Vector-CM Stable Equilibrium Analysis

Azzédine Touzni, Lang Tong, Raúl A. Casas, and C. Richard Johnson, Jr.

***Abstract*—The vector-constant modulus (VCM) criterion is an extension of the constant modulus (CM) criterion [2] introduced recently for equalization of channels involving Gaussian sources [1], [5]. In this letter, we analyze the behavior of VCM for arbitrary source distributions and combined channel-receiver impulse responses of finite dimension. We begin by pointing out the difference between the VCM and CM cost functions and showing that the VCM criterion can be expressed as a composite criterion combining the CM cost function and a penalty term involving cross-correlations of the equalizer output. We continue by providing conditions for noise-free channels, under which VCM admits stable minima corresponding to zero-forcing (ZF) channel receivers. We find that for sub-Gaussian sources, the VCM and CM criteria share the same global minima. For Gaussian and super-Gaussian sources, however, it appears that only ZF receivers corresponding to input/output transmission delays at the extremes of the range of possible delays are truly stable equilibria of VCM.**

***Index Terms*—Blind equalization, source shaping, vector-CM criterion.**

I. Introduction

Blind linear estimation of Gaussian sources finds application in a wide range of signal processing problems such as channel equalization, source separation, and sensor-array processing. For example, in digital communications, when source shaping [3] is used for transmission of high-order constellations to achieve optimum transmitted signal power, this signal has an approximately Gaussian distribution. In this case, well-known criteria like the constant modulus (CM) criterion fail to equalize the channel [2].

The vector constant modulus (VCM) cost function was recently introduced in [1] as an extension of the well-known CM cost function. [1], [5] give intuitive considerations emphasizing the ability of VCM to equalize data with Gaussian source distributions.

The contribution of this letter is to provide an analytical description of the robustness of the VCM cost function. More precisely, we derive the conditions under which VCM admits stable minima corresponding to zero-forcing (ZF) combined channel-receiver responses, in a noise-free scenario, for arbitrary source distributions and a channel-receiver response of finite dimension. It appears that the robustness of VCM with respect to Gaussian source distributions stems mainly from the fact that VCM can be seen as a composite criterion involving the CM cost function and a cross-correlation constraint in the time domain.

Manuscript received August 16, 1999. This work was supported in part Institut National de Recherche en Informatique et Automatique (INRIA), Rocquencourt, France, NSF Grant CCR-9804019 and NSF Grant ECS-9528363. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stan Reeves.

The authors are with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: touzni@ee.cornell.edu; ltong@ee.cornell.edu; raulc@ee.cornell.edu; johnson@ee.cornell.edu).

Publisher Item Identifier S 1070-9908(00)00922-6.

We show that VCM admits stable ZF receivers associated with input/output transmission delays at the extremes of the range of possible delays for Gaussian and super-Gaussian sources. Surprisingly, the condition of stability of these extrema, based on the kurtosis of the source, depends on the length of the channel-receiver response. In fact, the kurtosis of the source must be smaller than a constant, which decreases when the length of the channel increases. This constant is equal to 3 (when all the variables are real valued) in the limit case, where the channel length is infinite. For sub-Gaussian sources, it is shown that all ZF receivers, which are stable minima of the CM cost function, are also stable minima of the VCM cost function.

II. Problem Formulation

A. Criterion Definition

The relation between the received signal $\underline{x}(n)$ (the observations) and the source signal, in a noise-free scenario, is given by

$$\underline{x}(n) = H\underline{s}(n) \tag{1}$$

$$y(n) = \underline{f}^t\underline{x}(n). \tag{2}$$

where H denotes the channel-convolution matrix and $\underline{w}(n)$ corresponds to additive noise. We assume that the combined channel-receiver response $\underline{q} = H^t\underline{f}$ is an N-dimensional vector.

Linear estimates of $s_i(n)$, the ith component of the unknown source signal $\underline{s}(n)$, are provided by the minima of the VCM cost-function defined by

$$J_v^{(P)}(\underline{f}) \stackrel{\text{def}}{=} E\left\{\left(\sum_{k=0}^{P-1} y^2(n-k) - \rho\right)^2\right\} \tag{3}$$

where $\rho = E\{s^4\}/E\{s^2\}$ is the so-called dispersion constant introduced in the CM criterion. We assume, moreover, that the source is an independent and identically distributed (i.i.d.) zero-mean process of variance $E\{s^2\}$ and kurtosis $\kappa = E\{s^4\}/E\{s^2\}^2$, and that H is a full-column rank. All the variables are assumed to be real valued.

VCM will be compared to the well-known CM criterion, defined by

$$J_c(\underline{f}) = E\{(y^2(n) - \rho)^2\}. \tag{4}$$

Notice that the VCM cost function is equal to the CM cost function when $P = 1$. The choice of this parameter for the VCM cost function will be addressed in the sequel.

