Some Reflections on the 35 Years of the Constant Modulus Criterion

Ricardo Suyama, Romis Attux, Denis Fantinato and João Marcos T. Romano

Abstract—**This work, which is intended to be a celebration to the 35 years of the constant modulus criterion and to the impact it had in the development of our research group, presents both tutorial elements and a discussion of published and unpublished results that, hopefully, will generate new reflections and perspectives on this important topic.**

Keywords—*Unsupervised equalization, constant modulus criterion, CMA.*

I. INTRODUCTION

"To see a world in a grain of sand...", William Blake

This work is dedicated to Dominique Godard.

The problem of single-input / single-output (SISO) linear channel equalization can be stated as follows: a signal $s(n)$ is transmitted by means of a communication channel with impulse response $h(n)$. It is received as a distorted (and also possibly noisy) version $x(n)$. A filter, called *equalizer*, is used to compensate for the aforementioned distortions in order to recover an estimate of $s(n)$ (up to delay and scaling modifications).

Considering that the equalizer is a linear filter, there remains the task of adapting its free parameters. This is usually done by defining a criterion and optimizing an associated cost function. If the criterion includes information about samples of the desired signal (or other "strong" priors), it is said to be *supervised*. If it is solely based on more general statistical information, it is termed *unsupervised* or *blind*.

In this work, we discuss that which is probably the most studied blind approach: the constant modulus criterion, which was effectively proposed 35 years ago and still remains a challenging and prominent research topic. The work is a "strange" creature, in that it is neither exactly a tutorial nor a report of novel results, but a mixture of both, in a personal (and sentimental) view on a technique that has been of great importance in the shaping of us all as researchers and of our group. We planned it to be a tribute to the pioneer Dominique Godard and an invitation to all interested researchers to take part in the search for a deeper understanding of blind methods.

The work is structured as follows. In Section II, Godard's original proposal is recapitulated. In Section III, some remarks about the constant modulus cost are presented: the indicative curves of the null gradient, the polynomial formulation and a novel initialization heuristic. Section IV provides further investigation concerning Ding's minima. Some extensions are described in Section V. Finally, the conclusions of the work are summarized in Section VI.

II. THE CONSTANT MODULUS CRITERION

The constant modulus criterion first appeared in a paper by Dominique Godard, called "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", published in 1980 [1]. A central point of the proposal is to have the removal of intersymbol interference (ISI) (and noise) independently of the data constellation and of phase recovery. This explains the generic cost function below:

$$J_p(\mathbf{w}) = E\left[\left(R_p - |y(n)|^p\right)^2\right], \tag{1}$$

where $\mathbf{w}$ is the equalizer parameter vector, $y(n)$ is the equalizer output and $E[\cdot]$ is the statistical expectation operator.

Notice that $J_p(\cdot)$ is a sort of dispersion metric, and that it focuses exclusively on the modulus of the equalizer output $y(n)$. The value of R_p is shown to be:

$$R_p = \frac{E[|s(n)|^{2p}]}{E[|s(n)|^p]} . \tag{2}$$

When $p = 2$, one obtains the constant modulus (CM) criterion, being the name associated with the work of Treichler and Agee [2]. For constant amplitude modulation schemes, like the PSK modulation, the calculation of R_2 is trivial.

Using the stochastic gradient, the constant modulus algorithm (CMA) is obtained from the steepest descent approach (henceforth we will consider only real-valued signals and systems):

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu[R_2 - y^2(n)]y(n)\mathbf{x}(n) \tag{3}$$

where $\mathbf{x}(n)$ is the input vector and $\mathbf{w}(n)$ is the equalizer parameter vector at the *n-th* iteration.

In his original work, Godard showed that the criterion was consistent in that it had only zero-forcing solutions for the case of a doubly-infinite structure (i.e. in a condition of perfect inversion). Later, due to the work of Ding, Li, Johnson, Regalia and others, the properties of the CM criterion for the case of finite equalizers were studied in depth [3-6].

In this work, we present a personal perspective of the CM criterion, which will comprise views regarding the CM cost function developed by the research group, like the curves associated with the null gradients and the polynomial formulation (including a novel initialization heuristic), an original appreciation of Ding minima and, finally, an overview of some CM extensions along its 35 years of existence.

