Which Constant Modulus Criterion is Better for Blind Adaptive Filtering : CM(1,2) or CM(2,2) ?

M.Bellanger
CNAM, 292 rue Saint-Martin, 75141 Paris cedex 03, France (bellang@cnam.fr)

ABSTRACT

The two most important constant modulus criteria are studied and compared, exploiting recently obtained results. A theoretical analysis of the performance is provided and excess output MSE figures are derived. The answer to the title question is found to depend on the output error power. In applications where the lower bound for the output signal-to-noise ratio is small, typically less than 8 dB, the CM(2,2) criterion can be employed. Otherwise, the CM(1,2) criterion is preferable. This result can be of great help to system designers, to select the constant modulus criterion that best suits their application and their performance objectives.

I – INTRODUCTION

Constant modulus algorithms can be used in adaptive filtering for channel equalization, interference cancellation or source separation, in applications like data communications, frequency modulation broadcasting or radar. Two criteria are realistic for implementation in systems, they are denoted by CM(1,2) and CM(2,2) and they are associated with the following cost functions

$$J_{CM12} = E[(1-|y_n|)^2] \qquad (1)$$

and

$$J_{CM22} = E[(1-|y_n|^2)^2] \qquad (2)$$

where y_n denotes the complex-valued filter output.

The theoretical analysis of the performance of adaptive filters based on these criteria is difficult, due to the modulus operation involved. Most of the work on this topic has been carried out with the CM(2,2) criterion, whose cost function admits a fourth order series development and can exploit appropriate statistical tools, like the kurtosis [1].

In contrast, no such possibility exists for the CM(1,2) criterion, which has received little attention, in spite of its attractiveness for implementation as pointed out in [2]. The corresponding coefficient updating equation for the adaptive filter is

$$H_{n+1} = H_n + \delta(y_n/|y_n| - y_n)X_n^* \qquad (3)$$

with H_n the coefficient vector and X_n^* the complex conjugate of the input data vector. This is just like the equation of the LMS algorithm with step size δ, except that the reference sequence is $y_n/|y_n|$. There is no high order term in the updating equation and an RLS-like algorithm can be used to accelerate the convergence and improve performance.

The purpose of the present paper is to exploit a set of critical recent results to refine existing estimations, assess the performance of the CM(1,2) criterion and provide a comparison of the two constant modulus criteria. The first of these recent results concerns the collinearity property : the optimal coefficient vector obtained with the CM(2,2) criterion is almost collinear with the vector obtained with the minimum mean square error (MMSE) criterion, when the output error power is small [3]. Next, an upper general bound was derived for the excess output error [4]. Then, a simple and efficient approximation of the CM(2,2) cost function, which paved the way for an accurate estimation of the excess MSE, was introduced [5]. Finally, a link was established between the two criteria [6].

The paper is organized as follows. The Suyama-Attux approximation of the CM(2,2) cost function is recalled in section 2 and an extension is derived in section 3. In section 4, the connection with the CM(1,2) cost function is developed and performance estimations are derived. Simulation results are given in section 5. In the final discussion, the application areas of the two criteria are explicited and commented upon.

2 – THE SUYAMA-ATTUX APPROXIMATION

The reasoning is as follows. Since the filter coefficient vectors obtained with the MMSE and CM(2,2) criteria are nearly collinear when the output error is small, since the excess MSE is bounded by the square of the output error, which means that it is very small in this case, then the CM(2,2) cost function can be expressed accurately in terms of J_{MMSE}, the cost function associated with the MMSE criterion. However, there is a fundamental difference between the two cost functions, namely J_{CM22} is insensitive to the sign of the coefficient vector. Thus, the correct approximation is the product

$$J_{CM22} \approx K\, J_{MMSE}(H)\, J_{MMSE}(-H) = K\, P \qquad (4)$$

with K a scalar to be determined. It is the starting point to derive the relationship between the coefficient vectors.