B. Example

From the definition (3), it is not clear that the VCM cost function admits stable minima leading to the estimation of a Gaussian source. In order to illustrate the difference between the CM and VCM cost functions, we consider an example of a

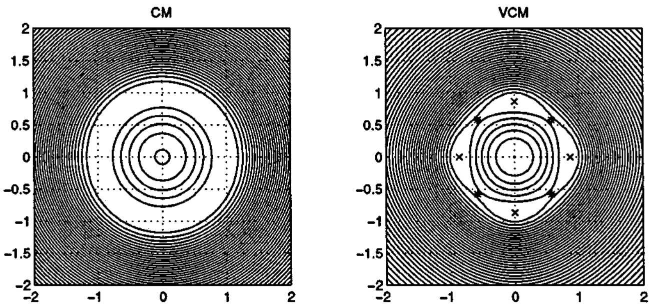

Fig. 1. VCM versus CM cost function for $N = 2$, $P = 2$, and $\kappa = 3$.

two-tap, channel-receiver impulse response $\underline{q} = (q_1, q_2)$ with $P = 2$ and a Gaussian source ($\kappa = 3$). We will characterize the extrema of each cost function.

When the input signal is Gaussian, $J_c(\underline{f})$ is equivalent to an output power constraint. The CM cost function admits infinitely many minima, which are solutions of the equation $q_1^2 + q_2^2 = 1$. In spite of the Gaussianity of the input signal, the VCM cost function admits distinct extrema, which are split into four global minima $\underline{q} = (0, \pm\sqrt{3/4})$ and $\underline{q} = (\pm\sqrt{3/4}, 0)$ leading to the estimation of source $s(n)$ or $s(n-1)$ and four unstable extrema corresponding to saddle points given by $\underline{q} = (\pm\sqrt{2/9}, \pm\sqrt{2/9})$, as is shown in Fig. 1.

III. Interpretation

According to the relation (5) below, the VCM criterion (3) can be expressed as a composite criterion involving the CM cost function and a penalty measure based on the cross-correlation of the square of the outputs.

$$J_v^{(P)}(\underline{f}) = PJ_c(\underline{f}) + 2\sum_{k=1}^{P-1}(P-k)E\left\{y^2(n)y^2(n-k)\right\} + (1-P)\rho^2. \quad (5)$$

Composite criteria involving the CM cost function and decorrelation-like regularization constraints on the outputs are usually introduced in order to avoid possible local minima of the CM cost function, when all global impulse-response parameterizations $\underline{q}$ are not achievable (i.e. when the channel-impulse matrix H is not full-column rank) under the assumption that the source is sub-Gaussian ($\kappa < 3$). The idea, based on a noise-free scenario, is to penalize the minima of the CM cost function, which differs from ZF receivers [7], [4].

The question addressed herein, concerning VCM, is rather different. We investigate the effect of the cross-correlation constraint, giving the composite VCM cost function, in order to derive the conditions on stability of ZF receivers when the source has an arbitrary distribution. Note that this question has not been investigated previously in the literature (to our knowledge) for other similar composite criteria.

Before addressing this problem, let us consider the CRIMNO cost function defined in [4] by

$$J_r(\underline{f}) = \lambda_0 J_c(\underline{f}) + \sum_{k=1}^{P-1}\lambda_k E\{y(n)y(n-k)\}^2 \quad (6)$$

for which the coefficients $\lambda_0, \ldots, \lambda_{P-1}$ are free parameters not explicitly defined. Clearly, the CRIMNO and VCM criteria are comparable, the difference between them being mainly in the squaring introduced in the penalty measure.

TABLE I
Classification of the Extrema of VCM for $N = 3$.

q_1	q_2	q_3
0	0	0
$\pm\sqrt{3/5}$	0	0
0	0	$\pm\sqrt{3/5}$
0	$\pm\sqrt{3/5}$	0
$\mp 3\sqrt{94}/47$	$\pm 3\sqrt{94}/94$	$\pm 3\sqrt{94}/94$
$\pm 3\sqrt{94}/94$	$\pm 3\sqrt{94}/94$	$\mp 3\sqrt{94}/47$
$-3\sqrt{94}/94$	$3\sqrt{94}/47$	$3\sqrt{94}/94$
$3\sqrt{94}/94$	$-3\sqrt{94}/94$	$-3\sqrt{94}/47$

When the source is Gaussian, the CM cost function is equivalent to an output power constraint. Therefore, $J_r(\underline{f})$ is basically a decorrelation criterion with a norm constraint on $\underline{q}$. It has been emphasized by several authors that output decorrelation is a sufficient criterion for equalization when the channel-matrix response is full-column rank [9]. Therefore, we realize that conditions on (λ_k) can be derived to guarantee the stability of all possible ZF receivers in the noise-free scenario for sub-Gaussian and Gaussian sources when H is full-column rank.