III. THE CM COST FUNCTION: SOME REMARKS

If one consider real-valued signals and systems and binary (+1/-1) independent and identically distributed (i.i.d.) signals, the CM cost function becomes:

$$J_{CM} = E\left[(1 - y^2(n))^2\right]$$
$$= 1 - 2E[y^2(n)] + E[y^4(n)]. \tag{4}$$

Taking the gradient of (4) and equaling it to the null vector, one obtains:

$$\mathbf{Rw} = E[y^3(n)\mathbf{x}(n)] \tag{5}$$

where $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}(n)^T]$ is the autocorrelation matrix of the signal vector $\mathbf{x}(n)$. Let us now recall the form of the Wiener-Hopf equations:

$$\mathbf{Rw} = \mathbf{p} \tag{6}$$

where $\mathbf{p} = E[\mathbf{x}(n)d(n)]$. It can be seen that the role of the Bussgang estimator [7] is to use, in a certain sense, the correlation between $y^3(n)$ and $\mathbf{x}(n)$ as a substitute for the cross-correlation between $s(n - d)$, being d an equalization delay, and $\mathbf{x}(n)$. There are two points here worthy of attention: $y(n)$ can, potentially, approximate the transmitted signal at any delay, which means that the CM criterion will tend to accommodate itself to the structure of the channel [8]. Another point is that the fact that $y(n)$ depends on $\mathbf{w}$ creates an equation of third order that is difficult to treat. Let us try to analyze it in more detail for a "simple" – but non-trivial – case, that of a two-tap equalizer.

Firstly, it can be shown that (4) becomes:

$$\begin{aligned}
J_{CM}(\mathbf{w}) = {} & 1 - 2E[w_0^2 x^2(n) + 2w_0 w_1 x(n)x(n-1) \\
& + w_1^2 x^2(n-1)] + E[w_0^4 x^4(n) \\
& + 4w_0^3 w_1 x^3(n)x(n-1) \\
& + 6w_0^2 w_1^2 x^2(n)x^2(n-1) \\
& + 4w_0 w_1^3 x(n)x^3(n-1) + w_1^4 x^4(n-1)].
\end{aligned} \tag{7}$$

The equations for the gradient are:

$$\begin{aligned}
\frac{\partial J_{CM}}{\partial w_0} = {} & E[-4w_0 x^2(n) - 4w_1 x(n)x(n-1) \\
& + 4w_0^3 x^4(n) + 12w_0^2 w_1 x^3(n)x(n-1) \\
& + 12w_0 w_1^2 x^2(n)x^2(n-1) \\
& + 4w_1^3 x(n)x^3(n-1)] = 0
\end{aligned} \tag{8}$$

and

$$\begin{aligned}
\frac{\partial J_{CM}}{\partial w_1} = {} & E[-4w_1 x^2(n-1) - 4w_0 x(n)x(n-1) \\
& + 4w_1^3 x^4(n-1) + 12w_0 w_1^2 x(n)x^3(n-1) \\
& + 12w_0^2 w_1 x^2(n)x^2(n-1) \\
& + 4w_0^3 x^3(n)x(n-1)] = 0.
\end{aligned} \tag{9}$$

This system of equations is quite tricky, and we will not attempt to solve it directly. However, it is possible to use both equations to eliminate either the term w_0^3 or the term w_1^3, creating a second-order equation on the respective variable, which can be solved using Bhaskara's formula. Figure 1 shows both equations as curves and their intersection generates all the minima, the saddle points and the maximum at the origin.
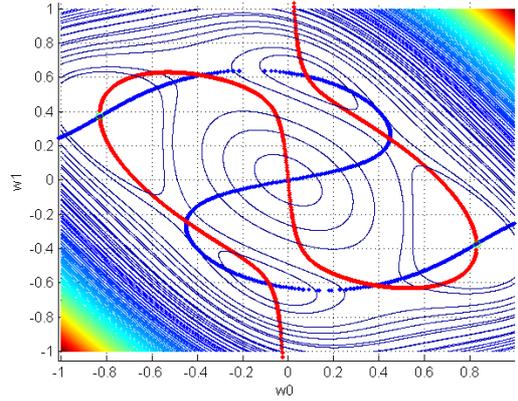

Fig. 1. Contours of the CM Cost Function and both Curves that Generate a Null Gradient

Perhaps an additional substitution could lead to formulas for generating the values of w_0 and w_1, but it would probably bring little insight regarding the nature of these solutions due to its complexity. The symmetry of these curves, anticipated in the structure of (8) and (9), seems to reveal this nature more elegantly albeit in a less formal way.

