

Two Contributions Derived from a Polynomial Formulation of the Constant Modulus Criterion

Denis G. Fantinato, Romis Attux, Celso de Sousa Jr., Ricardo Suyama, Aline Neves and J.M.T. Romano

Abstract—From the classical blind CM criterion, we derive a MSE-based polynomial formulation that lead to two contributions: a lower bound of the CM criterion, which works as an equalizability index, and an initialization heuristic for the CMA. The results indicate the validity of the index as an analytical tool and as a practical performance assessment metric in the context of inverse problems. For the heuristic, the results reveal that it is capable of outperforming the classical center-spike method and a random initialization approach.

Keywords—Blind equalization, constant modulus criterion, Equalizability index, CMA initialization.

I. INTRODUCTION

The problem of blind equalization holds a prominent place in modern adaptive signal processing theory. This is in part due to the theoretical relevance of this problem as general formulation of unsupervised inverse tasks in the time domain, but is also related, to a great extent, to the myriad of important applications thereof: echo cancelation, geophysical data mining and digital communications [1].

One of the most used strategies to deal with this problem is that based on the constant modulus (CM) criterion [2], which, along with its associated algorithm, the CMA, was an object of intense study for the last three decades [3][4]. A key result derived from these analyses is the strong connection with the Shavi-Weinstein criterion [3] - an important point of contact, both in terms of providing a unified view of blind approaches and of establishing links with blind source separation (BSS) theory. Other interesting connections were established with the Wiener criterion [1], a supervised formulation based on a mean-squared error (MSE) measure. They reveal, in simple terms, that there is a proximity between good-quality CM solutions and Wiener solutions (considering different delays for the reference or desired signal). However, there remain central theoretical and practical issues associated with the CM formulation that require significant further clarification, which is, in a certain sense, a consequence of the complexity of the CM cost function. In light of this, in this work, we propose a polynomial formulation of the CM cost function in terms of an MSE metric, which will give rise to two main contributions: first, the derivation of a lower bound for the CM cost function, which will work as an attainable

performance measure for channel deconvolution or, in other words, as a blind equalizability index; second, the proposal of a novel initialization heuristic capable of improving the global convergence performance of the CMA in comparison, for instance, with the standard center-spike initialization.

The first contribution is associated to the concept of attainable performance or equalizability, given a channel and an equalizer order. This measure can be obtained in an unsupervised fashion and yields a value that, to a certain extent, follows the attainable residual MSE level for the supervised case. It proves itself to be a potential analytical tool / practical performance assessment index in the context of equalization and linear inverse problems in a broad sense.

The second contribution is related to convergence aspects of CMA: it is well-known that an inadequate initialization of the filter coefficients can lead to local convergence [4][5]. In light of this, based on the solution of the polynomial formulation of the CM cost function, the novel initialization heuristic seeks to avoid as much as possible these situations and provide initial conditions more promising than those obtained using canonical approaches.

This work is organized as follows. In Section II we derive a polynomial formulation of the CM criterion, following constrained and unconstrained approach. Based on the latter strategy, the CM lower bound, which works as a blind equalizability index, is presented in Section III. A number of simulation results are shown in Section IV to illustrate the relevance of the proposal. In Section V, the novel initialization heuristic is introduced and, next, tested with the aid of simulations in Section VI. Finally, the conclusions are summarized in Section VII.

II. POLYNOMIAL FORMULATION OF THE CM CRITERION

The main idea underlying the classical CM criterion resides in the penalization of deviations from a quadratic term with respect to the equalizer output $y(n)$ around a fixed constant value R_2 . It can be expressed in terms of the minimization of the following cost function:

$$J_{CM}(\mathbf{w}) = E \left[(|y(n)|^2 - R_2)^2 \right], \quad (1)$$

where $R_2 = E [|s(n)|^4] / E [|s(n)|^2]^2$, $E[\cdot]$ denotes statistical expectation, \mathbf{w} is the parameter vector of a finite impulse response (FIR) equalizer with N coefficients, $s(n)$ the transmitted signal and $y(n)$ is the equalizer output signal. More specifically, the equalizer output can be defined as $y(n) = \mathbf{w}^H \mathbf{x}(n)$, in which $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]$

