A GENERAL MAXIMUM LIKELIHOOD CLASSIFIER FOR MODULATION CLASSIFICATION

C. Le Martret, D. Boiteau***

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

ABSTRACT

This paper deals with maximum likelihood (ML) classification of digital communication signals. We first propose a new approximation of the average likelihood function. Then we introduce the General Maximum Likelihood Classifier (GMLC) based on this approximation which can be applied to a wide range of classification problem involving finite mean power signals. Derivation of this classifier equations are given in the case of linear modulations with an application to the *M* PSK / *M'* PSK problem. We show that the new tests are a generalization 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 of that exploit cyclostationarity property of digital 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 and the Maximum Likelihood (ML) method. In the first one, some parameters based on statistics such as histograms, moments, ... are estimated from the observation and used as discriminating features (see [7] for reference). In the ML approach, quasi-optimal rules are derived from the development of the Average Likelihood Function (ALF) of the signal. However, these developments [1-3] (for log ALF) are only valid for a baseband pulse of duration equal to the symbol period.

We propose here the General Maximum Likelihood Classifier (GMLC) based on a new approximation of the ALF which can be applied to a wide range of classification problem involving finite mean power signals. In the case of modulation classification, this classifier applies to any baseband pulse and particularly, no restriction on the duration is needed. We can show that this approach provides tests which are a generalization of those given in [1-3] for linear modulations. Moreover, it gives a theoretical foundation for many heuristic classification systems using higher-order cyclic statistics such as the one presented in [5].

This paper is organized as follows. In section 2, we present the GMLC by introducing the ALF approximation. In section 3 we apply GMLC to linear modulation and derive equations for the *M* PSK / *M'* PSK problem. A conclusion is given in part 4.

2 THE GENERAL MAXIMUM LIKELIHOOD CLASSIFIER (GMLC)

2.1 Problem statement

Given a set of N_s reference complex signals $\{s_c(t), c = 1, \cdots, N_s\}$ and an observation $r(t)$ of one of these signals corrupted by white gaussian noise $n(t)$:

$$r(t) = s_c(t) + n(t), \quad t \in [0, L] \tag{0}$$

the classification problem we want to solve is to find out what reference signal was emitted.

We assume that each signal $s_c(t)$ is characterized by a set of random parameters Ω_c and one can classically define the ALF of $r(t)$ under c hypothesis as [2][1]

$$ALF_c = \mathbb{E}_{\Omega_c}\left[\exp\left(\frac{1}{N_0}\int_0^L \mathrm{Re}\left[r(t)\bar{s}_c(t)\right]dt\right)\right] \tag{1}$$

where $\mathbb{E}_{\Omega_c}$ denote the expectation with respect to (w.r.t.) the random parameters, $\mathrm{Re}[\cdot]$ the real part and $\bar{z}$ the conjugate of *z*. The optimal classifier in the Maximum Likelihood sense is then given by:

---

[1] We consider here the continuous case but the extension to discrete case is straightforward.

$$\hat{c} = \arg\max_{c=1,\cdots,N_S} \{ALF_c\}. \quad (2)$$

Unfortunately there are no closed form expression for (1) and this is a challenging problem that have been studied for many years to find the best approximations of the ALF while keeping tractable tests.

2.2 Approximation of the ALF

Most of the solutions proposed to approximate the ALF consist in applying as soon as possible the expectation w.r.t. to the random parameters and then make a power series expansion (PSE) of the result. For instance, this approach have been used in modulation classification [1-3] and sub-optimal tests have been obtained under restriction on the pulse duration. In that case, after the expectation w.r.t. to the symbols, the functional still remains to much complicated and two or three successive PSE are necessary to obtain tractable tests. In some cases these tests can only be given by empirical considerations and validate by extensive simulation[2].

In order to overcome these problems we propose here another way to approximate (1) by making first the PSE of the exponential and then applying the expectation. This approximation is the foundation of the GMLC (section 3) but it can apply to any ML estimation problem involving signals of type (0) such as synchronization (timing and phase estimation).