A. Polynomial Formulation of the CM Criterion and the Derived Lower Bound

In [9], a polynomial formulation of the CM criterion was proposed. It is based on considering, in (4), $\upsilon(n) = y^2(n) = w_0^2 x^2(n) + 2w_0 w_1 x(n)x(n-1) + w_1^2 x^2(n-1)$, which generates a constrained problem based on the cost function:

$$J_{CM} = E\left[(1 - \upsilon(n))^2\right]. \tag{10}$$

If the constraint regarding the weights is relaxed, it is possible to modify $\upsilon(n)$ so that it becomes:

$$\begin{aligned}
\upsilon_r(n) = {} & g_0 x^2(n) + g_1 x(n)x(n-1) \\
& + g_2 x^2(n-1).
\end{aligned} \tag{11}$$

The solution to the problem becomes then analytical, a Wiener solution with $\mathbf{R}_r = E[\mathbf{x}_r(n)\mathbf{x}_r(n)^T]$ and $\mathbf{p}_r = E[\mathbf{x}_r(n)]$, where

$$\mathbf{x}_r(n) = [x^2(n) \quad x(n)x(n-1) \quad x^2(n-1)]^T. \tag{12}$$

This means that $J_{CM} \geq 1 - \mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r$, being equality reached, for sure, when perfect channel inversion is possible, i.e., it is possible to have $J_{CM} = 0$.

The lower bound $1 - \mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r$ is an interesting result, as it seems to work as a sort of blind estimator of "invertibility" of the channel (or the mixing system, if a MIMO case is considered). By performing estimates exclusively over the received signal, the index $1 - \mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r$ provides the user with a notion of the performance an unsupervised algorithm will attain in the case of interest, *independently of any a priori choice of the equalization delay* [9].

This formulation also allows a beautiful geometrical interpretation concerning any search process over the CM cost function: it is as if the algorithm were performing a search over a Wiener-like solution having to obey a constraint that relates g_0, g_1 and g_2. We may use, for instance, $g_1^2 = 4g_0 g_2$, resulting in the surfaces shown in Figure 2. In a certain sense, this is a sort of attainable set [3] associated with
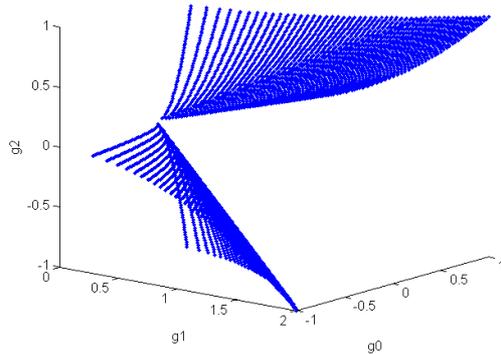
Fig. 2. Domain of Attainable Solutions

the CM for the two-tap case. Any version of the constant modulus algorithm will aim to reach the minimum of a paraboloid, but it will not be allowed to leave this surface.

B. A Novel Initialization Heuristic

The unconstrained Wiener solution for the polynomial filter is a useful tool to initialize the CMA, as shown in [9]. Here we propose a new initialization heuristic based thereon, and we shall compare it to the center-spike method, to a random initialization and to the heuristic proposed in [9], here referred to as *Crossed Terms Heuristic* (CTH). We will only consider the two-tap linear equalizer case for the sake of simplicity.

Firstly, let us keep in mind that, in this scenario, $g_0 = w_0^2$, $g_1 = 2w_0w_1$ and $g_2 = w_1^2$. Our idea is to use $sign(g_1)$ to define whether w_0 and w_1 have equal or different signs (w_0 will be assumed positive without loss of generality) and to create the following solutions:

$$w_{0,1} = \sqrt{g_0}, \qquad w_{1,1} = sign(g_1)\sqrt{g_2}$$

$$w_{0,2} = \sqrt{g_0}, \qquad w_{1,2} = \frac{g_1}{2\sqrt{g_0}} \qquad (13)$$

$$w_{0,3} = \frac{|g_1|}{2\sqrt{g_2}}, \qquad w_{1,3} = sign(g_1)\sqrt{g_2}$$

Finally, the initial solution will be an average of the three vectors.