Denis G. Fantinato, Romis Attux, Celso de Sousa Jr. and J.M.T. Romano, School of Electrical and Computer Engineering, University of Campinas, Campinas-SP, Brazil, E-mails: {denisgf,attux}@dca.fee.unicamp.br, celso.de_sousa_junior@yahoo.com.br, romano@dmo.fee.unicamp.br. Aline Neves and R. Suyama, Engineering, Modeling and Applied Social Science Center, Federal University of ABC, Santo André - SP, Brazil, E-mails: {aline.neves,ricardo.suyama}@ufabc.edu.br This work was partially supported by CAPES and CNPq.

is the equalizer input vector and $(\cdot)^H$ denotes Hermitian transposition.

From an analytical standpoint, it would be theoretically interesting if one could express the CM cost function as an MSE-like expression, i.e., in accordance with the Wiener criterion [1]:

$$J_{Wiener}(\mathbf{w}) = E \left[|s(n-d) - y(n)|^2 \right], \quad (2)$$

where d is a given delay. Indeed, this will be achieved if a polynomial formulation of (1) is made explicit.

A. The Constrained Case

Supposing that the transmitted signal is composed by unit-norm samples, it is possible to interpret Eq. (1) as a MSE function - in the form of Eq. (2) - between a desired signal $s(n) = R_2 = 1$ and the output $v_c(n) = |y(n)|^2$ of a polynomial Volterra filter [6], here denoted as θ_c , assuming the form:

$$J_{CM}(\mathbf{w}) = E \left[(v_c(n) - 1)^2 \right]. \quad (3)$$

This polynomial filter θ_c presents a quadratic nature by definition, and its parameters are subject to the constraint imposed by the values of \mathbf{w} - the subscript 'c' refers to this constraint. For a better understanding, we will present, as a simple example, the case of a two-tap equalizer (which can be straightforwardly extended to the N -tap case) and real-valued signals and parameters. Taking $v_c(n) = y^2(n)$ and expanding its terms, we have, using vector notation, the following:

$$\begin{aligned} v_c(n) &= \begin{bmatrix} w_0^2 & 2w_0w_1 & w_1^2 \end{bmatrix} \begin{bmatrix} x^2(n) \\ x(n)x(n-1) \\ x^2(n-1) \end{bmatrix} \\ &= \theta_c^T \xi(n), \end{aligned} \quad (4)$$

where θ_c is the polynomial filter parameter vector - which depends on \mathbf{w} - and $\xi(n)$ is the input vector in the Volterra domain.

It is also possible to extend the polynomial formulation of the CM criterion to include complex signals if their intrinsic properties are considered. The cost function defined in (3) will remain the same, but the expansion of the polynomial filter output $v_c(n)$ slightly differs from the case in (4), since now we consider $|y(n)|^2 = y(n)y^*(n)$, where the $*$ denotes complex conjugation, instead of $y^2(n)$ in the real case. For example, if we take a two-tap complex equalizer, $v_c(n)$ will also depend on θ_c and $\xi(n)$, although now they will be complex and defined as:

$$\begin{aligned} \theta_c &= \begin{bmatrix} |w_0|^2 & w_0^*w_1 & w_0w_1^* & |w_1|^2 \end{bmatrix}^T; \\ \xi(n) &= \begin{bmatrix} |x(n)|^2 \\ x(n)x^*(n-1) \\ x^*(n)x(n-1) \\ |x(n-1)|^2 \end{bmatrix}; \quad \theta_c, \xi(n) \in \mathcal{C}. \end{aligned} \quad (5)$$

It is important to note that the complex-valued condition increases the dimension of θ_c and $\xi(n)$, and the constraints for θ_c in respect to \mathbf{w} are more restrictive.

As we have shown, the Eq. (3) defines a polynomial formulation of the CM criterion for both real- and complex-valued signals. It is possible to go further by admitting a

greater freedom of choice of the parameters of the polynomial filter θ_c .

