On the misbehavior of constant modulus equalizers for improper modulations

Angela Sara Cacciapuoti and Francesco Verde

***Abstract*— In this paper, the constant modulus (CM) cost function is analyzed under the general assumptions that improper modulation schemes of practical interest are employed and the baseband equivalent of the channel impulse response is complex-valued. This study allows one to determine a broad family of undesired minima of the CM cost function, which do not lead to perfect symbol recovery in the absence of noise. The results developed herein generalize and subsume as a particular case existing studies of the CM cost function, which exclusively consider real-valued binary modulations.**

I. Introduction and preliminaries

LET[1] us consider a passband digital communication system employing linear modulation with symbol period T_s, and transmitting over a linear time-invariant (LTI) finite-impulse response (FIR) channel whose baseband impulse response $c_a(t)$ spans $L_c > 1$ symbol periods. After filtering and ideal carrier-frequency recovery, the complex envelope $r_a(t)$ of the received continuous-time signal is fractionally sampled at rate N/T_s, with $N > 1$ denoting the oversampling factor, and N consecutive samples, which are taken within the kth ($k \in \mathbb{Z}$) symbol interval, are collected in the vector $\mathbf{r}(k) \triangleq [r^{(0)}(k), r^{(1)}(k), \ldots, r^{(N-1)}(k)]^T \in \mathbb{C}^N$, with $r^{(\ell)}(k) \triangleq r_a(k\,T_s + \ell\,T_s/N)$ for $\ell \in \{0, 1, \ldots, N-1\}$.

Blind channel equalization consists of designing a linear FIR (L-FIR) equalizer, which is able to extract the desired symbol $s(k-d)$ (with d denoting the equalization delay) by jointly counteracting intersymbol interference (ISI) and noise, without making use of training sequences. Denoting with L_e the equalizer length (expressed in symbol intervals), the input-output relationship of an L-FIR equalizer is $y(k) = \mathbf{f}^H \mathbf{z}(k)$, where $\mathbf{f} \in \mathbb{C}^{NL_e}$ collects all the equalizer's parameters and $\mathbf{z}(k) \triangleq [\,\mathbf{r}^T(k), \mathbf{r}^T(k-1), \ldots, \mathbf{r}^T(k-L_e+1)\,]^T \in \mathbb{C}^{NL_e}$ can be expressed (see, e.g., [1]) as

$$\mathbf{z}(k) = \mathbf{C}\,\mathbf{s}(k) + \mathbf{v}(k)\,, \tag{1}$$

where $\mathbf{s}(k) \triangleq [s(k),\, s(k-1),\, \ldots,\, s(k-K+1)]^T \in \mathbb{C}^K$, the matrix $\mathbf{C} \in \mathbb{C}^{(NL_e)\times K}$ is Sylvester depending on the vectors $\mathbf{c}(q) \triangleq [\,c^{(0)}(q),\, c^{(1)}(q),\, \ldots,\, c^{(N-1)}(q)\,]^T \in \mathbb{C}^N$, with $K \triangleq L_e + L_c - 1$ and $c^{(\ell)}(q) \triangleq c_a(q\,T_s + \ell\,T_s/N)$ and, finally, the vector $\mathbf{v}(k) \in \mathbb{C}^{NL_e}$ accounts for noise. The following customary assumptions will be considered hereinafter:

A1) $\mathbf{C}$ is full-column rank, i.e., $NL_e \geq K$ and $\text{rank}(\mathbf{C}) = K$;

A2) $s(n) \in \mathbb{C}$ is a sequence of independent and identically distributed (i.i.d.) zero-mean random variables, whose kurtosis $\kappa_s \triangleq \text{E}[|s(n)|^4] - 2\,\text{E}^2[|s(n)|^2] - \left|\text{E}[s^2(n)]\right|^2 < 0$.

In the absence of noise, assumption A1 guarantees [1] the existence of L-FIR zero-forcing (ZF) designs for $\mathbf{f}$, i.e., solutions so that $\mathbf{C}^H\mathbf{f} = \mathbf{e}_d$, with $d \in \{0, 1, \ldots, K-1\}$. As regards assumption A2, the condition $\kappa_s < 0$ imposes that the transmitted symbols are "sub-Gaussian" [1], which is the case commonly encountered in digital communications.

