly from the definition of self-

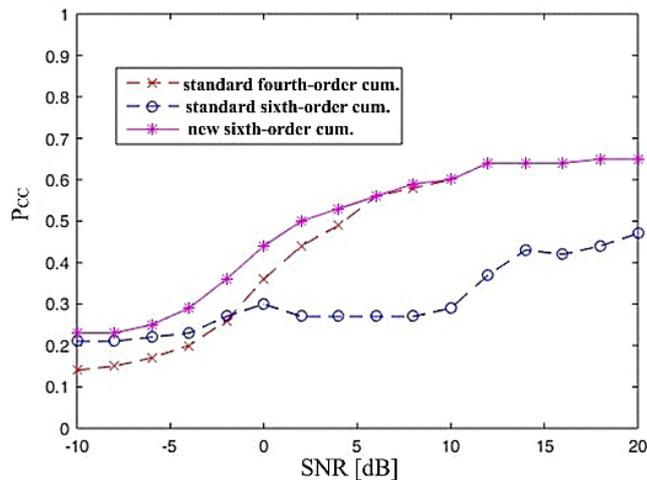


Fig. 14. Correct classification probability in {QPSK, PAM-64, PAM-32, PAM-16, PAM-8, PAM-4, BPSK} scenario, $N = 2000$.

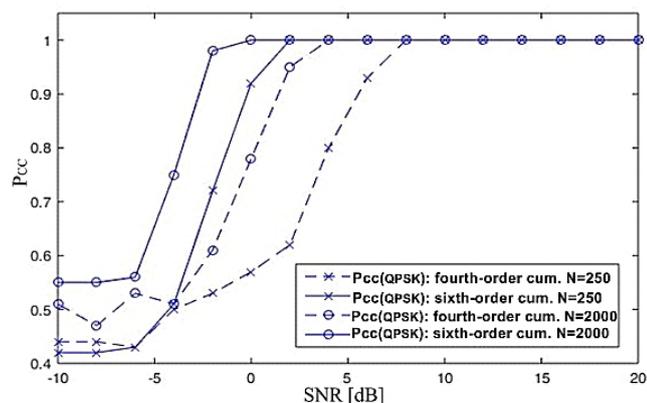


Fig. 15. Correct classification probability for complex (QPSK) signals, in seven – constellation candidates' scenarios.

normalized cumulants, as presented in Sec. 2). For more complex channel models additional channel estimation procedure needs to be introduced in simulated algorithms, like [9], [10], representing one of the most important tasks for future work and considerations of proposed novel two-stage AMC, along with the conditions of interference and other channel imperfections.

Achieved performance improvement with the AMC algorithm proposed in this paper comes at the cost of some added numerical complexity: $(N+1)$ new multiplications and $(N+1)$ new additions are introduced for the calculation of offset and conversion of the standard into the unbiased form of sixth-order cumulants, in comparison with classical sixth-order cumulants, as described with (18) and Fig. 10. Still, those introduce no practical change in the overall complexity of the AMC algorithm, meaning that it remains very competitive in terms of numerical and memory resources, and inference time [5]. Finally, the algorithm proposed in this paper can be expanded further with more complex classifiers (like various neural network structures), in applications where introducing more significant overall computational complexity represents an acceptable cost for providing even higher classification performance.