B. The Unconstrained Case

An interesting approach can be obtained if we relax the constraint on the values of \mathbf{w} shown in θ_c and make the polynomial filter parameters completely "free" - they will be hereinafter denoted simply as θ . Thus, if we turn our attention again to the two-tap real- and complex-valued equalizer example, the polynomial filter θ in each case will be defined, respectively, as:

$$\begin{aligned} \theta &= \begin{bmatrix} \theta_0 & \theta_1 & \theta_2 \end{bmatrix} \in \mathcal{R} \text{ or} \\ \theta &= \begin{bmatrix} \theta_0 & \theta_1 & \theta_2 & \theta_3 \end{bmatrix} \in \mathcal{C}. \end{aligned} \quad (6)$$

Independently of the situation, be it real- or complex-valued, the output of the unconstrained polynomial filter, defined as $v(n) = \theta^T \xi(n)$, can also give rise to an MSE-based cost function. In addition to that, if we consider the autocorrelation matrix \mathbf{R}_ξ and the cross-correlation vector \mathbf{p}_ξ for $d(n) = 1$ in the Volterra domain to be defined as

$$\mathbf{R}_\xi = E \left[\xi(n)\xi^H(n) \right]; \quad \mathbf{p}_\xi = E \left[\xi(n) \right], \quad (7)$$

it is possible to express the unconstrained cost function for the polynomial formulation as follows:

$$\begin{aligned} J_{LB}(\theta) &= E \left[(v(n) - 1)^2 \right] \\ &= 1 - \theta^H \mathbf{p}_\xi - \mathbf{p}_\xi^H \theta + \theta^H \mathbf{R}_\xi \theta. \end{aligned} \quad (8)$$

This new cost $J_{LB}(\theta)$, differently from the constrained polynomial cost (3), does not strictly follow the behavior of the CM cost (1). This new expression, nonetheless, is able to provide interesting information regarding the equalization task at hand, as will be seen in the next section.

III. THE CM LOWER BOUND - A BLIND EQUALIZABILITY INDEX

The unconstrained cost $J_{LB}(\theta)$ allows the polynomial filter θ to assume any set of coefficients. By doing this, it is expected that the minimization of $J_{LB}(\theta)$ lead to a MSE value lower than or equal to that attained in the constrained case (3). This means that the unconstrained formulation originates a lower bound to the CM cost function. Interestingly, the unconstrained case directly corresponds to a nonlinear - since $\xi(n)$ is in Volterra domain - Wiener filtering problem, but the dependence with respect to free parameters θ remains linear. This allows us to obtain a closed-form solution that minimizes $J_{LB}(\theta)$ as a consequence of the Wiener-Hopf equations [1]:

$$\theta_o = \mathbf{R}_\xi^{-1} \mathbf{p}_\xi. \quad (9)$$

It is important to notice that the solution for θ_o depends on second-order statistics in \mathbf{p}_ξ and fourth-order statistics in \mathbf{R}_ξ with respect to the received signal $x(n)$. The minimum MSE value can be straightforwardly obtained from Wiener filtering theory [1]:

$$J_{LB}(\theta_o) = 1 - \mathbf{p}_\xi^H \mathbf{R}_\xi^{-1} \mathbf{p}_\xi, \quad (10)$$

which constitutes the CM cost lower bound, since

$$J_{CM}(\mathbf{w}) \geq J_{LB}(\boldsymbol{\theta}_o). \quad (11)$$

At this point, a few remarks should be made concerning this CM lower bound $J_{LB}(\boldsymbol{\theta}_o)$. Firstly, as shown in Eq. (10), this value can be obtained in a totally unsupervised fashion, since it depends only on the statistics that form \mathbf{R}_ξ and \mathbf{p}_ξ . Secondly, as the simulations will show, the value $J_{LB}(\boldsymbol{\theta}_o)$ is a measure that can be employed as an unsupervised achievable performance evaluation metric. Therefore, ideally, by using (10), it is possible to extract from the statistical structure of the input signal a value that indicates how “equalizable” is a given channel for a certain equalizer order.