To blindly suppress ISI and noise, one can resort to the constant modulus (CM) criterion [1], where the vector $\mathbf{f}$ is chosen such as to minimize the cost function

$$J_{\text{cm}}(\mathbf{f}) \triangleq \text{E}[(\gamma_s - |y(k)|^2)^2]\,, \tag{2}$$

where $\gamma_s \triangleq \text{E}[|s(n)|^4]/\text{E}[|s(n)|^2]$ denotes the dispersion coefficient of the transmitted symbol sequence. Detailed studies of the stationary points of $J_{\text{cm}}(\mathbf{f})$ have been carried out in [2], [3], [4], under the customary assumption that the transmitted symbol sequence $s(n)$ is a *proper* [5] complex random process, i.e., its conjugate correlation function $R_{ss^*}(n,m) \triangleq \text{E}[s(n)\,s(n-m)] = 0$ for any $n, m \in \mathbb{Z}$ (examples of proper complex modulation schemes are PSK and QAM ones). In particular, under assumptions A1 and A2, it was demonstrated in [2], [3] that all the local minima of $J_{\text{cm}}(\mathbf{f})$ are *desired* ones, i.e., they are global and enable perfect recovery of the transmitted symbols in the absence of noise. However, in many cases of practical interest the symbol sequence $s(n)$ is an *improper* [5] random process, i.e., $R_{ss^*}(n,m) \neq 0$ for some $n, m \in \mathbb{Z}$. The simplest examples of improper modulations are all the real-valued ones. When real-valued modulations are considered, a common assumption [6], [7], [8], [9] is that the baseband channel impulse response (CIR) and the additive noise are also real-valued: in this case, the CM cost surface essentially exhibits the same characteristics of $J_{\text{cm}}(\mathbf{f})$ given by (2) when the symbol sequence is a proper complex process. A noticeable exception is [10], wherein the transmission of BPSK symbols over a complex baseband channel is considered. In this case, it was argued in [10] that, besides containing desired local minima, the infinite-length CM linear equalizer also exhibits *undesired* global

Manuscript received October 10, 2006; revised December 28, 2006. This work is partially supported by Italian National project Wireless 8O2.16 Multi-antenna mEsh Networks (WOMEN) under grant number 2005093248. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Richard Kozick.

The authors are with the Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni, Università degli Studi di Napoli Federico II, via Claudio 21, I-80125 Napoli, Italy [e-mail: (angelasara.cacciapuoti, f.verde)@unina.it].

Digital Object Identifier S 0000-0000(00)00000-0.

[1]Upper- and lower-case bold letters denote matrices and vectors; $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate, the transpose and the Hermitian operations; $|a|$ and $\angle a$ denote the magnitude and the phase of $a \in \mathbb{C}$; $\|\mathbf{a}\|$ is the Euclidean norm of $\mathbf{a} \in \mathbb{C}^m$; $\mathbf{0}_m \in \mathbb{R}^m$, $\mathbf{O}_{m\times n} \in \mathbb{R}^{m\times n}$ and $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ denote the null vector, the null and the identity matrix; $\mathbf{e}_\ell \in \mathbb{R}^m$ denotes the vector with 1 at the $(\ell+1)$th entry, for $\ell \in \{0, 1, \ldots, m-1\}$, and zero elsewhere; $\text{E}[\cdot]$ and $j \triangleq \sqrt{-1}$ denote statistical averaging and imaginary unit.

$$\widetilde{J}_{\text{cm}}(\mathbf{q}) \triangleq J_{\text{cm}}(\mathbf{C}\,(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{q}) = \kappa_s \sum_{\ell=0}^{K-1} |q_\ell|^4 + 2\left(\underbrace{\sigma_s^2\,\|\mathbf{q}\|^2}_{\text{E}[|y(k)|^2]}\right)^2 + \Big|\underbrace{\sigma_s^2\,\mathbf{q}^H\mathbf{J}^*\mathbf{q}^*}_{\text{E}[y^2(k)]}\Big|^2 - 2\,\gamma_s\,\underbrace{\sigma_s^2\,\|\mathbf{q}\|^2}_{\text{E}[|y(k)|^2]} + \gamma_s^2 \quad (3)$$

$$\widetilde{\mathbf{g}}(\mathbf{q}) = 2\,\kappa_s\,\widetilde{\boldsymbol{\Sigma}}(\mathbf{q})\,\mathbf{q} + 4\,\sigma_s^4\,\|\mathbf{q}\|^2\,\mathbf{q} + 2\,\sigma_s^4\,(\mathbf{q}^T\mathbf{J}\,\mathbf{q})\,\mathbf{J}^*\mathbf{q}^* - 2\,\gamma_s\,\sigma_s^2\,\mathbf{q} \quad (4)$$

$$\widetilde{\boldsymbol{\mathcal{H}}}(\mathbf{q}) = 4\,\kappa_s\,\widetilde{\boldsymbol{\Sigma}}(\mathbf{q}) + 4\,\sigma_s^4\,\|\mathbf{q}\|^2\,\mathbf{I}_K + 4\,\sigma_s^4\,\mathbf{q}\,\mathbf{q}^H + 4\,\sigma_s^4\,\mathbf{J}^*\mathbf{q}^*\mathbf{q}^T\mathbf{J} - 2\,\gamma_s\,\sigma_s^2\,\mathbf{I}_K \quad (5)$$

minima, which do not lead to perfect symbol recovery in the absence of noise. However, the analysis carried out in [10] cannot be directly extended to multidimensional real-valued modulations, as well as to others improper complex-valued modulation schemes of practical interest. Therefore, the aim of this letter is to improve and generalize the results of [10], by determining a broad family of undesired local minima of (2) under the more general assumptions:

