LABORATOIRE

INFORMATIQUE, SIGNAUX ET SYSTÈMES
DE SOPHIA ANTIPOLIS
UMR 6070

OPTIMAL STEP-SIZE CONSTANT MODULUS ALGORITHM

Vicente Zarzoso, Pierre Comon

Projet ASTRE

Rapport de recherche
ISRN I3S/RR–2004-23–FR

Septembre 2004

RÉSUMÉ :

L'algorithme a module constant (CMA) est la methode iterative la plus repandue pour l'egalisation aveugle de canaux de communication. Son implantation implique le choix d'une longueur de pas, constante, dont les performances dependent etroitement. Ce type d'algorithme, qui avait sa raison d'etre lorsque la puissance des calculateurs embarques etait limitee, se justifie difficilement; pourtant il continue a etre largement utilise. Dans cette contribution, on calcule le pas conduisant au minimum absolu du critere selon une direction de recherche a chaque iteration, et on montre a travers des simulations informatiques que l'algorithme CMA a pas optimal converge evidemment bien plus vite que sa version a pas constant, mais aussi qu'il permet dans la plupart des cas d'eviter le piege des minima locaux. De plus l'accroissement de complexite est relativement faible.

MOTS CLÉS :

CMA, Egalisation aveugle SISO, non gaussien

ABSTRACT:

The constant modulus algorithm (CMA) is the most widespread iterative method for blind equalization of digital communication channels. Its implementation typically involves a constant step-size parameter, on which the algorithm's performance strongly depends. This type of algorithm, whose use was justified when computational power was limited, is much less justified nowadays; yet, it is still widely employed. In this Letter, the step size with globally minimizes the cost function along the search direction is calculated analytically at each iteration. Illustrative experiments demonstrate that the resulting optimal step-size CMA converges much faster and is more robust to initialization than its constant step-size counterpart. These benefits are achieved at only a modest increase in computational complexity.

KEY WORDS :

CMA, SISO Blind Equalization, non Gaussian

Optimal Step-Size Constant Modulus Algorithm†

Vicente Zarzoso[1*] and Pierre Comon[2]

[1] Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool L69 3GJ, UK
vicente@liv.ac.uk

[2] Laboratoire I3S, Les Algorithmes – Euclide-B, BP 121, 06903, Sophia Antipolis, France
comon@i3s.unice.fr

Abstract

The constant modulus algorithm (CMA) is the most widespread iterative method for blind equalization of digital communication channels. Its implementation typically involves a constant step-size parameter, on which the algorithm's performance closely depends. In the past, this type of algorithm was justified on the grounds of limited computational power, but it continues to be widely employed nowadays despite its well-known drawbacks, such as convergence to local extrema and trade-off between convergence rate and accuracy. These shortcomings can be alleviated with the technique developed in this paper, whereby the step size leading to the absolute minimum of the CM criterion along the search direction is calculated algebraically at each iteration. Experiments demonstrate that the resulting optimal step-size CMA (OS-CMA) avoids the local minima more often and converges in fewer iterations than its constant step-size counterpart or the recently proposed recursive least squares CMA (RLS-CMA). The potential increase in complexity introduced by the optimal step-size calculation is relatively modest, and can easily be afforded by the computational power currently available.

Index Terms

Blind equalization, constant modulus algorithm, optimum adaption coefficient, algebraic methods, steepest descent minimization.

† Submitted to *IEEE Transactions on Communications*.

* Supported through a Post-Doctoral Research Fellowship awarded by the Royal Academy of Engineering of the UK.

I. Introduction

An important problem in digital communications is the recovery of the data symbols transmitted through a distorting medium. The constant modulus (CM) criterion is arguably the most widespread blind channel equalization principle [1], [2]. The CM criterion generally presents local extrema — often associated with different equalization delays — in the equalizer parameter space [3]. This shortcoming renders the performance of gradient-based implementations, such as the well-known constant modulus algorithm (CMA), very dependent on the equalizer impulse response initialization. Even when the absolute minimum is found, convergence can be severely slowed down for initial equalizer settings with trajectories in the vicinity of saddle points [4], [5]. The constant value of the step-size parameter (or adaption coefficient) must be carefully selected to ensure a stable operation while balancing convergence rate and final accuracy (misadjustment or excess mean square error). The stochastic gradient CMA (SG-CMA) drops the expectation operator and approximates the gradient of the criterion by a one-sample estimate, much in the LMS fashion. This rough approximation generally leads to slow convergence and poor misadjustment, even if the step size is carefully selected.