In order to express the MMSE solution, the reference signal samples s_n are assumed to have a constant and unit modulus and the output error is defined by

$$e_n = s_n - y_n = s_n - H^t X_n \qquad (5)$$

with H the filter coefficient vector. The cost function is

$$J_{MMSE} = E[|e_n|^2] = E[|s_n - y_n|^2] \qquad (6)$$

Now, the CM(2,2) cost function also must be expressed in terms of the output error. Consider the quantity

$$A = (s_n - y_n)(s_n^* - y_n^*)(s_n + y_n)(s_n^* + y_n^*) \qquad (7)$$

where y_n^* is the complex conjugate of y_n. To show the relationship with CM(2,2), it can also be written as

$$A = (1-|y_n|^2)^2 - (s_n y_n^* - s_n^* y_n)^2 \quad (8)$$

Returning to equation (7) and using definition (5) for the output error sequence, the quantity A is expressed by

$$A = 4|e_k|^2 - 4|e_k|^2 Real\{s_{k-d} e_k^*\} + |e_k|^4 \quad (9)$$

Then, the following series development is obtained for the cost function

$$J_{CM22} = 2E[|e_n|^2] - 4E[|e_n|^2 Real\{s_n e_n^*\}] + E[|e_n|^4] + 2E[Real\{(s_n e_n^*)^2\}] \quad (10)$$

As concerns the product P in (4), it takes the form

$$P = E[|e_n|^2]E[(2s_n - e_n)(2s_n^* - e_n^*)] \quad (11)$$

and, invoking the constant modulus property of s_n

$$P = 4E[|e_n|^2] + (E[|e_n|^2])^2 - 4E[|e_n|^2]E[Real\{s_n e_n^*\}] \quad (12)$$

Finally, the relation between the two cost functions is

$$J_{CM22} = \frac{1}{2} J_{MMSE}(H) J_{MMSE}(-H) + 2E[|e_n|^2]E[Real\{s_n e_n^*\}] - 4E[|e_n|^2 Real\{s_n e_n^*\}] + E[|e_n|^4] + 2E[Real\{(s_n e_n^*)^2\}] - \frac{1}{2}(E[|e_n|^2])^2 \quad (13)$$

Now, in the next section it is shown that the term $E[Real\{s_n e_n^*\}]$ is a second order term. Thus, in the above expression of J_{CM22}, all the terms are 4th order terms with respect to the error sequence e_n, except the first one and a first order approximation is

$$J_{CM22} \approx \frac{1}{2} J_{MMSE}(H) J_{MMSE}(-H) \quad (14)$$

In order to obtain the optimal coefficient vector, a derivation operation is carried out. To that purpose, the cost function (6) is rewritten by expanding $|e_n|^2$, which yields (15)

$$J_{MMSE}(H) = 1 + H^{*t} R_{X^*X} H - H^{*t} P_{sX^*} - H^t P_{sX^*}^*$$

with $R_{X*X} = E[X_n^* X_n^t]$ and $P_{sX^*} = E[s_n X_n^*]$

Next, the gradient of the MMSE cost function is

$$Grad[J_{MMSE}(H)] = 2(R_{X*X} H - P_{sX^*}) \quad (16)$$

After substitution of (15) into (14), derivation and some algebraic manipulations, it appears that the MMSE and CM(2,2) coefficient vectors are collinear and related by

$$H_{CM22} = \frac{2Real\{H_{CM22}^{*t} P_{sX*}\}}{1 + H_{CM22}^{*t} R_{X*X} H_{CM22}} H_{MMSE} \quad (17)$$

Now, the ratio $a = H_{CM22} / H_{MMSE}$ can be expressed in terms of the output mean square error denoted E_0

$$E_o = 1 - H_{MMSE}^{*t} R_{X*X} H_{MMSE} = 1 - H_{MMSE}^{*t} P_{sX*} \quad (18)$$

Substituting equation (18) into equation (17) leads to

$$a = \frac{2a[1-E_0]}{1 + a^2[1-E_0]} \quad (19)$$

Hence, the first order approximation of the proportionality factor

$$a = H_{CM22} / H_{MMSE} \approx 1 - E_o / 2 \quad (20)$$

Thus, the MMSE and CM(2,2) coefficient vectors are collinear and close when the MMSE output error is small and the first order approximation (14) is valid. As concerns the CM (2,2) output error power, it is given by the classical expression

$$E_{CM22} = E_o + (H_{CM22}^{*t} - H_{MMSE}^{*t}) R_{X*X} (H_{CM22} - H_{MMSE}) \quad (21)$$

Then, using (20) and (18),

$$E_{CM22} = E_0 + (1-a)^2(1-E_0) \quad (22)$$

and the approximation with two terms is

$$E_{CM22} \approx E_0 + E_0^2 / 4 \quad (23)$$

This estimation is well below the general bound given in [4], which is approximately E_0^2.