A3) besides fulfilling assumption A2, $s(n)$ is an improper [5] random process, with second-order moments $\sigma_s^2 \triangleq \text{E}[|s(n)|^2]$ and $\zeta_s(n) \triangleq \text{E}[s^2(n)] \neq 0$, $\forall n \in \mathbb{Z}$, whose improper nature arises from the conjugate symmetric relationship $s^*(n) = e^{j\,2\pi\beta n}\,s(n)$, $\forall n \in \mathbb{Z}$;

A4) the CIR $c_a(t)$ is complex-valued.

A large number of digital modulation schemes satisfy assumption A3, including ASK, BPSK, DBPSK, offset QPSK (OQPSK), offset QAM (OQAM), MSK and its variant Gaussian MSK (GMSK) (see [11] for a detailed discussion). Specifically, real modulation schemes, such as ASK, BPSK and DBPSK, fulfill assumption A3 with $\beta = 0$, whereas for complex modulation formats, such as OQPSK, OQAM, and MSK-type, it results that $\beta = 1/2$.

II. Analysis of the CM cost function

In the absence of noise, accounting for (1) and invoking assumptions A1, A2, A3 and A4, after tedious but straightforward algebraic manipulations (see also [1]), it can be shown that minimization of (2) with respect to $\mathbf{f}$ is equivalent to minimization of (3), reported at the top of the page, with respect to the *combined channel-equalizer* vector $\mathbf{q} \triangleq \mathbf{C}^H\mathbf{f} = [q_0, q_1, \ldots, q_{K-1}]^T \in \mathbb{C}^K$, where $\kappa_s = \sigma_s^2\,(\gamma_s - 3\,\sigma_s^2) < 0$ is the kurtosis of $s(n)$, whereas $\mathbf{J} \triangleq \text{diag}[1, e^{-j\,2\pi\beta}, \ldots, e^{-j\,2\pi\beta(K-1)}] \in \mathbb{C}^{K\times K}$ is a diagonal unitary matrix, i.e., $\mathbf{J}\,\mathbf{J}^* = \mathbf{J}^*\mathbf{J} = \mathbf{I}_K$. More precisely, if $\overline{\mathbf{q}} \in \mathbb{C}^K$ is a local minimum of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, then, under assumption A1, $\overline{\mathbf{f}} = \mathbf{C}\,(\mathbf{C}^H\mathbf{C})^{-1}\overline{\mathbf{q}} + \mathbf{f}_{\mathcal{N}}$, where $\mathbf{f}_{\mathcal{N}} \in \mathbb{C}^{NL_e}$ is an arbitrary vector belonging to the null space of $\mathbf{C}^H$, is a local minimum of $J_{\text{cm}}(\mathbf{f})$. Furthermore, observe that $\sigma_s^2\,\|\mathbf{q}\|^2$ in (3) represents the mean-output-energy $\text{E}[|y(k)|^2]$ of the equalizer output $y(k) = \mathbf{q}^H\mathbf{s}(k)$, whereas $\sigma_s^2\,\mathbf{q}^H\mathbf{J}^*\mathbf{q}^*$ coincides with the second-order moment $\text{E}[y^2(k)]$. It is noteworthy that, compared with expressions of the CM cost functions commonly encountered in the literature, eq. (3) is more general. Specifically, when both the transmitted symbols and the CIR are real-valued (i.e., $\beta = 0$ and $\mathbf{C} \in \mathbb{R}^{(NL_e)\times K}$), the cost function (3) ends up to that studied in [6], [7], [8], [9]. In fact, in this case, the combined channel-equalizer vector $\mathbf{q}$ turns out to be real-valued, too, i.e., $\mathbf{q} \in \mathbb{R}^K$, and, consequently, the second and third summand in (3) can be grouped together. Additionally, the CM cost function (3) is different from that studied in [2], [3]. Indeed, in these papers, it is assumed that the transmitted symbols are proper complex: in this case, the kurtosis of $s(n)$ assumes the form $\kappa_s = \sigma_s^2\,(\gamma_s - 2\,\sigma_s^2)$ and, most important, the third summand in (3) disappears, i.e., $\text{E}[y^2(k)] = 0$, $\forall k \in \mathbb{Z}$. Henceforth, the basic difference between (3) and the expressions of the CM cost functions considered in [2], [3], [6], [7], [8], [9] stems from the fact that, under assumptions A3 and A4, the third summand in (3), which arises as a consequence of the improper nature of $s(n)$, is nonzero and different from the second one.

