A GENERAL MAXIMUM LIKELIHOOD FRAMEWORK FOR MODULATION CLASSIFICATION

D. Boiteau, C. Le Martret***

* Centre d'Etudes de Systemes et de Techniques Avancees, 35170 Bruz, France - cesta@mail35.galeode.fr
** Centre d'ELectronique de l'ARmement, 35170 BRUZ, France - lemartrc@celar.fr

ABSTRACT

This paper deals with modulation classification. First, a state of the art is given which is separated into two classes: the pattern recognition approach and the Maximum Likelihood (ML) approach. Then we propose a new classifier called the General Maximum Likelihood Classifier (GMLC) based on an approximation of the likelihood function. We derive equations of this classifier in the case of linear modulation and apply them to the M PSK / M' PSK problem. We show that the new tests are a generalisation of the previous ones using ML approach, and don't need any restriction on the baseband pulse. Moreover the GMLC provides a theoretical foundation for many empirical classification systems including those systems that exploit cyclostationary property of modulated signals.

1 - INTRODUCTION

In this paper, we address the problem of digital modulation classification. Classically, one can consider two main approaches: the Pattern Recognition (PR) approach and the Maximum Likelihood (ML) approach. In the first one some statistical representations of the signal or of some of its parameters are extracted from the observed signal and used as discrimating features. In the ML approach, quasi-optimal rules are derived from the development of the average log-likelihood function of the signal. However, these developments [4] are only valid for baseband pulse of duration equal to the symbol period.

In this paper, we present the General Maximum Likelihood Classifier (GMLC) based on a new ML approximation. In application to modulation classification, this classifier applies to any baseband pulse and particularly no restriction on the duration is needed. This approach provide new tests which are generalisation of the ones given in [4] for linear modulation classification. Moreover, it gives a theoretical ML framework to higher-order cyclic moment and cumulant based classifiers such as the General Search Algorithm presented in [38-40].

This paper is organised as follow. In part 2, we present a state of the art of linear modulation classification. In part 3, we present the GMLC and its application to linear modulation classification where the MPSK/M'PSK case is detailed. A conclusion is given in part 4.

2 - STATE OF THE ART

2.1 - The Pattern Recognition approach

In the PR approach, two main classes can be separated, since they use statistics of extracted parameters or statistics of signal itself.

In the first class, the extracted parameters are conveniently chosen as the parameters supporting the symbol information and the discriminating features are some approximation of their density function. For PSK signals, the phase density function has been approximated by histogram in [7-9], by histogram of the difference of two consecutive phase measured at synchronous time in [10-14], and by statistical (higher order) moments in [15-24]. For QAM and ASK signal, the density of the modulus of the signal has been approximated by histogram in [7-8,14,19,22-23] and higher order moments in [10,24-26].

In the second class, the discriminating features are statistics of the signal itself or its constellation representation (after symbol period and synchronisation estimation).

In [27-31] the constellation classification is addressed. In [27-29] the first maximum or the first zero of the characteristic function of the constellation indicates the number of states in QAM constellations. In [30], the measurement of the Hellinguer distance between the density function of any two constellations is used. In [31], a cluster analysis is applied on the observed constellation in addition to 4th order cumulants.

In [32-33], a discriminating features is composed as a combination of fourth and second order moments of the "stationarized" signal (which leads to a fourth order cumulant in [32]).

In [34-43], explicit exploitation of the cyclostationarity of the signal are proposed (see [34,44] for a good introduction to the cyclostationary theory and link with the Wigner-Ville time frequency representation [43]). First classifiers operate in the frequency domain. In [35], cycle spectral moment functions of order two is used in a correlation based approach to discriminate between PSK and OQPSK signals. As it can be shown that a large class of linear modulated signals have identical second order statistics [34], researchers have since explored higher orders for further discrimination. In [36-"Mth law" classifier-,37,41], presence or absence of spectral lines in the spectrum of signals at the output of different order nonlinearities[1] is used as a feature for the classification system. In the temporal domain, [42] propose a discriminating feature composed as a combination of fourth and second order cumulant at cycle frequency $1/T$ to discriminate between 4 PSK and 16 QAM. Recently, in [38-40], a general search algorithm has been proposed were the discriminating feature is the set of maxima (evaluated over different temporal delay vectors) of all cycle cumulants of different orders for cycle frequencies associated with a given signal. The deduced system is based on a maximum correlation classification of the extracted features since theoretical reference can be obtained. With the same approach, optimal ML classifiers have been theoretically explored in [45] in application to the general class of cyclostationary signals.