3 – REFINING THE ESTIMATIONS

In order to obtain higher order terms in the above approximations, the fourth order terms in equation (13) have to be included in the estimation procedure and they have to be related to $(E[|e_n|^2])^2$ in the vicinity of the optimum. To begin with, the statistics of the variable $u_n = s_n e_n^*$ have to be analysed and the mean is

$$m_u = E[u_n] = E[s_n e_n^*] = 1 - E[s_n X_n^{*t} H^*] \quad (24)$$

Next, invoking the collinearity in the vicinity of the optimum, that is taking $H = aH_{MMSE}$, we obtain

$$m_u = E[s_n(s_n^* - y_n^*)] = 1 - a(1-E_0)$$
$$m_u = E_0 + (1-a)(1-E_0) \quad (25)$$

Comparing (22) and (25), it is observed that, as long as the ratio a is smaller than unity, m_u is greater than the output error power E_{CM22} and the difference is

$$m_u - E[|e_n|^2] = a(1-a)(1-E_0) \quad (26)$$

The real and imaginary parts of e_n^* are uncorrelated, zero-mean variables and they have the same power. It is worth pointing out that m_u is real and the rotation of e_n^* performed by s_n has introduced a small bias on the real part. Considering that the optimum is close to the first order approximation, the following approximation is made

$$E[Real\{(s_n e_n^*)\}] \approx E[|e_n|^2]3/2 \qquad (27)$$

A similar approach can be taken to study the other fourth order terms in the expression (13) of the CM(2,2) cost function. Finally, summing up all the terms, the following global approximation is obtained

$$J_{CM22} \approx \frac{1}{2} J_{MMSE}(H) J_{MMSE}(-H) + 3(E[|e_n|^2])^2 \qquad (28)$$

Now, using equation (12) and the product P, the symmetrical second order approximation of the cost function is

$$J_{CM22} \approx \frac{1}{2} P + \frac{3}{16} P^2 \qquad (29)$$

The second order term in the above expression generates additional terms in the expression of the coefficient vector ratio, which is, after a number of simplifications

$$a = \frac{H_{CM22}}{H_{MMSE}} \approx \frac{2a[1-E_0]+6E_0}{1+a^2[1-E_0]+6E_0} \qquad (30)$$

The second order approximation is

$$a \approx 1 - \frac{1}{2} E_o + \frac{7}{8} E_0^2 = 1 - \frac{1}{2} E_0 [1 - \frac{7}{4} E_0] \qquad (31)$$

The corresponding output error power is obtained by substituting (31) into (22) which yields

$$E_{CM22} \approx E_0 + \frac{1}{4} E_0^2 [1 - \frac{9}{2} E_0] \qquad (32)$$

Since the excess mean square error in (32) cannot take on negative values, it is clear that this second order approximation can only be valid for $E_0 < 2/9$.

4 – STUDY OF THE CM(1,2) CRITERION

The cost function (1) can be related to the cost function of the CM(2,2) criterion, using the expression

$$J_{CM12} = E[\frac{(1-|y_n|^2)^2}{(1+|y_n|)^2}] \qquad (33)$$

which leads to (34)

$$J_{CM12} = \frac{1}{4} J_{CM22} + E[(1-|y_n|)^3] - \frac{1}{4} E[(1-|y_n|)^4]$$

Obviously, no series development with a finite number of terms can be found for this function. However, a fourth order approximation can be readily obtained since the term $(1-|y_n|)$ is linked to the error sequence, through

$$1 - |y_n| \approx Real\{s_n^* e_n\} \qquad (35)$$

Therefore, the last term in the right-hand side of equation (34) is a fourth order term while the middle term reflects the difference between the coefficient vectors corresponding to the two criteria. In order to assess that difference, an analysis of the output error signal is performed.