In Table I, we have the global convergence rate and the average CM cost obtained for all four methods, considering the channel with transfer function $H(z) = 1 + \alpha z^{-1} + \beta z^{-2}$, being α and β varied from 0 to 2, in steps of 0.041. It can be seen that, in terms of the global convergence rate, the classical center-spike method performs poorer than all considered methods. On the other hand, the CTH and the average of the three vectors perform similarly, almost reaching 95% of the global convergence rate, with a slight advantage for the CTH.

TABLE I.   GLOBAL CONVERGENCE RATE AND AVERAGE CM COST.

Initialization	Conv. Rate	Avg. CM Cost
Center-Spike	31.99%	0.4599
Random	39.61%	0.4558
CTH	94.63%	0.4044
3 Vectors Avg.	93.88%	0.4046

In terms of the average CM cost, we observe that the higher the global convergence rate, the lower the CM cost. Indeed, it is clear that there is a reduction of the average CM cost from the center-spike and random methods to the CTH and the 3 vectors average methods.

IV.   DING MINIMA REVISITED

In the process of consolidating the analysis of the CMA, it was necessary to pass from the original analysis in the combined space performed by Godard [1] to an analysis in the parameter space. In the parameter space, Ding et al. [3] made an important analytical contribution in showing that for an all-pole channel, the CM criterion, aside from the global optimum (perfect inversion), would have "shallow" local minima, unable to reduce any intersymbol interference.

In some circles, this led to the belief that local minima in blind criteria were necessarily a drawback, even with the clarifying work of Johnson et al. [5]. To us, it soon becomes clear that this was not true: indeed, the Wiener solution is unique for a given equalization delay, but, for different delays, different degrees of performance are achieved [10]. The local minima of the CMA, for FIR channels, are often close to Wiener solutions associated with different delays, which can hardly be considered a drawback.

Another point is that there are two original results of our group that put Ding minima under a new light: they occur in the Wiener criterion if the delay is a degree of freedom and the polynomial formulation exposed above guarantees, for all-pole channels, perfect initialization, making their effect null. Both points will now be looked at in more detail.

A. Ding Minima in the Wiener Criterion

Let us consider the equalization of an AR(N) channel given by the following transfer function:

$$H(z) = \frac{1}{1 + \alpha z^{-N}}$$
$$= 1 - \alpha z^{-N} + \alpha^2 z^{-2N} - \alpha^3 z^{-3N} + \cdots \qquad (14)$$

For this particular channel, the values of the autocorrelation $r(n)$ of the channel output $x(n)$ are given by

$$r(n) = \begin{cases} \dfrac{(-\alpha)^n}{1 - \alpha^2} & n = 0, N, 2N, \ldots \\ 0 & \text{otherwise} \end{cases} \qquad (15)$$

Evaluating the Wiener solution for this AR(N) channel, one would observe that the inverse of the correlation matrix $\mathbf{R}$ can be computed as

$$\mathbf{R}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \alpha \\ 0 & 1-\alpha^2 & 0 & \ddots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & \ddots\, 0 & 0 & 1-\alpha^2 & 0 \\ \alpha & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad (16)$$

and the cross-correlation vectors $\mathbf{p}$ for the first $N+1$ values of the equalization delay are given by

$$\mathbf{p}_{d=0} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}^T$$
$$\mathbf{p}_{d=1} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}^T \qquad (17)$$
$$\vdots$$
$$\mathbf{p}_{d=N} = \begin{bmatrix} -\alpha & 0 & \cdots & 0 & 1 \end{bmatrix}^T.$$

These vectors give rise to the following Wiener solutions:

$$\begin{aligned}
\mathbf{w}_{d=0} &= [1 \quad 0 \quad 0 \quad \cdots \quad \alpha]^T \\
\mathbf{w}_{d=1} &= [0 \quad 1-\alpha^2 \quad 0 \quad \cdots \quad 0]^T \\
&\vdots \\
\mathbf{w}_{d=N} &= [0 \quad \cdots \quad 0 \quad 0 \quad 1-\alpha^2]^T
\end{aligned} \tag{18}$$

The first solution is the ideal case. The solutions for $0 < d < N$ have a single intermediate non-zero coefficient and, consequently, do not reduce ISI. If the delay is further increased, the pattern of bad solutions will repeat itself, with different non-zero coefficients.