Theoretical optimality in a ML sense of correlation based classifiers using temporal or spectral moment functions has only been demonstrated at order 2 [46]. In part 3, we extend this result to higher order moment functions.

2.2 - The Maximum Likelihood approach

In the ML approach, preliminary results can be found in [1] in application to the BPSK vs. QPSK modulation classification. Since this paper, many extensions (concerning the general MPSK and MQAM modulation classification problem [2-5], the special case of staggered modulation OQPSK [3], and multiple hypothesis modulation classification [6]) have been published exploiting the same theoretical developments. Recently alternative developments [47] have been proposed in application to MPSK classification in a coherent and synchronous environment with square baseband pulse. In this section, we focus on the development given in [1-6] since they are not only valid in this special environment.

In [1-6], the authors assume that the baseband pulse is of duration equal to the symbol period. This restriction allows them to develop the average log likelihood function under a small signal to noise ratio, and after some tedious calculus, they propose [4] 3 tests, noted q_M for PSK and QAM signal classification and alternated tests noted p_M for QAM signal classification [5]. These tests are expressed in terms of Mth order moments of a stationary match-filtered version of the observation. In [3], a table is presented which gives the value of possible M - generally obtained as the lowest order for which classification is possible - for many binary classification problems.

In addition to extensive simulations which show the quasi-optimality of the q_M tests, a very interesting study is presented in [4] which links these tests to many other PR based systems. As main result and with some simple extensions, it demonstrates that all the PR systems described in the first class in § 2.1 can be viewed as implementation of equivalent or simplified and sub-optimal versions of the q_M-tests (for phase) and p_M (for modulus). Furthermore, the study demonstrated the link between q_M-test and what is called "Mth law" classifiers which have already be referred in §2.1 as classifiers exploiting the cyclostationary property of the signal [36]. Then, it appears that a general ML framework must exist which globally justified the different approaches we have listed above, including correlation based classifier using temporal or spectral moment functions as discriminating features. This point is the object of the 3rd part of this paper.

3 - A NEW GENERAL MAXIMUM LIKELIHOOD APPROACH

3.1 - The General Maximum Likelihood Classifier (GMLC)

Let us define the set $\{s(t,\underline{\theta}_C), C=1,\dots,N_C, t\in Z\}$ of N_C reference signals where the vector $\underline{\theta}_C$ is the set of the parameters describing the signal C. The classification problem we want to solve is the following: given the complex observation

$$r(t)=s(t,\underline{\theta}_C)+n(t),\ 0\le t\le N$$

where $n(t)$ is a white gaussian noise with power spectral density N_0, find the number C of the emitted signal

When the parameters $\underline{\theta}_C$ can be probabilised one can classically define the Averaged Likelihood Function (ALF) of $r(t)$ under C hypothesis as[2]

$$\mathrm{ALF}_C=\mathbb{E}_{\underline{\theta}_C}\left[\exp\left(\frac{1}{N_0}\int_0^N \mathrm{Re}\left[r(t)s^*(t,\underline{\theta}_C)\right]dt\right)\right] \qquad (1)$$

where $\mathbb{E}_{\underline{\theta}_C}$ denote the expectation over $\underline{\theta}_C$, $\mathrm{Re}[\cdot]$ the real part and z^* the conjugate of z. The optimal classifier in the Maximum Likelihood sense is then given by:

$$\hat{C}=\underset{C=1\cdots N_C}{\arg\max}\left(\mathrm{ALF}_C\right). \qquad (2)$$

Unfortunately there are no closed form expression for (1). Some authors [1-6] have proposed some approximations of (1) that give sub-optimal tests under the restriction that the pulse is of duration T. Here we present a new approximation of (1) which generalises the previous one's without no restriction on the pulse function. This classifier will be named "General Maximum Likelihood Classifier" (GMLC).