The most attractive feature of the SG-CMA lies in its simplicity, and indeed its use was rightly justified on the grounds of the limited computational power available in the past. For the same reasons, strategies proposed to alleviate the drawbacks of this type of algorithms were then deemed to be impractical. However, the enhanced computing capabilities of current digital signal processors (DSPs) enable the practical implementation of some of these strategies. Block (or fixed-window) methods obtain a more precise gradient estimate from a batch of channel output samples, improving convergence speed and accuracy [6]. Tracking capabilities are preserved as long as the channel remains stationary over the observation window. The block-gradient CMA (simply denoted as CMA hereafter) is particularly suited to burst-mode transmission systems. Unfortunately, the multimodal nature of the CM criterion sustains the negative impact of local extrema on block implementations. The recently proposed recursive least squares CMA (RLS-CMA) [7], which operates on a sample-by-sample basis, also proves notably faster and more robust than the SG-CMA. The derivation of the RLS-CMA relies on an approximation to the CM cost function in stationary or slowly varying environments, where block implementations may actually prove more efficient in exploiting the available information (the received signal burst). Moreover, the problems posed by local extrema are not addressed by the RLS approach.

Analytical solutions to the minimization of the CM criterion were developed in [8], [9]. After solving a linearized LS problem, these methods require to recover the right structure of the solution space when multiple equalization solutions exist. The structuring process is equivalent to the joint

diagonalization of the matrix set associated with the unstructured solution space. In the general case, this joint diagonalization can be achieved through a matrix QZ iteration for which convergence proof has not yet been found. In addition, special modifications are required for input signals with a one-dimensional (i.e., binary) alphabet [8]–[10].

Hence, the CM criterion is a rational function in several variables (the equalizer filter taps) whose exact minimization is computationally too expensive. Simpler alternatives based on gradient-descent or recursive implementations also present some deficiencies, as commented above. A judicious compromise consists of performing consecutive one-dimensional absolute minimizations of the cost function. This technique, known as exact line search or steepest descent, is generally considered inefficient [11]. However, it was first observed in [12] that the value of the adaption coefficient that leads to the absolute minimum of most blind cost functions along a given search direction can be computed algebraically. Although never implemented, it was conjectured that the use of this algebraic optimal step size could not only accelerate convergence but also avoid local extrema in some cases. The goal of the present paper is the theoretical development and experimental evaluation of the optimal step-size CMA (OS-CMA) derived from this idea. It is demonstrated that the OS-CMA shows higher immunity to local extrema than the CMA and RLS-CMA, and converges in fewer iterations at an affordable overall computational cost.

After briefly reviewing CM-based equalization in Section II, the OS-CMA is derived and analyzed in Section III. The experiments of Section IV evaluate the performance of the OS-CMA relative to the CMA and RLS-CMA. Conclusions are drawn in Section V.

II. Constant Modulus Equalization

Zero-mean data symbols $\{s_n\}$ are transmitted at a known baud-rate $1/T$ through a time dispersive channel with impulse response $h(t)$. The channel is assumed linear and time-invariant (at least over the observation window), with a stable, causal and possibly non-minimum phase transfer function, and comprises the transmitter pulse-shaping and receiver front-end filters. Assuming perfect synchronization and carrier-residual elimination, fractionally-spaced sampling by a factor of P yields the discrete-time channel output

$$\mathbf{x}_n = \sum_k \mathbf{h}_k s_{n-k} + \mathbf{v}_n \tag{1}$$

in which $\mathbf{x}_n = [x(nT),\, x(nT + T/P),\, \ldots,\, x(nT + T(P-1)/P)]^{\mathrm{T}} \in \mathbb{C}^P$, $x(t)$ denoting the continuous-time baseband received signal. Similar definitions hold for $\mathbf{h}_k$ and the additive noise $\mathbf{v}_n$. Eqn. (1) represents the so-called single-input multiple-output (SIMO) signal model, and reduces to the single-input single-output (SISO) model for $P = 1$. The SIMO model is also obtained if spatial

diversity (e.g., an antenna array) is available at the receiver, with or without time oversampling, and can easily be extended to the multiple-input (MIMO) case.

To recover the original data symbols from the received signal, a linear equalizer is employed with finite impulse response spanning L baud periods $\mathbf{f} = [\mathbf{f}_1^{\mathrm{T}}, \mathbf{f}_2^{\mathrm{T}}, \ldots, \mathbf{f}_L^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^D$, $D = PL$, $\mathbf{f}_k = [f_{k,1}, f_{k,2}, \ldots, f_{k,P}]^{\mathrm{T}} \in \mathbb{C}^P$, $k = 1, \ldots, L$. This filter produces the output signal $y_n = \mathbf{f}^{\mathrm{H}}\tilde{\mathbf{x}}_n$, where $\tilde{\mathbf{x}}_n = [\mathbf{x}_n^{\mathrm{T}}, \mathbf{x}_{n-1}^{\mathrm{T}}, \ldots, \mathbf{x}_{n-L+1}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^D$. The equalizer vector can be blindly estimated by minimizing the CM cost function [1], [2]:

$$J_{\mathrm{CM}}(\mathbf{f}) = \mathrm{E}\big\{\big(|y_n|^2 - \gamma\big)^2\big\} \tag{2}$$

where $\gamma = \mathrm{E}\{|s_n|^4\}/\mathrm{E}\{|s_n|^2\}$ is a constellation-dependent parameter. The CMA is a gradient-descent iterative procedure to minimize the CM cost. Its update rule reads

$$\mathbf{f}' = \mathbf{f} - \mu\mathbf{g} \tag{3}$$

where $\mathbf{g} \stackrel{\mathrm{def}}{=} \nabla J_{\mathrm{CM}}(\mathbf{f}) = 4\mathrm{E}\big\{(|y_n|^2 - 1)y_n^*\tilde{\mathbf{x}}_n\big\}$ is the gradient vector at point $\mathbf{f}$, and μ represents the step-size parameter. In the sequel, we assume that a block of length N_d baud periods $\mathbf{x}_n$ is observed at the channel output, from which $N = (N_d - L + 1)$ vectors $\tilde{\mathbf{x}}_n$ can be constructed.