6. Conclusion

In this paper, a novel scheme for improving the performance of AMC based on sixth-order cumulants, in the form of a simple two-stage feature extraction structure is presented, which exploits benefits from both classical and theoretical (unbiased) sixth-order cumulant structures. Statistical properties of classical sixth-order cumulants of various PAM constellations were analyzed for the first time; their superior performance in distinguishing real from complex constellations was tested in computer-aided simulations, and explained through the presence of bias, which was reported for the first time for this class of signals. This presence of bias was also identified as an obstacle for mutual distinguishing of real constellations with classical sixth-order cumulants. The difference in the structure was explained between standard sixth-order cumulants and theoretical cumulants of the same order, which are relaxed from the presence of bias, thus being suitable for the classification of particular real constellations. This difference was then further used for the implementation of a simple mechanism for conversion between two cumulant structures in practical AMC. The same mechanism was incorporated in the novel two-stage AMC scheme. In computer simulations performance of a new classification algorithm was compared with the performance of a popular AMC algorithms based on standard sixth- and fourth-order cumulants, and better results with the new algorithm were reported in several simulation scenarios. Expectations in the practical use of tested algorithms were also discussed, in the context of SNR and sample size N values. Achieved new performance of the proposed algorithm is coming at

the price of (some) added complexity, which was also discussed, but making no significant impact on the overall complexity of sixth-order cumulant structures application in AMC, thus providing their well-known competitiveness to remain high. The proposed algorithm can be further combined with nowadays popular complex structures for enhanced feature extraction, like neural networks or deep learning methods, for potentially achieving even better performance. For future work, the proposed algorithm should be tested in propagation conditions that are not limited to the presence of noise only, but also consider the effects of multipath propagation, interference, and other real-world conditions.