Using the power series expansion of the exponential function in (1) and the identity $\mathrm{Re}[z]=(z+z^*)/2$ the ALF can be written

$$\mathrm{ALF}_C=1+\sum_{n=1}^{\infty}\alpha_n\lambda_{n,C},\quad \alpha_n=(n!(2N_0)^n)^{-1} \qquad (3)$$

$$\text{where}\quad \lambda_{n,C}=\mathbb{E}_{\underline{\theta}_C}\left[\int_0^N\left[r(t)s^*(t,\underline{\theta}_C)+r^*(t)s(t,\underline{\theta}_C)\right]dt\right]^n. \qquad (4)$$

Let us now define the vector $\underline{t}=[t_1,\dots,t_n]$ with $t_i\in[0,\dots,N]$ and the set of all possible couples $(P_j,\overline{P}_j)$, $j=1,\dots,2^n$, so that $P_j\cup\overline{P}_j=\{1,\dots,n\}$ and $P_j\cap\overline{P}_j=\emptyset$. Then (4) can be developed as:

$$\lambda_{n,C}=\mathbb{E}_{\underline{\theta}_C}\int_{\underline{t}}\sum_{j=1}^{2^n}\left\{\prod_{p\in P_j}r(t_p)s^*(t_p,\underline{\theta}_C)\cdot\prod_{q\in\overline{P}_j}r^*(t_q)s(t_q,\underline{\theta}_C)\right\}d\underline{t} \quad (5)$$

which reduces to:

---

[1] It can be easily shown that the Fourier transform of the nth power of the signal can be expressed as the integration of the nth spectral moment function over the $(n$-1)th spectral frequencies which can be viewed as some kind of projection on the cycle frequencies axis.

[2] In this definition, the term depending only of the received signal which is a common constant to all LLF is dropped. The term depending only of the signal $s(t,\underline{\theta}_C)$ of finite averaged power is here assumed, under a long observation time, independant of the random parameters $\underline{\theta}_C$, and is also dropped.

$$\lambda_{n,C} = \sum_{j=1}^{2^n} \int_{\underline{t}} R_r^j(\underline{t}) \cdot R_{s_C}^j(\underline{t}) d\underline{t} = \sum_{j=1}^{2^n} \Gamma_C^j \tag{6}$$

where $$R_r^j(\underline{t}) = \prod_{p \in P_j} \prod_{q \in \overline{P}_j} r(t_p) r^*(t_q), \tag{7}$$

$$R_{s_C}^j(\underline{t}) = \mathbb{E}_{\underline{\theta}_C} \prod_{p \in P_j} \prod_{q \in \overline{P}_j} s^*(t_p, \underline{\theta}_C) \cdot s(t_q, \underline{\theta}_C), \tag{8}$$

$$\Gamma_C^j = \int_{\underline{t}} R_r^j(\underline{t}) \cdot R_{s_C}^j(\underline{t}) d\underline{t}. \tag{9}$$

For each particular value of j, Γ_C^j can be interpreted as the measurement of the correlation (integration over $\underline{t}$) between the temporal moment function (8) of the reference, and the estimated one (7) of the observation. Note that by application of time-frequency duality Γ_C^j can be also interpreted as the correlation (integration over frequency domain) between reference and estimated spectral moment functions. Similar results have been obtained in [48] for MFSK modulation classification assuming a square pulse shape.

Then considering development[3] of (1) up to order Q with equations (6-9) the GMLC is given by

$$\hat{C} = \underset{C=1 \dots N_C}{\arg\max} \left(\sum_{n=1}^{Q} \alpha_n \lambda_{n,C} + \beta_C \right) \tag{10}$$

where β_C are constants that are adjust in order to minimise the Error Probability of the classifier. The value of Q is chosen at least to permit classification of the N_C signal.

Note at this point that no strong assumptions have been made on the signal model (see note 2) and the GMLC can be applied to every modulation type. In the next section we consider the linear modulation classification problem.