Furthermore, as there tends to exist a certain proximity between the good CM solutions and the best (supervised) Wiener solutions (2), it is possible to ask whether there is a close relationship between the CM lower bound value $J_{LB}(\boldsymbol{\theta}_o)$ and the attainable MSE performance of the supervised case, considering, which is quite interesting, the most suitable delay d . Hence, aside from being a theoretical indicative of the attainable performance under the CM criterion, the lower bound, as the simulations indicate, also is related to the attainable MSE level - this evokes again the notion of a blind equalizability index.

This discussion, which is speculative in some points, will become clearer with the aid of empirical analyses - this is the next step of this work.

IV. SIMULATION RESULTS FOR THE CM LOWER BOUND

In order to verify the soundness of the CM lower bound $J_{LB}(\boldsymbol{\theta}_o)$ in itself and as a blind equalizability index, it is necessary to compare its performance to those of the global optimal CM and (supervised) Wiener solutions (taking into account the equalization delay). For simulation effects, we estimate the minimum value of the CM cost function by initializing the CMA ($\mu = 0.0008$) at the best Wiener solution (in terms of the equalization delay), which can be considered a sufficiently reliable strategy to obtain global convergence.

Moreover, to test the hypotheses raised in the previous sections, the connections between the CM lower bound and the minimum CM and Wiener costs must hold for a variety of channel models. We consider, therefore, for real-valued systems, three scenarios: a first-, a second- and a third-order channel. For the complex-case, we analyze a first-order channel.

Aiming to cover a wide range of channels, we first consider a scenario with real-valued parameters in which the channel transfer function is given by $H(z) = 1 + \alpha z^{-1}$ (normalized to have unit norm), where α varies from 0 to 3, and the equalizer is a two-tap filter. The source is composed of 50000 independent and identically distributed (i.i.d.) BPSK samples and there is no additive noise. The obtained values for $J_{LB}(\boldsymbol{\theta}_o)$ and the minimum value of the CM and Wiener costs are illustrated in Fig. 1. The figure reveals that $J_{LB}(\boldsymbol{\theta}_o)$ is, as expected, a lower bound for the CM cost function. Also, it should be noticed that the CM lower bound tends to follow the general shape of both the CM and Wiener costs, i.e., it assumes lower

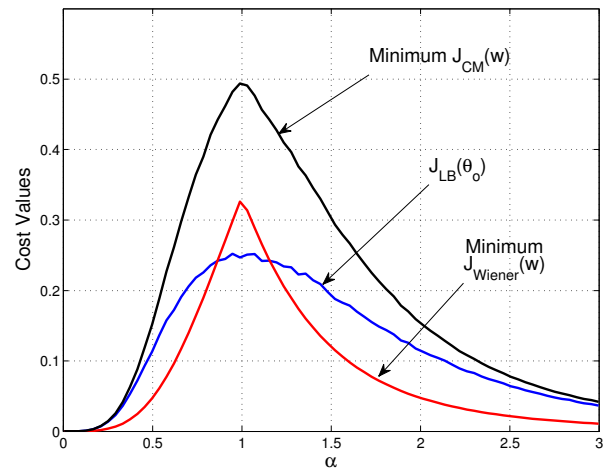


Fig. 1. Minimum cost values and CM lower bound for different 2-tap channel.

values for channels that are relatively “simpler” to equalize (or, in other words, channels that can lead to lower values of optimal MSE cost). On the other hand, it consistently achieves higher values for channels that are “harder” to equalize, which gives support to the aforementioned practical notion of its use as a blind equalizability index.

