MULTIPLE HYPOTHESIS MODULATION CLASSIFICATION BASED ON CYCLIC CUMULANTS OF DIFFERENT ORDERS

Pierre Marchand and Jean-Louis Lacoume

CEPHAG-ENSIEG
Domaine Universitaire
38402 Saint Martin d'Hères Cedex, France
e-mail: Pierre.Marchand@cephag.inpg.fr

Christophe Le Martret

CELAR
35170 Bruz, France

ABSTRACT

A multiple hypothesis modulation QAM classification task is adressed in this paper. The classifier is designed within the rigorous framework of decision theory. A characteristic feature is extracted from the signal, and is compared to the possible theoretical features in the maximum likelihood sense. This feature is composed of a combination between fourth-order and squared second-order cyclic temporal cumulants. No assumption about the power of the signal is made. It is shown that this uncertainty about the power of the signal does not affect the decision rule. As an application, we present simulated performance in the context of 4-QAM *vs* 16-QAM *vs* 64-QAM classification.

1. INTRODUCTION

The modulation classification process consists in determining the modulation type of an intercepted signal corrupted by noise and interferences. This is a challenging problem that is of course relevant in some military comunication systems and that has been investigated for several years. Once a signal is detected, the modulation type and its parameters have to be determined in order to select the proper demodulation scheme. One of the classical ways to tackle with the problem of modulation classification is based on maximum likelihood decision theory. It consists in processing the log-likelihood function of the signal (or an approximation thereof) and then comparing to an appropriate threshold. This approach has been proposed to classify QAM modulations [3] and PSK modulations [2]. A simpler way to derive a classifer structure relies on the pattern-recognition concept. This second approach consists in extracting "features" which are signatures of each specific signal format. In many cases, the pattern-recognition techniques are designed on a intuitive basis, and the attention is focused on practical implementation rather than on theoretical background. However, it is possible to formalize this approach by applying the rigorous framework of decision theory. By comparison to the methods proposed in ([3],[2]), the decision rule is then based on the likelihood function of the features rather than the approximated likelihood function of the signal. Such a classifier has been proposed in [9]. The decision rule is based on a vector containing cyclic cumulants (possibly of different orders, possibly at different cycle frequencies). This method is probably well designed for FSK signal classification, but can unfortunately not achieve QAM signal classification (reasons for this pitfall will be detailed in the sequel).

The purpose of this paper is to propose a classifier following the principle proposed in [9], but whose decision rule is based on a new feature, so that QAM signal classification is now possible. Our discriminating feature is a vector containing the samples of mixed-orders cyclic cumulants of the signal (namely, fourth- and second-order properly mixed cyclic cumulants). The idea of mixing different orders cumulants to classify QAM signals was first initiated in [5]. Then it was successfully applied to classify two QAM modulations in both cyclostationary [7] and stationary [6] contexts. This approach is quite because a novel statistic is built; it is neither a cumulant nor a moment, but is choosen in order to achieve the higher discrimination as possible. Generalizing the ideas of [7] and [6], we design in this paper a multiple hypothesis classifier in a rigorous decision theory framework. Besides, it should be underlined that unlike [9], [7], [6], we take here into account the lack of knowledge about the power of the noise. The classifier is then applied to the 4-QAM / 16-QAM / 64-QAM problem.

The paper is composed as follows. In section 2, we recall the expressions of fourth- and second order cyclic correlations for digital modulations. The theory of higher-order cyclostationarity has been developped mainly in [1] and [8]. We will adopt here a stochastic framework and notations well suited for complex-valued random processes. In section 3, the basic idea on which the classifier is based is explained, and the general structure of the algorithm is given. Simulations are provided in section 4.

2. PRELIMINARIES

2.1. Signals of interest

The modelling adopted in this paper is based on the stochastic theory of random processes. In our study, we are interested in *N*-QAM modulation classification (*i.e. N*-states Quadrature Amplitude Modulation). The discrete-time analytic signal representation for these modulations is:

$$x(t) = e^{2i\pi f_c t} \sum_k s_k q(t - kT_b - t_0) \tag{1}$$

where $t \in \mathbb{Z}$, $\{s_k = a_k + i.b_k\}_{k \in \mathbb{Z}}$ is a complex-valued, zero-mean, and *i.i.d.* symbol sequence with values in a N-dimensional set, T_b is the symbol duration, f_c is the carrier frequency, $q(t)$ is the real-valued pulse function, and t_0 is a non-random time shift. In this paper, we do not stationarize the signals, and consequently, time-dependency must be taken into account when expressing the temporal cumulants of (1). In other words, $x(t)$ is modeled as a cyclostationary process.