3.2 - GMLC tests for linear modulations

For linear modulations signals $s(t, \underline{\theta}_C)$ can be written

$$s(t, \underline{\theta}_C) = \sqrt{S} \sum_{k=0}^{N_s} S_{C,k} \cdot h(t - kT + t_0) e^{i\phi_0} \tag{11}$$

where S is the signal power, $h(t)$ is the baseband pulse with $\int h^2(t) dt = 1$, T is the symbol period and $\underline{\theta}_C = \{S_C, t_0, \phi_0\}$. The set $S_C = \{S_{C,k}, k = 1, \dots, N_s\}$ is an iid sequence of complex symbols which belong to a finite set depending of the modulation with $\mathbb{E}_{S_C}[S_{C,k}] = 0$ and $\mathbb{E}_{S_C}\left[|S_{C,k}|^2\right] = 1$. t_0 is the symbol timing offset and ϕ_0 is the carrier phase. In this paper, we suppose that the symbol period T and the pulse shape $h(t)$ are *a priori* known and using terminology of [4] we consider four different environments depending on the statistics of t_0 and ϕ_0:

- CS (coherent and synchronous): $t_0 = 0$ et $\phi_0 = 0$,
- NCS (non-coherent and synchronous): $t_0 = 0$ and ϕ_0 is a r.v. uniformly distributed over $[0, 2\pi[$,
- CA (coherent and asynchronous): $\phi_0 = 0$ and t_0 is a r.v. uniformly distributed over $[0, T[$,
- NCA (non-coherent and asynchronous): t_0 is a r.v. uniformly distributed over $[0, T[$ and ϕ_0 is a r.v. uniformly distributed over $[0, 2\pi[$.

GMLC tests for PSK and QAM signals in these environments can be systematically defined using (10) and (6).

Let us first introduce the following notations:

$$R_h^j(\underline{t}, t_0) = \prod_{p \in P_j} \prod_{q \in \overline{P}_j} h(t_p - k_p T + t_0) \cdot h^*(t_q - k_q T + t_0), \tag{12}$$

$$\overline{R}_h^j(\underline{t}) = \frac{1}{T} \int_T R_h^j(\underline{t}, t_0) dt_0, \tag{13}$$

$$m_{S_C}^j = \mathbb{E}_{S_C} \prod_{p \in P_j} \prod_{q \in \overline{P}_j} S_{C,k_p} \cdot S_{C,k_q}^*, \tag{14}$$

where (14) is one of the nth order moment of the symbols for signal C. In addition, if we define $l = \mathrm{card}(\overline{P}_j)$ then we note

$$\Delta^j = \tfrac{1}{2\pi} \int_0^{2\pi} \exp(i(n - 2l)\phi_0) d\phi_0 = \delta_{n-2l,0} \tag{15}$$

where δ is the Kronecker symbol. With these notations and replacing the signal model (11) in (8), we obtain

$$R_{s_C}^j(\underline{t}) = \sum_{k_1 \cdots k_n} m_{S_C}^j \Psi^j(\underline{t}) \tag{16}$$

where $\Psi^j(\underline{t})$ is given depending on the different cases by:

- CS: $$\Psi^j(\underline{t}) = R_h^j(\underline{t}, 0) \tag{17a}$$
- NCS: $$\Psi^j(\underline{t}) = R_h^j(\underline{t}, 0) \cdot \Delta^j \tag{17b}$$
- CA: $$\Psi^j(\underline{t}) = \overline{R}_h^j(\underline{t}) \tag{17c}$$
- NCA: $$\Psi^j(\underline{t}) = \overline{R}_h^j(\underline{t}) \cdot \Delta^j. \tag{17d}$$

Then, the correlation Γ_C^j becomes:

$$\Gamma_C^j = \sum_{k_1 k_2 \dots k_n} m_{S_C}^j \int_{\underline{t}} R_r^j(\underline{t}) \cdot \Psi^j(\underline{t}) d\underline{t} \tag{18}$$

Formula (14) (17) (18) give the way to implement the GMLC.

3.3 - Application to M PSK / M' PSK classification

In the binary classification case (C=1 or C=2) the GMLC test (10) is rewritten as:

$$\sum_{n=1}^{Q} \sum_{j=1}^{2^n} \sum_{k_1, \dots, k_n} \alpha_n \left(m_{S_1}^j - m_{S_2}^j \right) \cdot \int_{\underline{t}} R_r^j(\underline{t}) \cdot \Psi^j(\underline{t}) d\underline{t} \gtrless T_{opt} \tag{19}$$

where T_{opt} is the optimal threshold. The minimum value of Q is chosen in order to verify the two following conditions:

$$\Psi^j(\underline{t}) \neq 0 \tag{20}$$

and $$m_{S_1}^j \neq m_{S_2}^j. \tag{21}$$

It can be easily shown that for M PSK / M' PSK classification ($M' > M$) condition (21) is verified as soon as (14) contains terms of type $\mathbb{E}_{S_C}[S_{C,k}^M]$.