After PSE of the exponential in (1) the ALF can be re-written

$$ALF_c = 1 + \sum_{n=1}^{\infty} \alpha_n \lambda_{n,c}, \quad \alpha_n = \frac{1}{n!(2N_0)^n} \quad (3)$$

where $\lambda_{n,c} = \mathbb{E}_{\Omega_c}\left[\int_0^L r(t)\cdot \bar{s}_c(t) + \bar{r}(t)\cdot s_c(t)dt\right]^n$. (4)

If we define the vector $\underline{t} = [t_1,\ldots,t_n]$ with $t_i \in [0, L]$ and the set of all possible couples $(P_j, \overline{P}_j)$, $\mathrm{j} = 1,\ldots,2^{\mathrm{n}}$, such that $P_j \cup \overline{P}_j = \{1,\ldots,n\}$ and $P_j \cap \overline{P}_j = \varnothing$ then we can show that (4) can be developed as:

$$\lambda_{n,c} = \sum_{j=1}^{2^n} \int_{\underline{t}} \overline{R}_r^j(\underline{t})\cdot R_{s_c}^j(\underline{t})d\underline{t} = \sum_{j=1}^{2^n} \Gamma_c^j \quad (5)$$

where $$R_r^j(\underline{t}) = \prod_{p\in P_j}\prod_{q\in \overline{P}_j} \bar{r}(t_p)\cdot r(t_q), \quad (6)$$

$$R_{s_c}^j(\underline{t}) = \mathbb{E}_{\Omega c}\prod_{p\in P_j}\prod_{q\in \overline{P}_j} \bar{s}_c(t_p)\cdot s_c(t_q), \quad (7)$$

$$\Gamma_c^j = \int_{\underline{t}} \overline{R}_r^j(\underline{t})\cdot R_{s_c}^j(\underline{t})d\underline{t} \quad (8)$$

Firstly we can note that for each particular value of j the function $R_{s_c}^j(\underline{t})$ is an n-th order moment of the reference signal. Particularly, if $s_c(t)$ is cyclostationary it turns to be a n-th order cyclic moment. Secondly we can see that the integral Γ_c^j in (5) can be interpreted as the correlation (integration over $\underline{t}$) between the reference temporal moment (7) and the estimated one (6) using instantaneous values of the observation. Because of time-frequency duality this integral can also be interpreted as the correlation measurement between reference and estimated spectral moment functions. (Similar results have been obtained in [6] for M FSK modulation classification assuming a square baseband pulse). As a consequence we can say that:

- *i*) the ALF (3), with (4) can be viewed as the weighted sum of the correlations between all moments - at all orders - of the reference signal and those using instantaneous values of the observation.

- *ii*) for digital communication signals (which have cyclostationary property), the ML approach tells us that we have to consider all orders cyclic moments and gives us the optimal way to use them. This result is new compared to those using higher-order statistics based on heuristic approaches.

2.3 The GMLC

Then considering development of (1) up to order Q with equations (5-7) the GMLC is given by[3]

$$\hat{c} = \arg\max_{c=1,\ldots,N_S}\left\{\sum_{n=1}^{Q} \alpha_n \lambda_{n,c} + \beta_c\right\} \quad (9)$$

where Q is chosen at least to permit classification of the N_s signals and β_c are constants. These constants are introduced here to take into account the effects of the PSE truncation and are adjust in order to minimize the classifier error probability.

[2] As far as we know, the only classifier using exact log ALF given in [4] for M PSK signals with restrictive assumptions.

[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 work well also at high SNR.

3 GMLC FOR LINEAR MODULATIONS

3.1 Expressions of the GMLC

For linear modulations with mean power *P*, signals $s_c(t)$ can be written

$$s_c(t) = \sqrt{P} \cdot \sum_k d_{c,k} \cdot h(t - kT + t_0) e^{i\phi_0} \qquad (10)$$

where $h(t)$ is the baseband pulse, *T* the symbol period, $\{d_{c,k}\}$ the complex symbol sequence set, t_0 the symbol timing offset and ϕ_0 the carrier phase. In this paper, we suppose that the symbol period and the pulse shape are *a priori* known. Using terminology of [2] we can derive four GMLC tests depending whether t_0 and ϕ_0 are deterministic or random variables. In the deterministic context they can be set without restriction to zero. In the random context t_0 is a random variable (*rv*) uniformly distributed over $[0, T[$ and ϕ_0 a *rv* uniformly distributed over $[0, 2\pi[$. The four cases can be listed as follows[4]:

- CS (cohe. and sync.): $t_0 = \phi_0 = 0$, $\Omega_c = \{d_{c,k}\}$,
- NCS (non-cohe. and sync.): $t_0 = 0$, $\Omega_c = \{\phi_0, d_{c,k}\}$,
- CA (cohe. and async.): $\phi_0 = 0$, $\Omega_c = \{t_0, d_{c,k}\}$,
- NCA (non-cohe. and async.): $\Omega_c = \{t_0, \phi_0, d_{c,k}\}$.