The derivation of the cost function J_{CM12} with respect to the filter coefficient vector leads to the equation that defines the filter coefficient vector

$$E[(1/|y_n| - 1) y_n X_n^*] = 0 \qquad (36)$$

Therefore, the coefficient vector H_{CM12} is such that :

$$R_{X*X} H_{CM12} = E[y_n X_n^*] = E[\frac{y_n}{|y_n|} X_n^*] \qquad (37)$$

Derivation of equation (2) leads to a similar equation for the CM(2,2) criterion

$$R_{X^*X} H_{CM22} = E[y_n X_n^*] = E[|y_n|^2 y_n X_n^*] \qquad (38)$$

For the sake of completion, the same presentation for the MMSE criterion is

$$R_{X*X} H_{MMSE} = E[y_n X_n^*] = E[s_n X_n^*] \qquad (39)$$

Now, the equalizer output y_n can be related to the reference signal s_n by

$$y_n = s_n (1 + r_n) e^{j\theta_n} \qquad (40)$$

where r_n and θ_n are the deviations in amplitude and phase respectively. If the adaptive filter performs well, the output error is small and a first order approximation can be made

$$y_n \approx s_n + (r_n + j\theta_n) s_n \qquad (41)$$

and the MMSE condition (39) takes the form

$$E[(r_n + j\theta_n) s_n X_n^*] = 0 \qquad (42)$$

Introducing the approximation (41) into equation (38) leads to the following condition for the CM(2,2) criterion

$$E[y_n X_n^*] = E[(1 + 3r_n + j\theta_n) s_n X_n^*] \qquad (43)$$

Similarly, for the CM (1,2) criterion, using (37)

$$E[y_n X_n^*] = E[(1 + j\theta_n) s_n X_n^*] \qquad (44)$$

The magnitude deviation r_n is not involved in (44) while it comes with the factor 3 in (43). Therefore, the CM(1,2) equalizer coefficient vector is closer to the MMSE vector than the CM (2,2) vector. Now, if the MMSE E_0 is small, the 3 equalizers produce nearly the same output y_n and the variables r_n and θ_n can be considered to be the same. Then, substituting (42) into (43) leads to

$$R_{X*X} H_{CM22} = R_{X*X} H_{MMSE} + 2E[r_n s_n X_n^*] \qquad (45)$$

and, exploiting the results obtained in the previous section and introducing equation (20), the deviation term is approximated

$$2E\left[r_n s_n X_n^*\right] \approx -\frac{1}{2} E_0 R_{X*X} H_{MMSE} \qquad (46)$$

Next, turning to the CM(1,2) criterion and combining expressions (44), (42) and (46) leads to the following relation between the coefficient vectors

$$H_{CM12} \approx H_{MMSE}\left[1 + E_0 / 4\right] \qquad (47)$$

Accordingly, the output MSE is approximated by

$$E_{CM12} \approx E_0 + E_0^2 / 16 \qquad (48)$$

Clearly, the CM(1,2) equalizer comes very close to the MMSE solution when the output error power is small. The validity domain of the above estimations can be assessed by analysing the fourth order term in (34), which is subtracted from the fourth order terms contained in expression (13). Clearly, the validity domain of estimations (48) and (47) is likely to be large and this is confirmed by simulations.

5 – SIMULATION RESULTS

An equalizer is considered with QPSK uncorrelated and unit power symbols. The channel transfer function is

$$C(Z) = 1 + 0.1 c_1 e^{j\pi/4} Z^{-1} \qquad (49)$$

The equalizer output MSE is varied by changing the value of the real scalar c_1 The excess MSE values obtained for the CM(2,2) criterion with a gradient algorithm are shown in fig.1. The equalizer has N=3 coefficients and the MMSE solution is computed from theory, according to the definitions, with no delay on the reference sequence. The general bound is represented in the figure. For $c_1 = 9$ the output MMSE is $E_0 = 0.177$, which is smaller than the validity limit 2/9 = 0.222 in equation (32). In contrast, for $c_1 = 9.5$, $E_0 = 0.213$ and it is close to the limit. The results obtained with CM(1,2) are shown in fig.2. The estimation remains accurate for large values of the MMSE output error power, as predicted.