2.2. Cyclic multicorrelations

Let $C_{x,p+q,p}(t;\tau)$ be the $(p+q)$th-order cumulant-based correlation of the process $x(t)$, defined with p non-conjugated terms and q conjugated terms, as in [4]:

$$\begin{aligned} C_{x,p+q,p}(t;\tau) = \\ Cum[x(t), x(t+\tau_1), ..., x(t+\tau_{p-1}), \\ x^*(t-\tau_p), ..., x^*(t-\tau_{p+q-1})] \end{aligned} \tag{2}$$

Since $x(t)$ is almost-cyclostationary, there are at most countably many values of α for which the so-called $(p+q)$th-order cyclic correlation, defined by:

$$c^{\alpha}_{x,p+q,p}(\tau) = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=0}^{T-1} C_{x,p+q,p}(t;\tau) \exp(-i\alpha t) \ . \tag{3}$$

is non-zero.

Let us precise the expressions at order four ($p + q = 4$). We will consider the definition in which there are as many conjugated as non-conjugated terms ($p = q = 2$). Applying (3) to the process (1), it can be readily shown that the modulus of the cyclic tricorrelation (also called fourth-order temporal cumulant [1],[8]) is given by:

$$\begin{aligned} \left|c^{\alpha}_{x,4,2}(\tau)\right| &= \left|\frac{\mathfrak{C}_{s,4,2}}{T_b} \sum_{t=-\infty}^{t=+\infty} q(t)\, q(t+\tau_1)\, q(t-\tau_2) q(t-\tau_3) \right. \\ &\left. \cdot \exp(-i\alpha t)\right| \end{aligned} \tag{4}$$

where $\mathfrak{C}_{s,4,2}$ is the stationary fourth-order cumulant of the random sequence $\{s_k\}$: $\mathfrak{C}_{s,4,2} = Cum[s_k, s_k, s_k^*, s_k^*]$. In (4), it is necessary to consider the modulus, in order to avoid terms depending on t_0 and f_c, which are *a priori* both unknown.

Similarly, at order two, the modulus of the cyclic correlation of the process (1) is:

$$\left|c^{\alpha}_{x,2,1}(\tau)\right| = \left|\frac{\mathfrak{C}_{s,2,1}}{T_b} \sum_{t=-\infty}^{t=+\infty} q(t)\, q(t-\tau) \exp(-i\alpha t)\right| \tag{5}$$

where $\mathfrak{C}_{s,2,1} = Cum[s_k, s_k^*]$.

3. CLASSIFIER DESIGN

3.1. Problem statement

Classifying an observed signal $y(t)$ in one of M classes of possible modulation types mod_1,...,mod_M, may be formulated as an M-ary testing problem between H_1,...,H_M given by

$$H_i : y(t) = a \cdot x_i(t) + n(t) \quad i = 1, ..., M \tag{6}$$

where the modulation type of $x_i(t)$ is mod_i, and $n(t)$ is an additive white gaussian stationary noise with power unknown. The multiplicative factor a is introduced in order to formalize the lack of knowledge about the power of the signal (or, equivalently, the power of the noise). Thus, a is a random parameter independant of $b(t)$, and whose probability density function $p_A(a)$ is unknown.

Building a classifier on characteristic features rather than on the signal itself is equivalent to re-formulate the problem (6) as follows:

$$H_i : \widehat{\mathbf{r}}_T = u \cdot \mathbf{r}_i + \mathbf{e}_T \quad i = 1, ..., M \tag{7}$$

where $\mathbf{r}_i$ corresponds to a vector containing the theoretical characteristic features of $y(t)$. The vector $\widehat{\mathbf{r}}_T$ is the estimation of the features over T samples of the received signal $y(t)$. The vector $\mathbf{e}_T$ is the corresponding estimation error. The random parameter u with unkown probability function $p_U(u)$ is linked to the parameter a in a manner depending of the features choosen. The parameter u is independant of $\mathbf{e}_T$.