The different GMLC tests for PSK and QAM signals in these environments can be systematically defined using (3), (5-8).

Let us first introduce the following notations:

$$R_h^j(\underline{t}, t_0) = \prod_{p \in P_j} \prod_{q \in \overline{P}_j} \overline{h}(t_p - k_p T + t_0) \cdot h(t_q - k_q T + t_0), \qquad (11)$$

$$\widetilde{R}_h^j(\underline{t}) = \frac{1}{T} \int_T R_h^j(\underline{t}, t_0) dt_0, \qquad (12)$$

$$m_{d_c}^j = \mathbb{E}_{\{d_{c,k}\}} \prod_{p \in P_j} \prod_{q \in \overline{P}_j} \overline{d}_{c,k_p} \cdot d_{c,k_q}, \qquad (13)$$

where (13) is one of the *n*-th order moment of the symbols for signal *c*. In addition, if we define $\ell = \mathrm{Card}(P_j)$ then we set

$$\Delta^j = \tfrac{1}{2\pi} \int_0^{2\pi} \exp\big(i(n - 2\ell)\phi_0\big) d\phi_0 = \delta_{n-2\ell,0} \qquad (14)$$

where $\delta_{\cdot,\cdot}$ is the Kronecker symbol. With these notations and replacing the signal model (11) in (8), we obtain

$$R_{d_c}^j(\underline{t}) = \sum_{k_1 \cdots k_n} m_{d_c}^j \Psi^j(\underline{t}) \qquad (15)$$

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

- CS: $\Psi^j(\underline{t}) = R_h^j(\underline{t}, 0)$ (16a)
- NCS: $\Psi^j(\underline{t}) = R_h^j(\underline{t}, 0) \cdot \Delta^j$ (16b)
- CA: $\Psi^j(\underline{t}) = \widetilde{R}_h^j(\underline{t})$ (16c)
- NCA: $\Psi^j(\underline{t}) = \widetilde{R}_h^j(\underline{t}) \cdot \Delta^j$. (16d)

Then, the correlation Γ_c^j becomes:

$$\Gamma_c^j = \sum_{k_1 k_2 \ldots k_n} m_{d_c}^j \int_{\underline{t}} \overline{R}_r^j(\underline{t}) \cdot \Psi^j(\underline{t}) d\underline{t}. \qquad (17)$$

Formula (13) (16) (17) give the way to implement the GMLC. Note that the same kind of derivation could have been done in the same way with offset (or staggered) modulations.

3.2 Application to M PSK / M' PSK classification

In the binary classification case (*c*=1 versus *c*=2) the GMLC test (8) is rewritten as:

$$\sum_{n=1}^{Q} \sum_{j=1}^{2^n} \sum_{k_1, \ldots, k_n} (m_{d_1}^j - m_{d_2}^j) \int_{\underline{t}} \overline{R}_r^j(\underline{t}) \cdot \Psi^j(\underline{t}) d\underline{t} \lessgtr th \qquad (18)$$

where *th* is the optimal threshold. To get a non-trivial test, the minimum value of *Q* has to verify the two following conditions:

$$\Psi^j(\underline{t}) \neq 0 \qquad (19)$$

and

$$m_{d_1}^j \neq m_{d_2}^j. \qquad (20)$$

It can be easily shown that for *M* PSK / *M*' PSK classification ($M' > M$) condition (20) is verified as soon as (13) contains *M*-th order moments of symbols $d_{c,k}$.

Let us define the partition j_0 of *M* elements such that $\ell = 0$ or $\ell = M$, and $k_1 = \cdots = k_M$. In a CS environment (19) is always verified and the smallest value of *Q* for which (20) is verified is $Q = M$. In this case the test (18) reduces to:

$$\mathrm{Re}\left[\int_{\underline{t}} \overline{R}_r^{j_0}(\underline{t}) \cdot \sum_k R_h^{j_0}(\underline{t}, 0) d\underline{t}\right] \gtrless th. \qquad (21)$$

We can see here that (21) is the expression of the correlation over $\underline{t}$ between the *M*-th order baseband pulse moment and the one estimated using instantaneous values of the observation for all cycle frequencies k/T, $k \in Z$. In the particular case where $h(t)$ is of duration *T*, the test

[4] « cohe. » stands for coherent, « sync. » for synchronous and « async. » for asynchronous.