References

- [1] ELDEMERDASH, Y. A., DOBRE, O. A., ONER, M. Signal identification for multiple-antenna wireless systems: Achievements and challenges. *IEEE Communications Surveys & Tutorials*, 2016, vol. 18, no. 3, p. 1524–1551. DOI: 10.1109/COMST.2016.2519148
- [2] AL-NUAIMI, D. H., HASHIM, I. A., ZAINAL ABIDIN, I. S., et al. Performance of feature-based techniques for automatic digital modulation recognition and classification: A Review. *Electronics*, 2019, vol. 8, no. 12, p. 1–25. DOI: 10.3390/electronics8121407
- [3] ZHANG, T., SHUAI, C., ZHOU, Y. Deep learning for robust automatic modulation recognition method for IoT applications. *IEEE Access*, 2020, vol. 8, p. 117689–117697. DOI: 10.1109/ACCESS.2020.2981130
- [4] BOZOVIC, R., SIMIC, M. Spectrum sensing based on higher order cumulants and kurtosis statistics tests in cognitive radio. *Radioengineering*, 2019, vol. 29, no. 2, p. 464–472. DOI: 10.13164/re.2019.0464
- [5] PAJIC, M. S., VEINOVIC, M., PERIC, M., et al. Modulation order reduction method for improving the performance of AMC algorithm based on sixth-order cumulants. *IEEE Access*, 2020, vol. 8, p. 106386–106394. DOI: 10.1109/ACCESS.2020.3000358
- [6] HAZZA, A., SHOAIB, M., ALSHEBEILI, S. A., et al. An overview of feature-based methods for digital modulation classification. In *2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*. Sharjah (United Arab Emirates), 2013, p. 1–6. DOI: 10.1109/ICCSPA.2013.6487244
- [7] LEE, J. H., KIM, J., KIM, B., et al. Robust automatic modulation classification technique for fading channels via deep neural network. *Entropy*, 2017, vol. 19, no. 9, p. 1–11. DOI: 10.3390/e19090454
- [8] LEE, S. H., KIM, K.-Y., SHIN, Y. Effective feature selection method for deep learning-based automatic modulation classification scheme using higher-order statistics. *Applied Sciences*, 2020, vol. 10, no. 2, p. 1–14. DOI: 10.3390/app10020588
- [9] WU, H., SAQUIB, M., YUN, Z. Novel automatic modulation classification using cumulant features for communications via multipath channels. *IEEE Transactions on Wireless Communications*, 2008, vol. 7, no. 8, p. 3098–3105. DOI: 10.1109/TWC.2008.070015
- [10] ORLIC, V. D., DUKIC, M. L. Automatic modulation classification algorithm using higher-order cumulants under real-world channel conditions. *IEEE Communications Letters*, 2009, vol. 13, no. 12, p. 917–919. DOI: 10.1109/LCOMM.2009.12.091711
- [11] PENNACCHIO, A. A., LUSTOSA DA COSTA, J. P. C., BORDINI, V. M., et al. Eigenfilter-based automatic modulation classification with offsets for distributed antenna systems. In *XXXIV Simposio Brasileiro de telecomunicacoes e processamento*. Sinais, 2016, p. 260–261. DOI: 10.14209/SBRT.2016.8
- [12] LI, X., DONG, F., ZHANG, S., et al. A survey on deep learning techniques in wireless signal recognition. *Wireless Communications and Mobile Computing*, 2019, p. 1–12. DOI: 10.1155/2019/5629572
- [13] SWAMI, A., SADLER, B. M. Hierarchical digital modulation classification using cumulants. *IEEE Transactions on Communications*, 2000, vol. 48, no. 3, p. 416–429. DOI: 10.1109/26.837045
- [14] YANG, C., HE, Z., PENG, Y., et al. Deep learning aided method for automatic modulation recognition. *IEEE Access*, 2019, vol. 7, p. 109063–109068. DOI: 10.1109/ACCESS.2019.2933448
- [15] HUANG, S., LIN, C., ZHOU, K., et al. Identifying physical-layer attacks for IoT security: An automatic modulation classification approach using multi-module fusion neural network. *Physical Communication*, 2020, vol. 43, p. 1–10. DOI: 10.1016/j.phycom.2020.101180
- [16] ZHANG, Y., ANSARI, N., SU, W. Multi-sensor signal fusion-based modulation classification by using wireless sensor networks. *Wireless Communications and Mobile Computing*, 2015, vol. 15, no. 12, p. 1621–1632. DOI: 10.1002/wcm.2450
- [17] TEKBIYIK, K., EKTI, A. R., GORCIN, A., et al. Robust and fast automatic modulation classification with CNN under multipath fading channels. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. Antwerp (Belgium), 2020, p. 1–6. DOI: 10.1109/VTC2020-Spring48590.2020.9128408
- [18] SHI, Y., XU, H., JIANG, L., et al. Few-shot modulation classification method based on feature dimension reduction and pseudo-label training. *IEEE Access*, 2020, vol. 8, p. 140411 to 140425. DOI: 10.1109/ACCESS.2020.3012712
- [19] ZHANG, X., SUN, J., ZHANG, X. Automatic modulation classification based on novel feature extraction algorithms. *IEEE Access*, 2020, vol. 8, p. 16362–16371. DOI: 10.1109/ACCESS.2020.2966019
- [20] IEEE Standard for Ethernet - Amendment 1: Physical Layer Specifications and Management Parameters for 2.5 Gb/s and 5 Gb/s Operation over Backplane. In *IEEE Std 802.3cb-2018 (Amendment to IEEE Std 802.3-2018)*, 2019, p. 1–206, DOI: 10.1109/IEEESTD.2019.8604150
- [21] IEEE Standard for Ethernet - Amendment 4: Physical Layers and Management Parameters for 50Gb/s, 200Gb/s, and 400Gb/s Operation over Single-Mode Fiber. In *IEEE Std 802.3cn-2019 (Amendment to IEEE Std 802.3-2018 as amended by IEEE Std 802.3cb-2018, IEEE Std 802.3bt-2018, and IEEE Std 802.3cd-2018)*, 2019, p. 1–87. DOI: 10.1109/IEEESTD.2019.8937109
- [22] IEEE Standard for Ethernet Amendment 3: Physical Layer and Management Parameters for 25 Gb/s and 40 Gb/s Operation, Types 25GBASE-T and 40GBASE-T. In *IEEE Std 802.3bq-2016 (Amendment to IEEE Std 802.3-2015 as amended by IEEE Std 802.3bw-2015, and IEEE Std 802.3by-2016)*, 2016, p. 1–211. DOI: 10.1109/IEEESTD.2016.7572861
- [23] BASILE, C., CAVALLERANO, A. P., DEISS, M. S., et al., The US HDTV standard. *IEEE Spectrum*, 1995, vol. 32, no. 4, p. 36–45. DOI: 10.1109/6.375998
- [24] LI, X., WEI, J. L., BAMIEDAKIS, N., et al. Avalanche photodiode enhanced PAM-32 5 Gb/s LED-POF link. In *2014 The European Conference on Optical Communication (ECOC)*. Cannes (France), 2014, p. 1–3. DOI: 10.1109/ECOC.2014.6964013