The optimal classifier in the maximum likelihood sense is the one which decides "H_k true" if the conditional probability function of $\widehat{\mathbf{r}}_T$, $p_{\widehat{\mathbf{R}}_T \mid H_i}(\widehat{\mathbf{r}}_T \mid H_i)$ is maximum for $i = k$.

The method proposed in [9] uses a vector $\mathbf{r}_i$ consisting of true cyclic multicorrelations, e.g.

$$\mathbf{r}_i = \left[c^{\alpha_1}_{x_i,p_1+q_1,p_1}(\tau), c^{\alpha_2}_{x_i,p_2+q_2,p_2}(\tau)\right] \tag{8}$$

It should be noted that, since the noise $v(t)$ is stationary, if one includes in (8) only non-zero cycle frequencies, then the theoretical vector $\mathbf{r}_i$ for $y(t)$ is exactly the same for $x_i(t)$. Since the estimators of cyclic multicorrelations are consistent and asymptotic normal [8], it follows that under

H_i, $\mathbf{e}_T$ is asymptotically zero-mean, multivariate complex gaussian with covariance matrix Σ_i. We emphasized in section 2 that the cyclic multicorrelations of all the QAM signal are the same to within a multiplicative factor. Consequently, if the vectors $\mathbf{r}_i$ were defined as in (8), they would be all proportional, and in this case, a maximum likelihood classifier is of no interest. So we have to find out another feature.

3.2. The new feature

The feature we will use is defined as:

$$\mathbf{r}_i = \frac{\widetilde{\mathbf{r}}_i}{||\widetilde{\mathbf{r}}_i||}, \widetilde{\mathbf{r}}_i = \left[\left|c^{2\pi/T_b}_{x_i,4,2}(a\tau, b\tau, c\tau)\right| + \lambda \left|c^{2\pi/T_b}_{x_i,2,1}(\tau)\right|^2\right]. \tag{9}$$

where $(a, b, c) \in \mathbb{Z}$, $\lambda \in \mathbb{R}$ and $\tau = 0, ..., T_b$. Hence $\mathbf{r}_i$ is a $(T_b + 1)$-dimensional vector, consisting in a normalized combination of a cyclic tricorrelation and a squared cyclic correlation, both at the first cycle frequency $2\pi/T_b$. The tricorrelation is considered only on a domain restricted to a line containing the origin, and parametrated by $a, b,$ and c. One can qualify $\mathbf{r}_i$ as a "generalized fourth-order function", because the global order is four, but is it neither a cumulant nor a moment. This new fourth-order function gives raise to vectors $\mathbf{r}_i$ that are not proportional when i varies. The reasons for these choices are detailed in [5], [6], [7].

Moreover, the parameters $a, b, c,$ and λ will be adjusted to maximize the distance between the $\mathbf{r}_i$'s, so that the minimum achievable probability of error be as low as possible. Since the vectors $\mathbf{r}_i$ defined in (9) have all the same norm, it is clear that the distribution is optimum if their isobarycentre

$$\mathbf{s} = \frac{1}{M}(\sum_{i=1}^{M} \mathbf{r}_i) \tag{10}$$

is null. Consequently, the optimum parameters $a, b, c,$ and λ should ideally minimize the norm of the isobarycentre:

$$(a, b, c, \lambda)_{opt} = \arg\min ||\mathbf{s}||\,. \tag{11}$$

Since $||\mathbf{r}_i|| = 1\ \forall i$, one can easily see that this condition is equivalent to

$$\begin{aligned}(a, b, c, \lambda)_{opt} &= \arg\min \frac{2}{M(M-1)} \sum_{i=1}^{M} \sum_{j=i+1}^{M} \rho_{i,j} \\ &\triangleq \arg\min \rho \end{aligned} \tag{12}$$

where $\rho_{i,j} = \mathbf{r}_i \cdot \mathbf{r}_j$. Note that ρ can be interpreted as the mean of all the correlations that one can define in the set $\{\mathbf{r}_1, ..., \mathbf{r}_M\}$.