(21) can be simplified and becomes

$$\sum_k \mathrm{Re}\left[\left(\int_{(k-1)T}^{kT} \bar{r}(t)\cdot h(t-(k-1)T)dt\right)^M\right] \lessgtr th \qquad (22)$$

which is the q_M test defined in [2] in a CS environment.

In a NCS case condition (19) imposes to have $n = 2\ell$. Under this constraint it comes that the smallest value of Q for which (20) is verified is $Q = 2M$ with partitions 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 (18) reduces to:

$$\left|\int_{\underline{t}} \overline{R}_r^{j0}(\underline{t})\cdot \sum_k R_h^{j0}(\underline{t},0)d\underline{t}\right|^2 \gtrless th\,. \qquad (23)$$

This result is well known in the frequency domain where all the cyclic 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*, one can verify that (23) simplifies to the optimal q_M-test defined in [2] for NCS environment[5].

Similarly, the GLMC tests in asynchronous cases (CA and NCA) are obtained by taking into account the integration (12) which leads to:

$$\text{CA: } \int_L \mathrm{Re}\left[\int_{\underline{t}} \overline{R}_r^{j0}(\underline{t}).R_h^{j0}(\underline{t},t_0)d\underline{t}\right]dt_0 \gtrless th\,, \qquad (24)$$

$$\text{NCA: } \int_L \left|\int_{\underline{t}} \overline{R}_r^{j0}(\underline{t}).R_h^{j0}(\underline{t},t_0)d\underline{t}\right|^2 dt_0 \gtrless th\,. \qquad (25)$$

Note here that the GMLC gives a theoretical demonstration of test (25) that have been proposed in [2] by empirical considerations and that the CA case (24) which was not treated in [2] is apparently new.

In the frequency domain, (24) can be seen as the correlation over the spectral frequencies $\underline{f}$ between the *M*-th order moment of the reference and the one estimated using instantaneous values of the observation at null cycle frequency. In the same way, it can be shown that (25) is the summation over cycle frequencies k/T, $k \in Z$ of the correlation modulus over $\underline{f}$ between the *M*-th order cyclic moment of the reference and the one estimated using instantaneous values of the observation. This result is in accordance with the fact that cyclic spectra are differently affected by a phase term arising from t_0.

[5] In the case of QAM classification, we note here that the choice $n = M$ and $\ell = M/2$ with $k_1 = \cdots = k_M$ will lead to another quasi optimal test in the NCS environment which reduces to the p_M-test [3] for $h(t)$ of duration *T*.

4 CONCLUSION

In this paper, we have presented a new approximation of the ALF for random finite mean power signals. It shows that the ALF can be viewed as a weighted sum of the correlations between all moments - at all orders - of the reference signal and those estimated using instantaneous values of the observation. It also provides a general theoretical framework for many pattern recognition based systems exploiting cyclostationarity of the modulated signals for classification as well as detection that have been proposed for the past ten years. Although this approximation is applied in this paper to classification, it can also apply to any ML estimation problems involving telecommunication signals such as synchronization.

Then we have applied this approximation for the classification problem, which leads to the General Maximum Likelihood Classifier. For linear modulation classification problem, the GMLC doesn't need any restriction on the pulse duration. For a pulse of symbol duration q_M and p_M-tests obtained in [1-3] can be retrieved by formal calculus.

REFERENCE

[1] K. Kim, A. Polydoros, "Digital modulation classification: the BPSK versus QPSK case", IEEE MILCOM conf., 1988.

[2] C.Y. Huang, A. Polydoros, "Likelihood methods for MPSK modulation classification", IEEE Tr. on Communications, vol.43, n°2/3/4, april 1995.

[3] C.S. Long and al, "Further results in likelihood classification of QAM signals", IEEE MILCOM conf., 1994.

[4] P.C. Sapiano, J.D. Martin, "Maximum likelihood PSK classifier", IEEE MILCOM conf., 1996.

[5] C.M. Spooner, "Classification of co-channel communication signals using cyclic cumulants", 29th Asilomar conf.on SSC, 1996.

[6] B.F. Beidas, C.L. Weber, "Higher-order correlation-based approach to modulation recognition of digitally frequency-modulated signals", IEEE Journal on Selected Areas in Communications, vol.13, n°1, january 1995.

[7] D. Boiteau, C. Le Martret, "A general maximum likelihood framework for modulation classification", IEEE ICASSP conf., 1998.