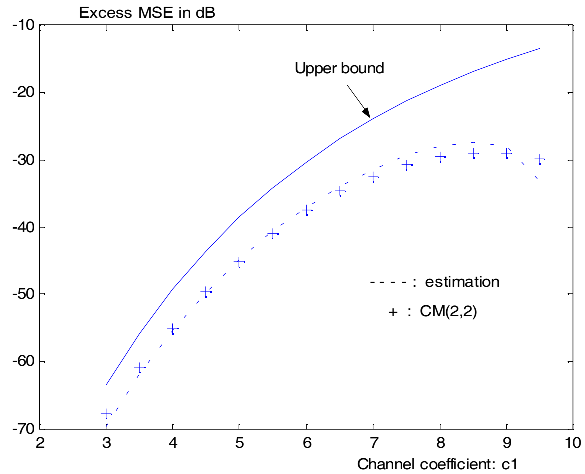

Fig.1. Excess MSE for the CM(2,2) criterion

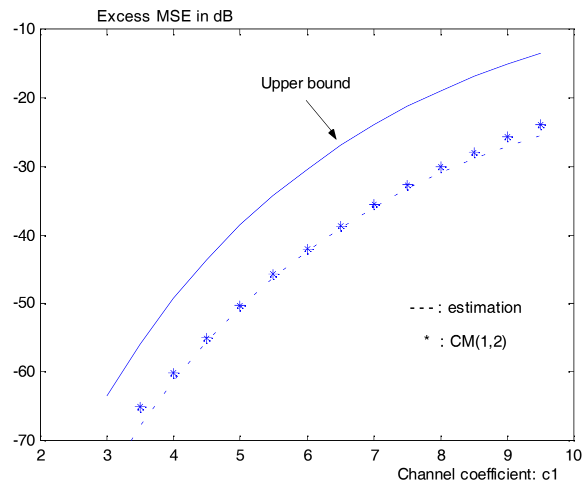

Fig.2. Excess MSE for the CM(1,2) criterion

6 – CONCLUSION

For small values of the mean square output error, the filter coefficient vectors associated with the MMSE, CM(2,2) and CM(1,2) criteria are nearly collinear. With respect to MMSE, the proportionality factor is less than one for CM(2,2) and greater than one for CM(1,2).

Considering (32) and (48), it appears that the excess MSE values obtained with the two criteria are equal if $E_0 = 1/6$ which corresponds to an output SNR close to 8 dB. This can serve as a delimitation for the application areas of the two criteria. In applications like QPSK data transmission, the lower bound for the SNR at the output of the equalizer is generally taken greater than 10 dB, to achieve a bit error rate smaller than 10^{-3} and then, criterion CM(1,2) can be recommended. The same applies to frequency modulation with the threshold effect.

From a research perspective, the above results illustrate the potential of the CM(1,2) criterion for blind adaptive filtering and more effort should devoted to study further the properties of this criterion and related algorithms.

References

[1] C.R.Johnson et al., « Blind Equalization Using the Constant Modulus Criterion : a Review », Proceedings of the IEEE, Vol.86, N°10, pp.1927-1950, 1998.

[2] O.W.Kwon, C.K.Un and J.C.Lee, « Performance of Constant Modulus Adaptive Filters for Interference Cancellation », Signal Proc.,V26-2, Feb.1992,pp.185-96.

[3] H.H.Zeng, L.Tong and C.R.Johnson, « An Analysis of Constant Modulus Receivers », IEEE trans. on Signal Processing, Vol.47, N°11, Nov.1999, pp.2990-2999.

[4] P.Schniter and C.R.Johnson,« Bounds for the MSE Performance of Constant Modulus Estimators », IEEE Transactions on Information Theory, Vol.46, N°7, November 2000, pp.2544-2560.

[5] R.Suyama,R.Attux,J.M.Romano and M.Bellanger, « On the Relationships Between Wiener and Constant Modulus Criteria for Adaptive Filtering », Proc. of Asilomar Conference, 10-12 Nov.2003.

[6] M.Bellanger, « A Simple Comparison of Constant Modulus and Wiener Criteria for Equalization with Complex Signals », Elsevier, Digital Signal Processing 14 (2004), pp.429-437.