3.3. Decision rule

In this section, we show that under the condition that the vectors $\mathbf{r}_i$ are normalized, the optimal decision rule for the M-ary testing problem defined in (7) is the same as in the case $u = 1$. The MV classifier maximizes the conditional probability function $p_i \triangleq p_{\widehat{\mathbf{R}}_T \mid H_i}(\widehat{\mathbf{r}}_T \mid H_i)$. It can be readily shown that

$$p_i = \int p_U(u) \cdot p_{\widehat{\mathbf{R}}_T \mid U}(\widehat{\mathbf{r}}_T | u, H_i) du. \tag{13}$$

Since $\mathbf{e}_T$ is asymptotically gaussian in (7), maximizing p_i is asymtotically equivalent to maximizing

$$p'_i = \int p_U(u) \cdot \exp\left(-\frac{||\widehat{\mathbf{r}}_T - u\mathbf{r}_i||^2}{\sigma^2}\right) du. \tag{14}$$

Note that to derive (14), we made the approximation $\Sigma_i = \sigma^2 I$, which of course leads to a convenient but suboptimal scheme. Now, it is obvious that if $f(t) > g(t)$, then $\forall t$, $\int f(t)dt > \int g(t)dt$. Hence, maximizing

$$p''_i = \exp\left(-\frac{||\widehat{\mathbf{r}}_T - u\mathbf{r}_i||^2}{\sigma^2}\right) \tag{15}$$

will guarantee that p'_i is maximized. If $||\mathbf{r}_i||$ is independant of i, it is straightforward to see that maximizing p''_i is equivalent to maximizing $\widehat{\mathbf{r}}_T \cdot \mathbf{r}_i$, which corresponds to the classical correlation receiver.

Finally, the classifier will decide that the modulation type is mod$_k$ whenever

$$\widehat{\mathbf{r}}_T \cdot \mathbf{r}_k > \widehat{\mathbf{r}}_T \cdot \mathbf{r}_i\ \forall i \neq k. \tag{16}$$

4. APPLICATION TO 4-PSK *VS*. 16-QAM *VS*. 64-QAM CLASSIFICATION

4.1. Computation of optimal parameters

We suppose now $M = 3$, mod$_1$ = 4-QAM, mod$_2$ = 16-QAM, and mod$_3$ = 64-QAM. In this case, the correlation coefficient defined in (12) becomes:

$$\rho = \frac{1}{3}(\mathbf{r}_1 \cdot \mathbf{r}_2 + \mathbf{r}_1 \cdot \mathbf{r}_3 + \mathbf{r}_2 \cdot \mathbf{r}_3) \tag{17}$$

This coefficient is easily computable using (9), (4) et (5) with:

$$\begin{array}{lll} \mathfrak{C}_{s,4,2} = -1 & \mathfrak{C}_{s,2,1} = 1 & \text{for 4-QAM} \\ \mathfrak{C}_{s,4,2} = -0.68 & \mathfrak{C}_{s,2,1} = 1 & \text{for 16-QAM} \\ \mathfrak{C}_{s,4,2} = -0.619 & \mathfrak{C}_{s,2,1} = 1 & \text{for 64-QAM} \end{array} \tag{18}$$

We also suppose that the pulse function of the modulations is given by: $q(t) = 1$ for $t = 0, ..., T_b - 1$ and $q(t) = 0$ elsewhere.

The exhaustive minimization of (17) is then performed thanks to the SIMPLEX algorithm, which lead to the following optimal parameters:

$$(a, b, c, \lambda)_{opt} = (0, 1, 1, -2.867) \tag{19}$$

and the corresponding correlation coefficient is given by:

$$\rho_{\min} = 0.28 \tag{20}$$

Note that this correlation coefficient is a little higher than the correlation coefficient exhibited in the case of a binary hypothesis test ($\rho_{\min} = -0.06$, see [7]). This will result in some inevitable degradation of the performance.

4.2. Simulations

The performance of the classifier has been simulated by Monte-Carlo runs.