In the next section, we show that behavior of the VCM cost function is, surprisingly, significantly different from the behavior of the CRIMNO criterion. This result stems from the fact that the decorrelation argument cannot be used for the VCM criterion (because $E\{y^2(n)y^2(n-k)\}$ differ from a decorrelation measure) to explain its ability to admit stable ZF solutions when the source is Gaussian.

IV. Stability of ZF Solutions

Next, we prove that VCM admits ZF extrema, and we derive the conditions of stability of these solutions.

According to the assumptions introduced in the introduction, characterization of the VCM extrema can be equivalently given in terms of the channel-receiver response $\underline{q}$. Let us consider the kth component of the gradient $\nabla J_v^{(P)}(\underline{q})$ given by

$$\frac{\partial J^{(P)}(\underline{q})}{\partial q_k} = 4PE\{s^2\}^2\left((\kappa-3)q_k^2 + (P+2)\|\underline{q}\|_2^2 - \kappa\right)q_k + 4(\kappa-3)E\{s^2\}^2\left(\sum_{j=1}^{P-1}(P-j)\omega_{k,j}^{(2)}\right)q_k + 8E\{s^2\}^2\left(\sum_{j=1}^{P-1}(P-j)\left(\sum_i q_i q_{i-j}\right)\omega_{k,j}^{(1)}\right) \quad (7)$$

where q_k is the kth component of $\underline{q}$. The terms $\omega_{k,j}^{(p)}$ (with $p = 1, 2$) are defined by q_{k+j}^p when $k-j < 0$, $q_{k-j}^p + q_{k+j}^p$ when $k+j \le P-1$ and $k-j \ge 0$, q_{k-j}^p when $k+j > P-1$, and 0 elsewhere. Since the contributions of $\omega_{k,j}^{(2)} q_k$ and $(\sum_i q_i q_{i-j})\omega_{k,j}^{(1)}$ of

(7) are equal to zero when $\underline{q}$ has only one nonzero component, the ZF solutions that are the extrema of (5) also correspond to the ZF solutions zeroing the gradient of the CM criterion.

The result concerning the stability of these solutions is summarized in the theorem that follows.

Theorem 1[1]*:* The cost function $J_v^{(P)}(\underline{q})$ admits extrema of the form $q_v^{(m)} = (0\ldots,0,\mu,0,\ldots,0)$, where μ is the mth component of $q_v^{(m)}$ for

$$\mu = 0, \quad \text{and} \quad \mu = \pm\sqrt{\frac{\kappa}{\kappa+P-1}}. \tag{8}$$

The stability of the previous solutions is related to the sign of the quadratic form $u^t\Psi(\underline{q})u$, where $\Psi(\underline{q})$ is the Hessian of $J_v^{(P)}(\underline{q})$ of dimension $N\times N$ of entries $\Psi_{k,j}(\underline{q}) = (\partial^2 J^{(P)}(\underline{q})/\partial q_k \partial q_j)$ and u is an arbitrary vector.

Under the assumption $P = N$ (i.e., when the averaging window in (3) matches the combined channel-receiver length), we have the following.

1) For $\underline{q}_v = (0\cdots 0)$, $\Psi(0) = -4PE\{s^4\}I_{N\times N} < 0$. Therefore, $\underline{q}_v$ is a maximum.
2) Let
$$\kappa_\star = 1 + 2\frac{P}{P-1}, \quad \text{where } P > 1. \tag{9}$$
When $\kappa > \kappa_\star$, $\Psi(\underline{q})$ admits positive and negative eigenvalues for $\underline{q} = \underline{q}_v^{(j)}$ with $j = 1,\ldots,P$, these solutions correspond to saddle points. When $\kappa = \kappa_\star$, $\Psi(\underline{q}) \geq 0$ for $\underline{q} = \underline{q}_v^{(1)}$ and $\underline{q} = \underline{q}_v^{(P)}$. All other solutions correspond again to saddle points.
3) Under the condition $3 \leq \kappa < \kappa_\star$, the Hessian $\Psi(\underline{q}) > 0$ for $\underline{q} = \underline{q}_v^{(1)}$ and $\underline{q} = \underline{q}_v^{(P)}$. These two vectors are thus stable minima of $J_v^{(P)}(\underline{q})$. When $\kappa > 3$, $\Psi(\underline{q})$ admits positive and negative eigenvalues for $\underline{q} = \underline{q}_v^{(j)}$ with $j \neq 1, P$. These solutions are thus saddle points. When $\kappa = 3$,, $\Psi(\underline{q}) \geq 0$ for $\underline{q} = \underline{q}_v^{(j)}$ with $j \neq 1, P$.
4) Under the condition $\kappa < 3$, $\Psi(q) > 0$ for $\underline{q} = \underline{q}_v^{(j)}$ with $j = 1,\ldots,P$. All these solutions correspond to global minima.