A vector $\overline{\mathbf{q}} \in \mathbb{C}^K$ is a stationary point of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ if it is a solution of $\widetilde{\mathbf{g}}(\mathbf{q}) \triangleq \nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})] = \mathbf{0}_K$, where $\nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})]$ denotes the complex gradient of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ with respect to $\mathbf{q}^*$. Accounting for (3), one obtains (4) at the top of the page, with $\widetilde{\boldsymbol{\Sigma}}(\mathbf{q}) \triangleq \text{diag}[|q_0|^2, |q_1|^2], \ldots, |q_{K-1}|^2] \in \mathbb{R}^{K\times K}$. A stationary point $\overline{\mathbf{q}}$ is a local minimum of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ if the Hessian matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\mathbf{q}) \triangleq \nabla_{\mathbf{q}}\left\{\nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})]\right\} \in \mathbb{C}^{K\times K}$ is positive definite for $\mathbf{q} = \overline{\mathbf{q}}$. Accounting for (4), one has (5), shown at the top of the page. A useful property of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ can be demonstrated. By virtue of (4), the cost function (3) can be rewritten as $\widetilde{J}_{\text{cm}}(\mathbf{q}) = \frac{1}{2}\,\mathbf{q}^H\widetilde{\mathbf{g}}(\mathbf{q}) - \gamma_s\,\sigma_s^2\,\|\mathbf{q}\|^2 + \gamma_s^2$, where the identity $\mathbf{q}^H\widetilde{\boldsymbol{\Sigma}}(\mathbf{q})\,\mathbf{q} = \sum_{\ell=0}^{K-1}|q_\ell|^4$ has been used. Thus, if $\overline{\mathbf{q}}$ is a stationary point of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, i.e., $\widetilde{\mathbf{g}}(\overline{\mathbf{q}}) = \mathbf{0}_K$, one obtains

$$\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}) = \gamma_s\,(\gamma_s - \sigma_s^2\,\|\overline{\mathbf{q}}\|^2)\,, \quad (6)$$

which allows one to readily calculate the value of the CM cost function at any stationary point. As a by-product, since $\widetilde{J}_{\text{cm}}(\mathbf{q})$ is a nonnegative function, it follows from (6) that the mean-output-energy corresponding to any stationary point cannot be greater than γ_s. On the basis of (4) and (5), Theorem 1, which is shown at the top of the next page and whose proof is reported in Appendix I, provides a family of local minima of (3). Some comments are now in order. First of all, observe that when $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},1}$, the equalizer output is given by $y(k) = \overline{\mathbf{q}}_{\text{min},1}^H\,\mathbf{s}(k) = e^{j\,\theta}s(k - i_1)$, i.e., except for an arbitrary phase rotation, perfect symbol recovery is guaranteed: in this case, the CM behaves as a blind ZF equalizer, which completely suppresses ISI. It is worth noting that, unlike the CM cost functions studied in [2], [3], [6], [7], [8], [9], the function (3) does not exhibit the ISI-free local minima (7) when the transmitted symbols $s(n)$ are "Gaussian" [1], i.e., $\kappa_s = 0 \Leftrightarrow \gamma_s = 3\,\sigma_s^2$, or "super-Gaussian" [1], i.e., $\kappa_s > 0 \Leftrightarrow \gamma_s > 3\,\sigma_s^2$, as well as when $s(n)$ is sub-Gaussian with $-\sigma_s^4 \le \kappa_s < 0 \Leftrightarrow 2\,\sigma_s^2 \le \gamma_s < 3\,\sigma_s^2$. Additionally, observe that, accounting for (6), it follows that $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1}) = \gamma_s\,(\gamma_s - \sigma_s^2)$. In contrast, when $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},2}$ or $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},3}$, the equalizer output is contaminated by ISI, since a particular linear combination of the two transmitted symbols

Theorem 1: The CM cost function (3) has local minima at the following vectors:

$$\overline{\mathbf{q}}_{\text{min},1} = e^{j\,\theta}\,\mathbf{e}_{i_1}\,, \quad \text{when } \sigma_s^2 \le \gamma_s < 2\,\sigma_s^2\,, \tag{7}$$

$$\overline{\mathbf{q}}_{\text{min},2} = e^{j\,\theta}\sqrt{\frac{\gamma_s}{\gamma_s+\sigma_s^2}}\cdot\left[\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}\right]\,, \text{when } \sigma_s^2 < \gamma_s < 2\,\sigma_s^2\,, \tag{8}$$

$$\overline{\mathbf{q}}_{\text{min},3} = e^{j\,\theta}\cdot\left[\rho\,\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}e^{j\,\pi\beta(i_2-i_1)}\sqrt{1-\rho^2}\;\mathbf{e}_{i_2}\right]\,, \text{when } \gamma_s = \sigma_s^2\,, \tag{9}$$

with $\theta \in [0, 2\pi)$, $\ell_{i_1,i_2} \in \mathbb{Z}$, $i_1 \ne i_2 \in \{0, 1, \ldots, K-1\}$ and $0 < \rho < 1$.