The analysis is now extended to a case in which the channel is a second-order system with transfer function $H(z) = 1 + \alpha z^{-1} + \beta z^{-2}$ (with subsequent unit-power normalization), where α and β vary from 0 to 2, and a three-tap equalizer is employed. The results are shown in Fig. 2. As in the previous

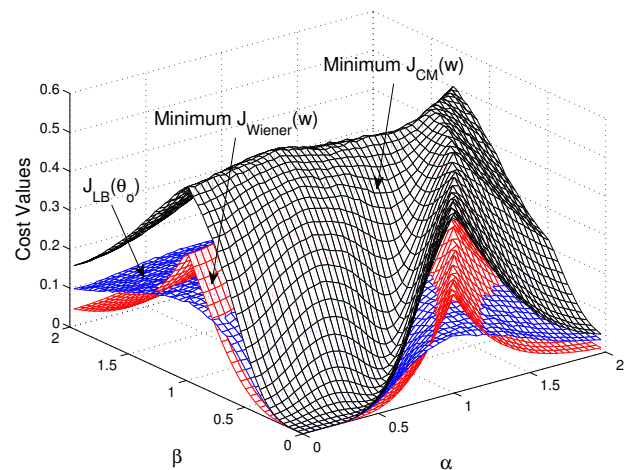


Fig. 2. Minimum cost values and CM lower bound for different 3-tap channel.

case, the CM lower bound held its validity and was shown to be consistent with the general pattern of the minimum MSE.

As a last test for real-valued scenarios, 500 third-order channels with coefficients randomly generated according to a uniform distribution from -1 to $+1$ and an equalizer with same order are considered. For each channel, the values for $J_{LB}(\boldsymbol{\theta}_o)$ and of the minimum CM and Wiener costs were obtained. In the histogram presented in Fig. 3, we show the frequency of the difference values between (i) the CM cost

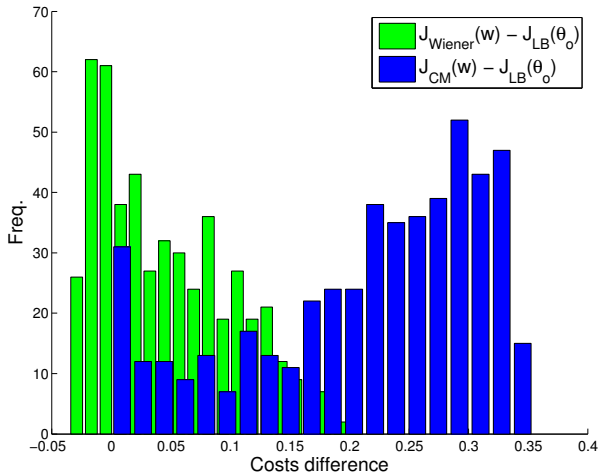


Fig. 3. Differences histogram for randomly generated 4-tap channel.

and $J_{LB}(\theta_o)$ and (ii) the Wiener cost and $J_{LB}(\theta_o)$. It is clear, from these differences, that the CM lower bound is again related in a consistent way to the minimum CM and Wiener costs. Notice, in particular, the proximity with respect to the supervised formulation, which is indicative of the potential relevance of the lower bound and an indicative of the potential of channel equalization.

For complex-valued systems and parameters, we concentrate our analysis on a complex channel of the type $H(z) = 1 + (\alpha + j\beta)z^{-1}$, with α and β varying from 0 to 2, an i.i.d. 4-QAM modulated source of 50000 samples and a two-tap complex equalizer. Fig. 4 shows the obtained cost values and the CM lower bound for this scenario. It is possible to say

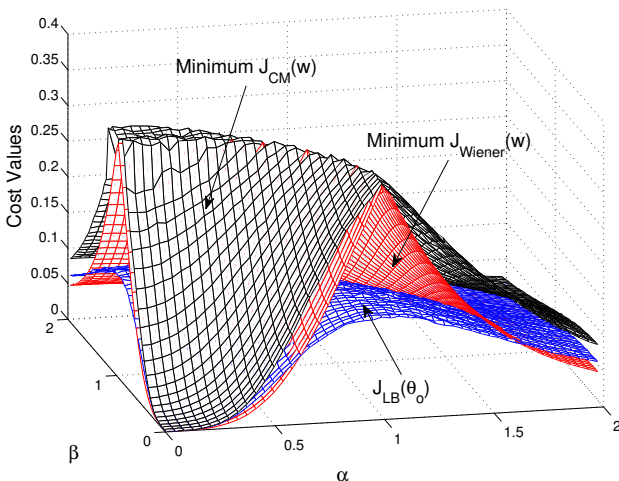


Fig. 4. Minimum cost values and CM lower bound for complex 2-tap channel.

that, since the complex-valued case admits a larger number of coefficients for θ , the solution for $J_{LB}(\theta)$ presents a higher degree of freedom in relation to the real-valued case. Hence, as we can see in Fig. 4, the lower bound $J_{LB}(\theta_o)$ attains lower values. Even so, the idea of the blind equalizability index is still feasible.