So, we have just proved that undesirable minima are not exclusively found in the CM criterion, but also in the classical Wiener paradigm. It is remarkable that the latter has also revealed a new class of bad minima, derived from the interval $0 < d < N$.

Let us now visualize the relation between CM and Wiener "bad minima" in a simple AR(1) case, by posing $\alpha = 0.6$, and a MA(1) equalizer. Figure 3 establishes a comparison between CM and Wiener local minima. In this rather simple case, there are only ideal and bad minima, both of which can be identified from the contours. As expected, it is shown that the bad Wiener minima are collinear with the CM local minima. It is still possible to prove that the magnitude of the non-zero element of the Wiener minima decays exponentially with the equalization delay.
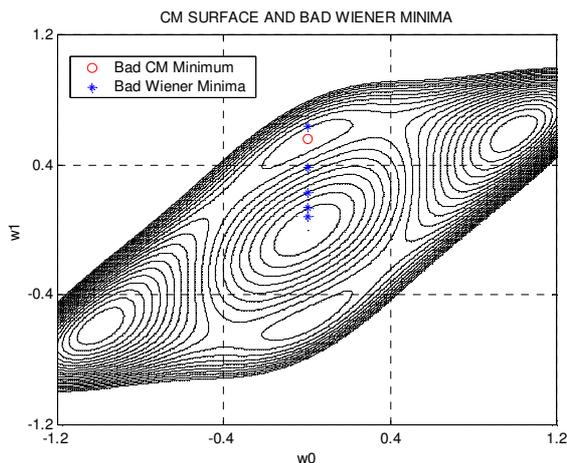

Fig. 3. CM Surface and Wiener Local Minima

B. Perfect Initialization

We will appeal to the reader's intuition to avoid unnecessary mathematical complications. If the channel is a stable and causal all-pole filter, this means that it can be inverted with a FIR filter. In this case, in the absence of noise (which is the standard hypothesis in this work), it will be possible to have $J_{CM} = 0$. This means that the lower bound $1 - \mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r$ presented in Section III-A will have to be null, and the derived relaxed solution will be, as a matter of fact, equivalent to the constrained solution. Therefore, by calculating the relaxed solution and using one of the vectors in (13), one will have a good estimate of the channel inverse in a blind fashion. Hence, the CMA will not be prone to local convergence in this particular case. Naturally, it will be necessary to have enough data to produce a good initial estimate, but, from a theoretical standpoint, Ding minima become innocuous within the CMA framework. Let us present an example analogous to the channel Ding used in his classical work [3]:

$$H(z) = \frac{1}{1 + 0.6z^{-1}} \tag{19}$$

The signal is composed of i.i.d. binary samples (+1/-1). Using only 10 samples, one already obtains $g_0 = 1$, $g_1 = 1.2$ and $g_2 = 0.36$, which is directly "decoded" – based on any vector of (13) – as $\mathbf{w} = [1 \quad 0.6]^T$.

V. SOME EXTENSIONS

So far, we presented some of the CM criterion perspectives studied by our research group. However, since Godard's proposal, there were several other remarkable CM-related efforts that are worth mentioning – many of them can be viewed as extensions of the CM criterion, as they are able to enlarge its scope and/or to improve its performance. For the sake of brevity, we will restrain ourselves to list a few of them, as follows.

We start by the *Least-Squares CMA* (LS-CMA), an extension of the original CMA [11]. The main idea is to employ an optimization method capable of accelerating the convergence rate of the algorithm when dealing with highly-dynamic environments. This is accomplished through the employment of the Gauss' method [11], which uses information of the Jacobian for more precise steps in the adaptation process. Other variations of this methods include the *Orthogonalized-CMA* (O-CMA) [12] based on the orthogonalization of the observed signals in order to increase convergence speed; the QR-CMA [13], which also employs a data prewhitening step through the QR decomposition in order to simplify the CMA iteration and increase convergence speed; and the *Recursive Least-Squares CMA* (RLS-CMA) [14], which explores an approximation of the CM cost function in order to enable the use of the RLS algorithm for the coefficients adaptation.

Following a different approach to increase the convergence speed, in [15][16] the authors present the *Optimal Step-size CMA*, or shortly, OS-CMA. This algorithm aims at a most effective choice for the step-size μ in the CMA, which can greatly contribute to a fast convergence and, very conveniently, to avoid the convergence to local solutions in certain cases [16]. Basically, by adopting the minimization of a cost function in a single search direction, a third-order degree polynomial in function of μ is obtained, whose roots can be extracted through procedures like the Cardano's formula. The root capable of attaining the lowest value of the cost function will be the optimum step-size. Hence, a few iterations of the CMA with the optimum step-size can quickly lead the algorithm to a global solution.