$s(k-i_1)$ and $s(k-i_2)$, with $i_1 \ne i_2 \in \{0, 1, \ldots, K-1\}$, is extracted in these cases. Henceforth, different from [2], [3], [6], [7], [8], [9], the cost function (3) exhibits local minima that do not lead to perfect ISI suppression, even in the absence of noise. In particular, it is worth noting that, when $-2\,\sigma_s^4 < \kappa_s < -\sigma_s^4 \Leftrightarrow \sigma_s^2 < \gamma_s < 2\,\sigma_s^2$, the CM cost function exhibits the undesired local minima (8). In this case, relying on (6), it results that $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},2}) = \widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1})\cdot[\gamma_s/(\gamma_s+\sigma_s^2)] < \widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1})$, which shows that, surprisingly enough, the ISI-free local minima (7) are not global. On the other hand, when $\kappa_s = -2\,\sigma_s^4 \Leftrightarrow \gamma_s = \sigma_s^2$, a situation occurring when $s(n)$ is constant modulus, accounting for (6), the value of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ at the undesired local minima (9) is given by $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},3}) = 0$. In this case, $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1})$ turns out to be zero as well and, hence, both the desired (7) and undesired (9) local minima are global. It should be observed that, by setting $\beta = 0$, $\ell_{i_1,i_2} = 1$ and $\rho = \cos(\phi)$, with $\phi \in [0, 2\,\pi)$, the expression of the undesired local minima (9) ends up to that derived in [10, eq. (15)] for the case of an infinite length CM equalizer, under the simplifying assumption of BPSK symbols with unitary variance. Finally, it is noteworthy that, with reference to real-valued symbols (i.e., $\beta = 0$), it was shown in [12] that the undesired minima (8) and (9) disappear by minimizing the CM cost function (2), provided that the equalizer output $y(k)$ is a *not* strictly linear function of $\mathbf{z}(k)$, namely, $y(k) = \text{Re}\{\mathbf{f}^H\,\mathbf{z}(k)\}$. More generally, if the transmitted symbols are improper complex, one has to use widely-linear equalizing structures [5], whereby the equalizer output is given by $y(k) = \mathbf{f}^H\mathbf{z}(k) + \mathbf{g}^H\mathbf{z}^*(k)$ and the CM cost function is minimized with respect to both $\mathbf{f}$ and $\mathbf{g}$, where $\mathbf{g} \in \mathbb{C}^{NL_e}$ is not necessarily constrained to be equal to $\mathbf{f}^*$. This issue is the topic of our current research.

To corroborate our analysis, the results of a Monte Carlo computer simulation are now presented. We consider both QPSK and OQPSK modulations, with $s(n)$ taking equiprobable values in $\{\pm 1, \pm j\}$, and a noise vector $\mathbf{v}(k)$ in (1) modeled as a zero-mean complex proper white random process, with autocorrelation matrix $\mathbf{R_{vv}} \triangleq \text{E}[\mathbf{v}(k)\,\mathbf{v}^H(k)] = \sigma_v^2\,\mathbf{I}_{NL_e}$. The signal-to-noise ratio (SNR) at the equalizer input is defined as $\text{SNR} \triangleq [\sigma_s^2/(N\,\sigma_v^2)]\cdot\sum_{q=0}^{L_c-1}\|\,\mathbf{c}(q)\,\|^2$ and is set to 25 dB. The received signal $r_a(t)$ is fractionally sampled at rate $2/T_s$ and the $\mathcal{Z}$-transforms of the two 3rd-order polyphase components $\widetilde{c}^{\,(\ell)}(q)$ are given by $\widetilde{C}^{\,(\ell)}(z) = (1 - 0.5\,e^{j\,\theta_{1,\ell}}z^{-1})(1 - 1.2\,e^{j\,\theta_{2,\ell}}z^{-1})$, for $\ell \in \{0, 1\}$, where $\theta_{1,0} = 0.7\,\pi$, $\theta_{2,0} = \theta_{1,0} + \pi$, $\theta_{1,1} = \theta_{1,0} + 0.2\,\pi$ and $\theta_{2,1} = \theta_{2,0} + 0.2\,\pi$. The minimization of the CM cost function (2) is adaptively carried out by resorting to the stochastic gradient descent algorithm [1], where $\gamma_s = \sigma_s^2 = 1$, $L_e = 5$, double-spike initialization is used and the step-size is continuously adjusted to achieve fast convergence without compromising stability. For each of the 10^3 Monte Carlo trials carried out, both the symbol and noise sequences are randomly and independently generated. Fig. 1 reports the signal-to-interference-plus-noise ratio (SINR) at the output of the CM equalizer as a function of the number of iterations, when either QPSK or OQPSK modulations are employed at the transmitter; for the sake of comparison, it is also reported the SINR (which is the same for QPSK and OQPSK modulations) at the output of the minimal-norm L-FIR ZF equalizer (synthesized by assuming perfect knowledge of $\mathbf{C}$). Results show that, when $s(n)$ is a proper random sequence (QPSK) and, thus, all the local minima of $J_{\text{cm}}(\mathbf{f})$ are desired, the performance of the CM equalizer rapidly improves as the number of iterations increases and becomes comparable to that of the ZF equalizer. In contrast, when the transmitted symbols are improper (OQPSK), due to the presence of the undesired global minima (9), the curve of the CM equalizer quickly saturates to a value that is significantly less (of about 50 dB) than the output SINR of the ZF equalizer.