The simulations in both real- and complex-valued scenarios indicate that the blind equalizability index is consistent, which can be interesting for practical applications in digital communication and blind source separation (e.g. in underdetermined cases dealt with via kurtosis-based methods).

V. INITIALIZATION HEURISTIC

From a blind equalization standpoint, it is well-known that a suitable initialization for the CMA can potentially lead to the global minima of the CMA, being that, however, a difficult task in practice. As seen in Section IV, since the CM lower bound establishes connections with the performance of optimal CM filters, it is not unreasonable to expect that there be some sort of relationship between the unconstrained polynomial filter solution θ_o and the CM global minima. Therefore, if it is possible to have access to these shared characteristics, e.g., through a mapping between θ_o and \mathbf{w} , it should be possible to increase the CM global convergence rate.

Considering only the real-valued case, by comparing the constrained and unconstrained polynomial filter coefficients, θ and θ_c , respectively, it is possible to see that, as in the two-tap filter example given by Eqs. (6) and (4), there will be a correspondence between (i) θ_0 and w_0^2 , (ii) θ_1 and $2w_0w_1$, and (iii) θ_2 and w_1^2 . There are, hence, some terms of θ associated to quadratic terms of \mathbf{w} , as cases (i) and (iii), which we will denote as $\theta[w_l^2]$ for $l = 0, \dots, N$, and other θ terms related to cross-product of the filter coefficients \mathbf{w} , like in (ii), which will be denoted as $\theta[w_lw_m]$ for $l, m = 0, \dots, N$. If these considered terms from θ_c and θ remain close, they can be employed to provide an estimate of the CM optimal solution, which can be used as an initial condition to the CMA.

In order to better explore the relationship between θ_c and θ , we propose an initialization heuristic to obtain the initial coefficients \mathbf{w} for CMA, which can be expressed by the following algorithm:

Algorithm 1 Initialization Heuristic

```

 $w_l \leftarrow \sqrt{\max_l |\theta[w_l^2]|}$ 
for  $m$  from 0 to  $N$ ,  $m \neq l$  do
     $w_m \leftarrow \theta[w_lw_m] / (2w_l)$ 
end for
    
```

Basically, the algorithm searches for the largest magnitude of $\theta[w_l^2]$ to determine the associated filter coefficient w_l . After that, the remaining coefficients of \mathbf{w} are obtained from the related cross-product $\theta[w_lw_m]$. It is important to remark that the sign of the largest magnitude reference w_l is always assumed to be positive, which implies no loss of generality in view of the sign ambiguity inherent to Bussgang algorithms [1].

The complex-valued case presents a more intricate relationship between θ_c and θ and will not be considered in this work, being a perspective for future research.

VI. SIMULATION RESULTS FOR THE INITIALIZATION HEURISTIC

In order to test the initialization heuristic, we consider three different scenarios: channels of first-, second- and fifth-orders.

The equalizer is always assumed to have the same order of the channel. In all cases, the proposed heuristic is compared to two other methods: the standard center-spike initialization [5] and a random initialization based on a uniform coefficient distribution between -1 and $+1$. The performance of all methods is analyzed using the outcome of the CMA, with $\mu = 0.0008$, for each initialization, after 30000 iterations. Global convergence is assumed to have occurred when the distance between the taps obtained via CMA and the optimal taps - assumed to be the outcome of the CMA initialized at the best Wiener solution - is smaller than 0.1 (taking into account the sign ambiguity).