Simulations have been performed on synthetic data in white gaussian noise for two different signal to noise ratios (SNR=5 dB and 10 dB; for SNRs under 5 dB, the classifier exhibited poor performance, even for large sample sizes). The SNR is defined as follows:

$$SNR = 10 \log \frac{\sum_i x^2(i)}{\sum_i n^2(i)} \tag{21}$$

The number of transmitted symbols N_s varies from 64 to 4096 symbols, with $T_b = 10$ (*i.e.* $T = 640$ to 40960 samples). For each couple (SNR, N_s), 500 different signals (different symbol sequences and noise samples) are generated for each of the three modulations. The figure 1 gives the performance (probability of correct classification in %, or P_{cc}) obtained for different N_s and for a given SNR.

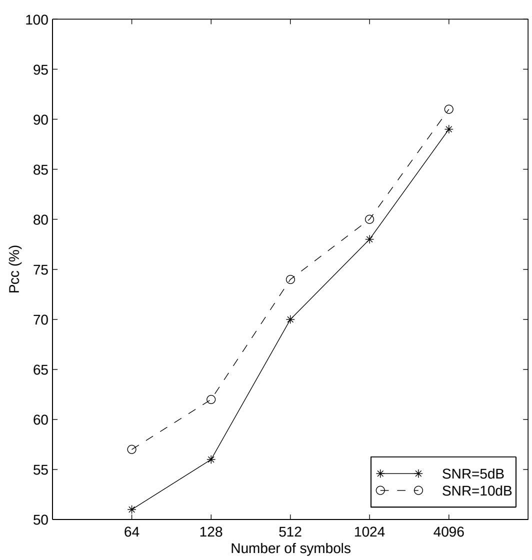

Figure 1: Performance of the classifier

5. DISCUSSION

As shown by the simulations, the classifier is not reliable for sample sizes less than 1024 symbols, for the SNRs tested here. Due to lack of space, the confusions matrices are not shown here, but if we look precisely at them, it is obvious that most of the classification errors are due to confusion between 16-QAM and 64-QAM. This is because the value of their fourth-order cumulants are very close (0.68 *vs.* 0.619). Theoretically, it could be worth to imagine a discriminating feature involving sixth-order statistics, but it may be not reasonable from a estimation point of view. However, to our knowledge, the feature introduced in this paper is the only one that can achieve QAM modulation classification in a cyclostationary context.

6. REFERENCES

[1] C.M. SPOONER, W.A. GARDNER, "The cumulant theory of cyclostationary time-series, Part I: Foundation; Part II: Development and applications," *IEEE Trans. on Sig. Proc.*, vol. 42-12, Dec. 1994.

[2] C.Y. HUANG, A. POLYDOROS, "Likelihood methods for MPSK modulation classification," IEEE Trans. on Com., No. 43-2/3/4, Feb./Mar./Apr. 1995.

[3] A. POLYDOROS, K.KIM, "On the detection and classification of quadrature digital modulations in broadband noise," IEEE Trans. on Com., vol. 38, Aug. 1990.

[4] P.O. AMBLARD, M. GAETA AND J.L LACOUME, "Statistics for complex variables and signals - Part II: Signals," Signal Processing, vol. 53-1, Aug. 1996

[5] P. MARCHAND AND D. BOITEAU, "Higher-order statistics for QAM signals: a comparison between cyclic and stationary representations," EUSIPCO '96, Trieste, Sept. 1996.

[6] D. BOITEAU, C. LE MARTRET, "Modulation classification by means of different orders statistical moments" MILCOM 5-7 Nov. 1997, Monterey, California.

[7] P. MARCHAND, C. LE MARTRET AND J.L. LACOUME, "Classification of linear modulations by a combination of different orders cyclic cumulants," IEEE Workshop on HOS, Banff, Canada, July 1997.

[8] A.V. DANDAWATÉ AND G.B. GIANNAKIS, "Asymptotic theory of mixed time averages and Kth-order cyclic-moment and cumulant statictics," IEEE Trans. on Inf. Theory., vol. 41-1, Jan. 1995.

[9] A.V. DANDAWATÉ AND G.B. GIANNAKIS, "Polyspectral analysis of non-stationary signals: system identification, classification and ambiguity functions," Higher-Order Statistics - J.L. Lacoume editor, Elsevier Science publishers, 1992.