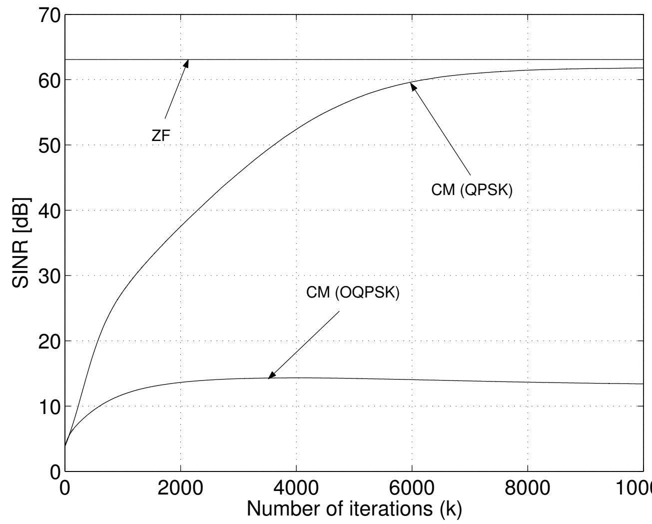

Fig. 1. SINR versus number of iterations.

Appendix I: proof of Theorem 1

We distinguish the following groups of stationary points.

G0. The only vector belonging to this group is $\overline{\mathbf{q}}_0 = \mathbf{0}_K$, which trivially fulfills $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{0}_K$.

G1. In this group, there are all the vectors satisfying $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{0}_K$ ensuring an ISI-free equalizer output, that is, they exhibit only one nonzero entry $\overline{q}_{i_1}$, with $i_1 \in \{0, 1, \ldots, K-1\}$, i.e., $\overline{\mathbf{q}} = \overline{q}_{i_1}\,\mathbf{e}_{i_1}$. Vector $\overline{\mathbf{q}} = \overline{q}_{i_1}\,\mathbf{e}_{i_1}$ satisfies $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{0}_K$ if and only if (iff) $\left[(\kappa_s + 3\,\sigma_s^4)\,|\overline{q}_{i_1}|^2 - \gamma_s\,\sigma_s^2\right]\overline{q}_{i_1}\,e^{j\,2\pi\beta i_1} = 0$; since $\overline{q}_{i_1}\,e^{j\,2\pi\beta i_1} \ne 0$ and $\kappa_s = \gamma_s\,\sigma_s^2 - 3\,\sigma_s^4$, this equation is

$$\begin{cases} \left[(\kappa_s+\sigma_s^4)\,|\overline{q}_{i_1}|^2+2\,\sigma_s^4\,(|\overline{q}_{i_1}|^2+|\overline{q}_{i_2}|^2)-\gamma_s\,\sigma_s^2\right]|\overline{q}_{i_1}|^2+\sigma_s^4\,(\overline{q}_{i_1}^*\,\overline{q}_{i_2})^2\,e^{j\,2\pi\beta(i_1-i_2)}=0 \\ \left[(\kappa_s+\sigma_s^4)\,|\overline{q}_{i_2}|^2+2\,\sigma_s^4\,(|\overline{q}_{i_1}|^2+|\overline{q}_{i_2}|^2)-\gamma_s\,\sigma_s^2\right]|\overline{q}_{i_2}|^2+\sigma_s^4\,(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}=0 \end{cases} \tag{10}$$

equivalent to $\gamma_s\,\sigma_s^2\,|\overline{q}_{i_1}|^2-\gamma_s\,\sigma_s^2=0 \Leftrightarrow |\overline{q}_{i_1}|=1$. Thus, the general expression of the vectors belonging to this group is $\overline{\mathbf{q}}_1=e^{j\,\theta}\,\mathbf{e}_{i_1}$, with $\theta\in[0,2\pi)$ and $i_1\in\{0,1,\ldots,K-1\}$.