Other extensions tried to deal with the inherent phase ambiguity present in the CM cost function. In this sense, in [17], the authors proposed a *Modified CMA* (MCMA), based on the following cost function

$$\begin{aligned}
J_{MCMA} = &E[(Re\{y(n)\}^2 - R_R^2)^2] \\
&+ E[(Im\{y(n)\}^2 - R_I^2)^2]
\end{aligned} \tag{20}$$

The idea was to force the real and imaginary parts of the recovered signal to lie on a reduced four-point constellation, defined by constants R_R and R_I. This approach is very similar to other algorithms, such as the *Reduced Constellation Algorithm* (RCA) [18], and the *Multimodulus Algorithm* (MMA) [19].

Some modifications of CMA and MMA were proposed in the literature for dealing with nonsquare constellations. For

instance, in [20] the authors propose the *Radius Directed Equalization* (RDE), which is based on the CM cost function, but considering several dispersion constants: the complex case, in this case, is divided into annular decision regions, and the dispersion constant is a function of the region of the complex plane in which the recovered sample is located. The same idea was also explored for the MMA [21]: the pair of constants R_R and R_I depends on the region in which the estimated symbol is located.

A more recent development, devoted specially to the case of high-order constellations, was presented in [22], and jointly explores the ideas of RDE and MMA, named *Regional Multimodulus Algorithm* (RMA). The idea is to consider a partition of the complex plane into regions, each one containing 4 symbols (the constellation is viewed as a collection of 4-QAM constellations). Then, the adaptation rule will depend on which region the estimated symbol is located, and is performed as if the signal pertains to the 4-QAM constant modulus of the corresponding region.

When dealing with non-Gaussian noise, such as impulsive noise, the CMA is known for losing its performance [23]. In that sense, the algorithm referred to as *Fractional Lower-Order Statistics CMA* (FLOS-CMA) propose the use of statistics different from the fourth- and second-order to aid the equalization process [23][24]. Indeed, the classic statistics combined to the impulsive behavior of the noise can culminate in larger disturbances in the optimization process, mainly observed in stochastic gradient-based algorithms. By adopting fractional lower-order statistics of the equalizer output – e.g., $\|y(n)\|^p$, being $0 < p < 2$ –, the FLOS-CMA, is able to overcome this difficulty and still guarantee a good performance. The value of p, in this case, is a fractional that will depend on the shape of the noise distribution (assumed to be an alpha-stable distribution) [23].

Finally, an extension of the polynomial formulation (Section III-A) considers the method of Lagrange multipliers to obtain the analytical solutions of the CM criterion [25]. This approach is easily conducted for a two-tap equalizer, however, for larger filters, the Lagrangian formulation involves the solution of a non-linear system, which is possible to be achieved through the employment of an iterative algorithm.

VI. CONCLUSIONS

In this work, we discussed the foundations of the CM criterion and of the CMA, and presented a "personal" view on an arbitrary selection of published and novel (to the best of our knowledge) results. It is intended as a celebration of the 35 years of Godard's work and as a tribute to a research topic of the utmost relevance to the development of our research group.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", *IEEE Trans. on Communications*, v. COM-28(11), pp. 1867-1875, 1980.

[2] J. Treichler and B. Agee, "New Aproach to Multipath Correction of Constant Modulus Signals", *IEEE Trans. Acoust., Speech and Signal Processing*, v. ASSP-31(2), pp. 459-472, 1983.

[3] Z. Ding, R. Kennedy, B. Anderson and C. Johnson, "Ill-Convergence of Godard Blind Equalizers in Data Communication Systems", *IEEE Trans. on Signal Processing*, v. 39, pp. 1313-1327, 1991.

[4] Y. Li and Z. Ding, "Convergence Analysis of Finite Length Blind Adaptive Equalizers", *IEEE Trans. on Signal Processing*, v. 43(9), pp. 2120-2129, 1995.

[5] C. R. Johnson Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown and R. A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review", *Proc. of the IEEE*, v. 86(10), pp. 1927-1950, 1998.