- [25] PENG, L., LIU, M., HELARD, M., et al. PN-PAM scheme for short range optical transmission over SI-POF - an alternative to Discrete Multi-Tone (DMT) scheme. *Journal of the European Optical Society-Rapid Publications*, 2017, vol. 13, no. 1, p. 1–12. DOI: 10.1186/s41476-017-0048-6
- [26] ALNWAIMI, G., BOUJEMAA, H., ARSHAD, K. Optimal packet length for free-space optical communications with average SNR feedback channel. *Journal of Computer Networks and Communications*, 2019, p. 1–8. DOI: 10.1155/2019/4703284
- [27] ORLIC, V. D., DUKIC, M. L. Properties of an algorithm for automatic modulation classification based on sixth-order cumulants. In *Proceedings of the XLIV ICEST 2009 Conference*. Veliko Tarnovo (Bulgaria), 2009, vol. 2, p. 635–638. ISBN: 978-954-438-796-9
- [28] GHAURI, S. A. KNN based classification of digital modulated signals. *IJUM Engineering Journal*, 2016, vol. 17, no. 2, p. 71–82. DOI: 10.31436/ijumej.v17i2.641
- [29] BLOCK, H. W., FANG, Z. A. Multivariate extension of Hoeffding's lemma. *The Annals of Probability*, 1988, vol. 16, no. 4, p. 1803–1820. [Online] Cited 2020-10-16. Available at: <http://www.jstor.org/stable/2243993>
- [30] SIMIC, M., STANKOVIC, M., ORLIC, V. D., *Simulation Code for Matlab*. [Online] Cited 2021-02-13. Available at: https://github.com/CumulantsAMC/Radioengineering/blob/main/Mcode_AMC_simulations_Simic_Radioengineering.m

About the Authors ...

Mirko SIMIC received his B. S. degree in Electrical Engineering from the University of Belgrade, Serbia in 2002.

He works in Vlatacom Institute, Belgrade, Serbia, as a System Architect. Primary responsibilities include delivery of projects related to radio & fiber telecommunications and border control systems. His research interests include terrestrial telecommunication, smart city, and cybersecurity.

Milos STANKOVIC received his Dipl. Ing. degree from the School of Electrical Engineering, Univ. of Belgrade, Serbia, in 2002. He received his Ph.D. degree in Systems and Entrepreneurial Engineering from the University of Illinois at Urbana-Champaign (UIUC), USA in 2009. In 2012 he joined the Innovation Center, School of Electrical Engineering, Belgrade, as an EU Marie Curie Research Associate Professor. In 2017 he joined the Singidunum University, Belgrade, as an Associate Professor. His interests include networked control systems, machine learning, dynamic game theory, optimization and decentralized decision making, big data analytics, Cyber-Physical Systems and the IoT.

Vladimir D. ORLIC (corresponding author) received his B. S. degree in 2007 and Ph.D. degree in 2011, in Electrical Engineering from the University of Belgrade, Serbia. He works as a Systems Architect in Vlatacom Institute, Belgrade, Serbia, responsible for project management in the domain of surveillance systems, along with scientific and development research. He has the formal title of the Scientific Associate. His research interests include communication systems, signal processing, algorithm development, security systems, and radio communications.