G2. In this group, there are all the vectors satisfying $\widetilde{\mathbf{g}}(\mathbf{q})=\mathbf{0}_K$ leading to an ISI-contaminated equalizer output, i.e., the number of their nonzero entries is greater than one. To prove the existence of undesired local minima, it is sufficient to focus attention on those solutions of $\widetilde{\mathbf{g}}(\mathbf{q})=\mathbf{0}_K$ possessing only two nonzero entries $\overline{q}_{i_1}$ and $\overline{q}_{i_2}$, with $i_1\neq i_2\in\{0,1,\ldots,K-1\}$, i.e., $\overline{\mathbf{q}}=\overline{q}_{i_1}\,\mathbf{e}_{i_1}+\overline{q}_{i_2}\,\mathbf{e}_{i_2}$. After some algebraic manipulations, it can been seen that $\overline{\mathbf{q}}$ fulfills $\widetilde{\mathbf{g}}(\mathbf{q})=\mathbf{0}_K$ iff the system (10) at the top of the page is satisfied. Since the quantities enclosed in square brackets are real-valued, fulfillment of (10) requires that $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ be a real number, which happens when $\measuredangle\overline{q}_{i_1}-\measuredangle\overline{q}_{i_2}=\pi\,\beta\,(i_1-i_2)+\pi\,\ell_{i_1,i_2}$ or $\measuredangle\overline{q}_{i_1}-\measuredangle\overline{q}_{i_2}=\pi\,\beta\,(i_1-i_2)+\frac{\pi}{2}+\pi\,\ell_{i_1,i_2}$, with $\ell_{i_1,i_2}\in\mathbb{Z}$. In these cases, system (10) can be split up into the two different systems

$$\begin{cases}\gamma_s\,|\overline{q}_{i_1}|^2+\sigma_s^2\,\delta_{i_1,i_2}\,|\overline{q}_{i_2}|^2=\gamma_s \\ \sigma_s^2\,\delta_{i_1,i_2}\,|\overline{q}_{i_1}|^2+\gamma_s\,|\overline{q}_{i_2}|^2=\gamma_s\end{cases}, \quad \text{with } \delta_{i_1,i_2}=1,3\,, \tag{11}$$

which involve only the magnitudes of $\overline{q}_{i_1}$ and $\overline{q}_{i_2}$. Specifically, it results that $\delta_{i_1,i_2}=3$ when $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ is positive, whereas one has $\delta_{i_1,i_2}=1$ when $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ is negative. By resorting to the Cramer's rule, it is easily seen that, if $\gamma_s\neq\sigma_s^2\,\delta_{i_1,i_2}$, the solution of system (11) is unique and is given by $|\overline{q}_{i_1}|^2=|\overline{q}_{i_2}|^2=\gamma_s/(\gamma_s+\sigma_s^2\,\delta_{i_1,i_2})$. On the other hand, when $\gamma_s=\sigma_s^2\,\delta_{i_1,i_2}$, system (11) admits an infinite number of solutions characterized by the relation $|\overline{q}_{i_1}|^2+|\overline{q}_{i_2}|^2=1$. In summary, the general expressions of the vectors belonging to this group are given by $\overline{\mathbf{q}}_{2,1}=e^{j\,\theta}\sqrt{\gamma_s/(\gamma_s+3\,\sigma_s^2)}\cdot[\mathbf{e}_{i_1}+(-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}]$ and $\overline{\mathbf{q}}_{2,2}=e^{j\,\theta}\sqrt{\gamma_s/(\gamma_s+\sigma_s^2)}\cdot[\mathbf{e}_{i_1}-j\,(-1)^{\ell_{i_1,i_2}}e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}]$, for $\sigma_s^2<\gamma_s<3\,\sigma_s^2$, whereas, for $\gamma_s=3\,\sigma_s^2$, one obtains $\overline{\mathbf{q}}_{2,3}=e^{j\,\theta}\cdot[\rho\,\mathbf{e}_{i_1}+(-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\sqrt{1-\rho^2}\,\mathbf{e}_{i_2}]$ and, finally, for $\gamma_s=\sigma_s^2$, one has $\overline{\mathbf{q}}_{2,4}=e^{j\,\theta}\cdot[\rho\,\mathbf{e}_{i_1}-j\,(-1)^{\ell_{i_1,i_2}}e^{j\,\pi\beta(i_2-i_1)}\,\sqrt{1-\rho^2}\,\mathbf{e}_{i_2}]$, with $\theta\in[0,2\pi)$, $\ell_{i_1,i_2}\in\mathbb{Z}$, $i_1\neq i_2\in\{0,1,\ldots,K-1\}$ and $0<\rho<1$.

At this point, to find the local minima of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, we have to study the positive definiteness of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}})$ given by (5), evaluated at each of the stationary points previously derived.

G0. Since $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_0)=-2\,\gamma_s\,\sigma_s^2\,\mathbf{I}_K$, the cost function $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local maximum at $\overline{\mathbf{q}}_0=\mathbf{0}_K$.