The bound (9) is the admissible maximum source kurtosis for which the VCM cost function admits global stable minima, leading to the perfect estimation of the input signal, indicating that the source could be Gaussian as well as super-Gaussian. For example, with $P = 2$, we have $\kappa_\star = 5$. Interesting behavior occurs when $3 \leq \kappa < \kappa_\star$, where the only stable ZF solutions are those leading to the estimation of the source $s(n)$ and $s(n-(P-1))$. When $\kappa = 3$, the stability of the zero-forcing receivers corresponding to other transmission delays is not well defined due to the presence of zero eigenvalues. This specific behavior, emphasizing the role of the two delays at the edges of the combined channel-equalizer response, stems from the property of symmetry of the constraint term appearing in the composite criterion (5). The constraint can be seen as cross-correlation penalization of $y(n)$ with respect to the past observations $y(n-k)$ for $k > 1$, or equivalently, a cross-correlation penalization $y(n-(P-1))$ with respect to future observations $y(n-k)$ for $k < 1$. Comparable behavior occurs, for instance, for linear-prediction methods that lead to the estimation of the source of delay zero (i.e. $s(n)$ [6], [8]).

Notice that the upper bound $\kappa_\star$ depends implicitly on the channel length. When $P \to +\infty$, we have $\kappa_\star \to 3$. This shows, in particular, that when the channel and receiver have a double infinite response, VCM does not admit stable ZF minima for Gaussian and super-Gaussian sources.

[1]The proof is available on request.

V. Other Extrema

In this section, we address the problem of the characterization of extrema that differ from ZF solutions. We focus on the special case of Gaussian sources (i.e. $\kappa = 3$).

The location of all VCM extrema and the analysis of their stability for arbitrary P and N (where N denotes the dimension of the channel-receiver response) is not a trivial problem. We give insight on the behavior of VCM through characterization of all the extrema for $N = 3$. Two different situations $P = N$ and $P < N$, have to be considered separately for the analysis of VCM.

For $P = N$, all the extrema of VCM are shown in the Table I. The stable solutions corresponding to ZF estimation of the source that are associated with delays at the edges of the channel-equalizer response are given in boldface. The Hessian for the ZF center-spike solution has one zero eigenvalue so that stability of this solution is not well defined, as is mentioned in Theorem 1. The other extrema are split into a global maximum $\underline{q} = 0$ and saddle points, and are thus unstable.

For $P < N$, it is easy to characterize for $N \geq 3$ subsets of extrema that do not correspond to ZF receivers. For example, for $N = 3$ and $P = 2$, $\underline{q}_v = (q_1, 0, q_3)$ with $q_1^2 + q_3^2 = 1/16$ denotes a continuum of solutions as extrema of the VCM cost function. However, the stability of these solutions for VCM are not clearly defined, in contrast with the CM cost function for which these solutions (up to a scale factor) correspond to a set of stable minima. In fact, the eigenvalues of the Hessian for these solutions are given by $0, 2+8q_3\sqrt{1-q_3^2}, 16$. Thus, the Hessian could be semi-positive definite or semi-negative definite. We get arrive at similar conclusions for the general case $N > 3$.

References

[1] V. Y. Yang and D. L. Jones, "A vector constant modulus algorithm for shaped constellation equalization," *IEEE Signal Processing Lett.*, vol. 5, Apr. 1998.

[2] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.

[3] M. V. Eyuboglu, G. D. Forney Jr., P. Dong, and G. Long, "Advanced modulation techniques for V. Fast," *Eur. Trans. Telecom.*, vol. 4, pp. 243–256, May/June 1993.

[4] Y. Chen, C. L. Nikias, and J. G. Proakis, "Blind equalization with criterion with memory nonlinearity," *Opt. Eng.*, vol. 31, June 1992.

[5] M. A. Haun and D. L. Jones, "The fractionally spaced vector constant modulus algorithm," in *ICASSP'99*, Phoenix, AZ, Mar. 15–19.

[6] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[7] S. Choi, H. Luo, and R. W. Liu, "An adaptive system for direct blind multi-channel equalization," in *SPAWC'97*, Paris, France, Apr. 1997.

[8] O. Macchi and A. Hachicha, "Self-adaptive equalization based on prediction principle," *GLOBECOM'86*, pp. 1641–1645.

[9] D. Gesbert, C. B. Papadias, and A. Paulraj, "Blind equalization of polyphase FIR channels: A whitening approach," *Asilomar'97*, pp. 1604–1608, Nov.