In a CS environment (20) is always verified and the smallest value of Q for which (21) is verified is $Q = M$ with partitions j_0 such that $\ell = 0$ or $\ell = M$, and $k_1 = \cdots = k_M$. In these cases the test (19) reduces to:

$$\mathrm{Re}\left[\int_{\underline{t}} R_r^{j_0}(\underline{t}) \cdot \sum_k R_h^{j_0}(\underline{t}, 0) d\underline{t} \right] \gtrless T_{opt} \tag{22}$$

We can see here that (22) is the expression of the correlation over $\underline{t}$ between the observation's M-th order moment function and the baseband pulse one for all cycle frequency k/T, $k \in Z$. In the particular case where $h(t)$ is of duration T, the test (22) can be simplified and becomes

$$\sum_k \mathrm{Re}\left[\left(\int_{(k-1)T}^{kT} r(t) h^*(t - (k-1)T) dt \right)^M \right] \lessgtr T_{opt} \tag{23}$$

which is the q_M test defined in [4] in a CS environment.

[3] The lowest is the SNR the more accurate is the truncated power series expansion. However it is well known that ML tests obtained by this way works well also at high SNR.

In a NCS case condition (20) imposes to have $n=2\ell$. Under this constraint it comes that the smallest value of Q for which (21) is verified is $Q=2M$ with partitions j_0 such that $\ell=M$, and $k_1=\cdots=k_M$, $k_{M+1}=\cdots=k_{2M}$ and $k_1 \neq k_{M+1}$. In that case the test (19) reduces to:

$$\left|\int_{\underline{t}} R_r^{j_0}(\underline{t}) \cdot \sum_k R_h^{j_0}(\underline{t},0) d\underline{t}\right|^2 \gtrless T_{opt}. \tag{24}$$

This result is well known in the frequency domain where all the cycle spectra at cycle frequency k/T are identically affected by a phase term arising from ϕ_0. As in the CS case, when $h(t)$ is of duration T, it can be shown that (24) simplify to the optimal q_M-test defined in [4] for NCS environment[4].

Similarly, the GLMC tests in asynchronous cases (CA and NCA) are obtained by taking into account the integration (13) which leads to:

$$\text{CA:} \quad \int_T \mathrm{Re}\left[\int_{\underline{t}} R_r^{j_0}(\underline{t}) . R_h^{j_0}(\underline{t},t_0) d\underline{t}\right] dt_0 \lessgtr T_{opt} \tag{25}$$

$$\text{NCA:} \quad \int_T \left|\int_{\underline{t}} R_r^{j_0}(\underline{t}) . R_h^{j_0}(\underline{t},t_0) d\underline{t}\right|^2 dt_0 \gtrless T_{opt} \tag{26}$$

where j_0 is the same in (25) as in (23) and in (26) as in (24).

Note here that the GMLC gives a theoretical demonstration of test (26) that have been proposed in [4] by empirical considerations.

In the frequency domain, it can be shown that (25) is the correlation over the spectral frequencies f between the observation's Mth order moment function and the reference's one at null cycle frequency. In the same way, it can be shown that (26) is the summation over cycle frequencies k/T, $k \in Z$ of the modulus of the correlation over f between the observation's cycle Mth order moment function and the reference's one. This result is in accordance with the fact that cycle spectra are differently affected by a phase term arising from t_0 [34,44].

4 - CONCLUSION

In this paper, we have first presented a state of the art of linear modulation classification. Some new developments in maximum likelihood theory are then given which shown that the likelihood function of a observation given a reference can be closely approximated by a measure of the correlation between empirical and true temporal (or spectral) higher order moment functions. A new General Maximum Likelihood Classifier is derived and applied to linear modulation classification. Deduced tests have been shown to be equivalent to the q_M and p_M-tests obtained in [1-6] under constrained baseband pulse and to provide a general theoretical framework for many pattern recognition based systems exploiting cyclostationarity of the modulated signals for classification as well as detection.