G1. The matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ turns out to be diagonal, with diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)\}_{i_1+1,i_1+1}=2\,\sigma_s^2\,\gamma_s$ and $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)\}_{i+1,i+1}=-2\,\sigma_s^2\,(\gamma_s-2\,\sigma_s^2)$, for $i\in\{0,1,\ldots,K-1\}-\{i_1\}$. Hence, if $\gamma_s\geq 2\,\sigma_s^2$, the diagonal entries of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ take on both positive and negative values and, consequently, the vector $\overline{\mathbf{q}}_1$ is a saddle point. On the other hand, in accordance with assumption A2, if $\gamma_s<2\,\sigma_s^2$, the diagonal entries of the diagonal matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ are all positive and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_1$.

G2. First, the matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})$ is nonsingular, with diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})\}_{i+1,i+1}=2\,\gamma_s\,(\kappa_s+2\,\sigma_s^4)/(\gamma_s+3\,\sigma_s^2)$, for $i\in\{i_1,i_2\}$, and $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})\}_{i+1,i+1}=-2\,\gamma_s\,(\kappa_s+2\,\sigma_s^4)/(\gamma_s+3\,\sigma_s^2)$, for $i\in\{0,1,\ldots,K-1\}-\{i_1,i_2\}$. It is apparent that, regardless of κ_s, the matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})$ cannot be positive definite since its diagonal entries take on both positive and negative values and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a saddle point at $\overline{\mathbf{q}}_{2,1}$. Second, it can be seen that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})$ turns out to be diagonal, with diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})\}_{i+1,i+1}=2\,\gamma_s\,\sigma_s^2$, for $i\in\{i_1,i_2\}$, and $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})\}_{i+1,i+1}=-2\,\gamma_s\,\kappa_s/(\gamma_s+\sigma_s^2)$, for $i\in\{0,1,\ldots,K-1\}-\{i_1,i_2\}$. If assumption A2 is fulfilled, i.e., $\kappa_s<0$, the diagonal matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})$ is positive definite since its diagonal entries are all positive and, hence, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_{2,2}$. Third, it can be verified that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})$ is diagonal, with diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})\}_{i+1,i+1}=6\,\sigma_s^4$, for $i\in\{i_1,i_2\}$, and $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})\}_{i+1,i+1}=-2\,\sigma_s^4$, for $i\in\{0,1,\ldots,K-1\}-\{i_1,i_2\}$. Since the diagonal entries of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})$ assume both positive and negative values, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a saddle point at $\overline{\mathbf{q}}_{2,3}$. Finally, it results that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,4})$ is a diagonal matrix, with positive diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,4})\}_{i+1,i+1}=2\,\sigma_s^4$, for $i\in\{0,1,\ldots,K-1\}$ and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_{2,4}$.

References

[1] C.R. Johnson, 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. IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.

[2] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 818–826, Apr. 1996.

[3] D. Liu and L. Tong, "An analysis of constant modulus algorithm for array signal processing", *Eurasip Journal on Signal Processing*, vol. 73, pp. 81–104, Jan. 1999.

[4] D. Jamali and T. Ogunfunmi, "Stationary points of the finite length constant modulus optimization", *Eurasip Journal on Signal Processing*, vol. 82, pp. 625–641, Apr. 2002.

[5] P.J. Schreier and L.L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Processing*, vol. 51, pp. 714–725, Mar. 2003.

[6] C.R. Johnson and B.D.O. Anderson, "Godard blind equalizer error surface characteristics: white, zero-mean, binary source case", *Int. J. Adaptive Contr. Signal Processing*, vol. 9, pp. 301–324, July-Aug. 1995.

[7] J.P. Leblanc, I. Fijalkow, and C.R. Johnson, "CMA fractionally spaced equalizers: stationary points and stability under iid and temporally correlated sources", *Int. J. Adaptive Contr. Signal Processing*, vol. 12, pp. 135–155, Mar. 1998.

[8] H.H. Zeng, L. Tong, and C.R. Johnson, "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Information Theory*, vol. 44, pp. 1523–1538, July 1998.

[9] H.H. Zeng, L. Tong, and C.R. Johnson, "An analysis of constant modulus receivers," *IEEE Trans. Signal Processing*, vol. 47, pp. 2990–2999, Nov. 1999.

[10] C.B. Papadias, "On the existence of undesirable global minima of Godard equalizers," in *Proc. of Int. Conf. on Acoustic, Speech, and Signal Processing*, Munich, Germany, Apr. 21-24, 1997, pp. 3941–3944.

[11] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Subspace-based blind channel identification of SISO-FIR systems with improper random inputs", *Eurasip Journal on Signal Processing, Special Issue on Signal Processing in Communications*, vol. 84, pp. 2021–2039, Nov. 2004.

[12] A. Shah, S. Biracree, R.A. Casas, T.J. Endres, S. Hulyalkar, T.A. Schaffer, and C.H. Strolle, "Global convergence of a single-axis constant modulus algorithm," in *Proc. of the Tenth IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, USA, Aug. 14-16, 2000, pp. 645–649.