In the first case, the channel has the transfer function $H(z) = 1 + \alpha z^{-1}$ (subsequently normalized), where α varies from 0 to 2 in steps of 0.041. Tab. I displays the global convergence rates for all three methods and also their associated average CM cost. From the results, the heuristic

TABLE I
GLOBAL CONVERGENCE RATE AND AVERAGE CM COST FOR FIRST-ORDER CHANNELS.

Init.	Global Conv. Rate	Average CM Cost
Heuristic	100%	0.2540
Center-Spike	46.94%	0.3238
Random	42.86%	0.3237

allowed the best possible performance to be reached in all cases, whereas the two other methods had a rate lower than 50%, which clearly indicates the efficiency of the proposal. An analysis of the obtained average CM cost is also significantly favorable to the new heuristic.

In the second scenario, the channel is of the form $H(z) = 1 + \alpha z^{-1} + \beta z^{-2}$ (with subsequent normalization), where α and β vary from 0 to 2, in steps of 0.041. The results are presented in Tab. II, and they also give support to the valid-

TABLE II
GLOBAL CONVERGENCE RATE AND AVERAGE CM COST FOR SECOND-ORDER CHANNELS.

Init.	Global Conv. Rate	Average CM Cost
Heuristic	78.22%	0.3243
Center-Spike	35.44%	0.3673
Random	37.61%	0.3809

ity of the proposed initialization procedure, as the heuristic presented a global convergence rate that is more than twice the rate obtained by the other approaches, as well as a better performance level in terms of average CM cost.

In the last test case, due to the dimension of the problem at hand, 100 normalized fifth-order channels had their coefficients randomly generated according to a uniform coefficient distribution between -1 and $+1$. The global convergence rates and their average CM costs, presented in Tab. III, once again indicate the practical relevance of the initialization heuristic.

TABLE III
GLOBAL CONVERGENCE RATE AND AVERAGE CM COST FOR FIFTH-ORDER CHANNELS.

Init.	Global Conv. Rate	Average CM Cost
Heuristic	40%	0.3950
Center-Spike	16%	0.4322
Random	13%	0.4440

VII. CONCLUSIONS

In this work, we discuss a polynomial formulation of the CM criterion, which by means of a constraint relaxation procedure, leads to a lower bound for the associated cost function. It is shown that this bound has the potential of serving as a sort of equalizability index i.e. a measure that indicates how well a certain FIR equalizer will perform, ideally, for a given received signal. Although this was not discussed here for the sake of space limitation, the same idea can be extended to handle the initialization process in underdetermined source separation based on fourth-order methods (e.g., methods based on kurtosis). The results showed that the index is useful from an analytical standpoint and is also potentially interesting in practice (e.g., in digital communications, space-time audio processing etc.).

Based on the polynomial formulation, we proposed an initialization heuristic for CMA, which showed a good performance in all tested scenarios, having reached better convergence rates and average CM cost than those obtained via the canonical center-spike methodology and a random selector.

As perspectives for future work, we indicate: (i) a detailed analysis of the lower bound and of the associated heuristic in the context of blind source separation, (ii) a more extensive treatment of the complex case, (iii) a search for modifications in the heuristic that lead to further improvements and (iv) studies for noisy channels.

ACKNOWLEDGEMENTS

The authors thank CAPES and CNPq for the financial support.

REFERENCES

- [1] J.M.T. Romano, R. Attux, C.C. Cavalcante, R. Suyama, "Unsupervised Signal Processing: Channel Equalization and Source Separation," *CRC Press*, 2010.
- [2] D.N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. on Comm.*, v. 28(11), pp. 1867-1875, 1980.
- [3] 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.
- [4] Z. Ding, R.A. Kennedy, B.D.O. Anderson, C.R. Johnson, Jr., "Ill-Convergence of Godard Blind Equalizers in Data Communication Systems," *IEEE Trans. on Communications*, v. 39(9), pp. 1313-1327, 1991.
- [5] G.J. Foschini, "Equalizing without altering or detecting data," *AT&T Technical Journal*, v. 64(8), pp. 1885-1911, 1985.
- [6] V.J. Mathews, G.L. Sicuranza, "Polynomial Signal Processing," *Wiley*, New York, 2000.