[6] P. A. Regalia, "On the Equivalence Between the Godard and Shalvi-Weinstein Schemes of Blind Equalization", *Signal Processing*, v. 73(1-2), pp. 185-190, 1999.

[7] J. M. T. Romano, R. R. F. Attux, C. C. Cavalcante and R. Suyama, *Unsupervised Signal Processing: Channel Equalization and Source Separation*. CRC Press, 2010.

[8] H. H. Zeng, L. Tong and C. R. Johnson Jr., "An Analysis of Constant Modulus Receivers", *IEEE Trans. on Signal Processing*, v. 47(11), pp. 2990-2999, 1999.

[9] D. G. Fantinato, R. R. F. Attux, C. Sousa Júnior, R. Suyama, A. Neves, J. M. T. Romano, "Two Contributions Derived from a Polynomial Formulation of the Constant Modulus Criterion", *Proc. of XXXI Simpósio Brasileiro de Telecomunicações*, 2013.

[10] H. H. Zeng, L. Tong and C. R. Johnson Jr., "Relationships Between the Constant Modulus and Wiener Receivers", *IEEE Trans. on Information Theory*, v. 44(4), pp. 1523-1538, 1998.

[11] B. G. Agee, "The Least-Squares CMA: A New Technique for Rapid Correction of Constant Modulus Signals", *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 11, pp. 953-956, 1986.

[12] R. Gooch and J. Lundell, "The CM Array: An Adaptive Beamformer for Constant Modulus Signals", *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 11, pp. 2523–2526, 1986.

[13] P. A. Regalia, "A Finite-Interval Constant Modulus Algorithm", *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 3, pp. 2285–2288, 2002.

[14] Y. Chen, T. Le-Ngoc, B. Champagne and C. Xu, "Recursive Least Squares Constant Modulus Algorithm for Blind Adaptive Array", *IEEE Trans. on Signal Processing*, v. 52(5), pp. 1452-1456, 2004.

[15] Vicente Zarzoso and Pierre Comon, "Semi-Blind Constant Modulus Equalization with Optimal Step Size", *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 3, pp. iii/577-iii/580, 2005.

[16] Vicente Zarzoso and Pierre Comon, "Optimal Step-Size Constant Modulus Algorithm", *IEEE Transactions on Communications*, v. 56(1), pp. 10-13, 2008.

[17] K. Wesolowski. "Self-Recovering Adaptive Equalization Algorithms for Digital Radio and Voiceband Data Modems". *European Conf. Circuit Theory & Design*, pp. 19-24, 1987.

[18] A. Benveniste and M. Goursat. "Blind Equalizers". *IEEE Trans. Commun.*, v. COM-32, pp. 871-883, 1984.

[19] J. Yang, J. J. Werner, G. A. Dumont, "The Multimodulus Blind Equalization Algorithm", *Proc. of International Conference on Digital Signal Processing*, v. 1, pp. 127-130, 1997.

[20] M. J. Ready and R. P. Gooch, "Blind Equalization Based on Radius Directed Adaptation", *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 3, pp. 1699-1702, , 1990.

[21] J. Yang, J. J. Werner, G. A. Dumont, "The Multimodulus Blind Equalization and Its Generalized Algorithms", *IEEE Journal on Selected Areas in Communications*, v. 20(5), pp. 997-1015, 2002.

[22] J. Mendes Filho, M. T. M. Silva, M. D. Miranda, V. H. Nascimento, "A Region-Based Algorithm for Blind Equalization of QAM Signals", *IEEE Workshop on Statistical Signal Processing*, pp. 685-688, 2009.

[23] M. Rupi, P. Tsakalides, E. Del Re, C. L. Nikias, "Constant Modulus Blind Equalization Based on Fractional Lower-Order Statistics", *Signal Processing*, v. 84(5), pp. 881-894, 2004.

[24] S.M. Shah, R. Samar, S. M. R. Naqvi, J. A. Chambers, "Fractional Order Constant Modulus Blind Algorithms with Application to Channel Equalisation", *Electronics Letters*, v. 50(23), pp. 1702-1704, 2014.

[25] D. G. Fantinato, R. R. F. Attux, R. Suyama, A. Neves, J. M. T. Romano, "A Volterra Filtering Approach for the Polynomial Formulation of the Constant Modulus Criterion", *Proc. of the International Telecommunications Symposium*, pp. 1-5, 2014.