REFERENCES

[1] K. Kim, A. Polydoros, MILCOM conf., 1988.
[2] A. Polydoros, K. Kim, IEEE Tr. on Communications, vol.38, n°8, 1990.
[3] C.Y. Hwang, A. Polydoros, MILCOM conf., 1992.
[4] C.Y. Huang, A. Polydoros, IEEE Tr. on Communications, vol.43, n°2/3/4, 1995.
[5] C.S. Long and al, MILCOM conf., 1995.
[6] K.M. Chung and al, 29th Asilomar conf. on SSC, 1996.
[7] F. Jondral, Signal Processing, n°9, 1985.
[8] L. Vergara Dominguez and al, Signal Processing, n°22, 1991.
[9] P.C. Sapiano and al, ICASSP conf, 1995.
[10] F.F. Lietke, Signal Processing, n°6, 1984.
[11] R.J. Mammone and al, ICC conf., 1987.
[12] S.H. Sue, S. Soliman, Radar Signal Processing, vol.37, Pt.F, n°6, 1990.
[13] S.H. Sue, S. Soliman, MILCOM conf., 1989.
[14] A.J. Nagy, MILCOM conf., 1994.
[15] J.E. Hipp, MILCOM conf., 1986.
[16] S. Soliman, S.Z. Hsue, IEEE Tr. on Communications, vol. 40, n°5, 1992.
[17] Y. Yang, S. Soliman, GLOBECOM conf., 1991.
[18] Y. Yang, S. Soliman, ICC conf., 1991.
[19] R. Lamontagne, ICANN conf., 1993
[20] K. Assaleh and al, MILCOM conf., 1992.
[21] K. Farrel, R.J. Mammone, MILCOM conf., 1993.
[22] E.E Azzouz, A.K. Nandi, IEEE Signal Processing, vol.47, 1995.
[23] A.K. Nandi, E.E Azzouz, IEEE Signal Processing, vol.56, 1997.
[24] C. Louis, P. Sehier, MILCOM conf., 1994.
[25] Y.T. Chan, L.G. Gadbois, Signal Processing, n°16, 1987.
[26] S.B. McMillan and al,ICASSP conf., 1990.
[27] S.L. Wood and al, ICASSP conf., 1988.
[28] S.L. Wood, J.R. Treisher, 22th Asilomar conf. on CSC, 1988.
[29] Q. Zhu and al, ICASSP conf., 1993.
[30] D.L. Dohono, X. Huo, IEEE SPAWC, 1997
[31] A. Swami, B. Sadler, IEEE SPAWC, 1997
[32] D. Boiteau, C. Le Martret, EUSIPCO conf., 1997.
[33] C. Le Martret, D. Boiteau, MILCOM conf. 1997.
[34] W.A. Gardner, *Statistical Spectral Analysis*, Prentice Hall, 1988.
[35] Y.K. Kim, C.L. Weber, MILCOM conf., 1989.
[36] J. Reichert, ICASSP conf., 1992.
[37] G. Feyh and al, MILCOM conf., 1994.
[38] W.A. Gardner, C.M. Spooner, MILCOM conf, 1988.
[39] C.M. Spooner, 2nd workshop on cyclostationary signals, 1994.
[40] C.M. Spooner, 29th Asilomar conf.on SSC, 1996.
[41] B. Rice and al, ICASSP conf., 1994.
[42] P. Marchand and al, workhop HOS, 1997.
[43] S-C Pei, T-Y Wang, ISSPA, vol. 1&2, 1987.
[44] W.A. Gardner, C.M. Spooner, IEEE Tr. on Signal Processing, vol.42, n°12, 1994
[45] A.V. Dantawate, G.B. Giannakis, SPIE, vol 1770, 1992.
[46] W.A. Gardner, IEEE Tr. on Communications, vol.36, n°8, 1988.
[47] P.C. Sapiano, J.D. Martin, MILCOM conf, 1996.
[48] B.F. Beidas, C.L. Weber, IEEE J. on Selected Areas in Communications, vol.13, n°1, 1995.

[4]In the case of QAM classification, we note here that the choice $n=M$ and $l=M/2$ with $k_1=\cdots=k_M$ will lead to another quasi optimal test in the NCS environment which simplify to the p_M-test [5] for $h(t)$ of duration T.