III. Optimal Step-Size CMA

A. Steepest-Descent Minimization

Steepest-descent minimization consist of finding the absolute minimum of the cost function along the line defined by the search direction (typically the gradient) [11]:

$$\mu_{\mathrm{opt}} = \arg\min_{\mu} J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g}). \tag{4}$$

In general, exact line search algorithms are unattractive because of their relatively high complexity. Even in the one-dimensional case, function minimization must usually be performed using costly numerical methods. However, as originally observed in [12], the CM cost $J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g})$ is a rational function in the step size μ. Consequently, it is possible to find the optimal step size μ_{opt} in closed form among the roots of a polynomial in μ. Exact line minimization of function (2) can thus be performed at relatively low complexity.

B. Algebraic Optimal Step Size: the OS-CMA

In effect, some algebraic manipulations show that the derivative of $J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g})$ with respect to μ is the 3rd-degree polynomial

$$p(\mu) = d_3\mu^3 + d_2\mu^2 + d_1\mu + d_0 \tag{5}$$

with real-valued coefficients given by

$$d_3 = 2\mathrm{E}\{a_n^2\}, \qquad d_2 = 3\mathrm{E}\{a_n b_n\}, \qquad d_1 = \mathrm{E}\{2a_n c_n + b_n^2\}, \qquad d_0 = \mathrm{E}\{b_n c_n\} \tag{6}$$

where $a_n = |g_n|^2$, $b_n = -2\mathbb{R}\mathrm{e}(y_n g_n^*)$, and $c_n = (|y_n|^2 - \gamma)$, with $g_n = \mathbf{g}^{\mathrm{H}}\tilde{\mathbf{x}}_n$. Alternatively, the coefficients of the OS-CMA polynomial can be obtained as a function of the sensor-output statistics as:

$$\begin{aligned} &d_3 = C_{\mathbf{gggg}}, \qquad d_2 = -3\mathbb{R}\mathrm{e}(C_{\mathbf{gggf}}) \\ &d_1 = 2C_{\mathbf{ffgg}} + \mathbb{R}\mathrm{e}(C_{\mathbf{fgfg}}) - \gamma C_{\mathbf{gg}}, \qquad d_0 = \mathbb{R}\mathrm{e}(\gamma C_{\mathbf{fg}} - C_{\mathbf{fffg}}) \end{aligned} \tag{7}$$

where $C_{\mathbf{abcd}} \overset{\mathrm{def}}{=} \mathrm{E}\{\mathbf{a}^{\mathrm{H}}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\mathbf{b}\mathbf{c}^{\mathrm{H}}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\mathbf{d}\} = \sum_{ijkl}\mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}a_i^* b_j c_k^* d_l$, and $C_{\mathbf{ab}} \overset{\mathrm{def}}{=} \mathbf{a}^{\mathrm{H}}\mathbf{R}_{\tilde{x}}\mathbf{b}$, with $\mathbf{R}_{\tilde{x}} = \mathrm{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\}$ denoting the sensor-output covariance matrix. The latter procedure needs to compute in advance the sensor-output covariance matrix $\mathbf{R}_{\tilde{x}}$ and 4th-order moments $\mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}$, $1 \leqslant i, j, k, l \leqslant D$. Coefficients (6)–(7) are derived in Appendix A.

Having obtained its coefficients through any of the above equivalent procedures, the roots of polynomial (5) can be extracted as explained in the next section. The optimal step size corresponds to the root attaining the lowest value of the cost function, thus accomplishing the *global* minimization of J_{CM} in the gradient direction. Once μ_{opt} has been determined, the filter taps are updated as in (3), and the process is repeated with the new filter and gradient vectors, until convergence. This algorithm is referred to as *optimal step-size CMA (OS-CMA)*.

To improve numerical conditioning in the determination of μ_{opt}, gradient vector $\mathbf{g}$ should be normalized beforehand. This normalization does not cause any adverse effects since the relevant parameter is the search direction $\tilde{\mathbf{g}} = \mathbf{g}/\|\mathbf{g}\|$. Accordingly, vector $\mathbf{g}$ is substituted by $\tilde{\mathbf{g}}$ to compute polynomial coefficients (6)–(7) and update rule (3).

C. Root Extraction

The roots of polynomial (5) can be found through standard analytical procedures such as Cardano's formula, or more efficient iterative methods [13], [14]. The MATLAB code of a general algorithm for extracting the roots of a 3rd-degree polynomial is given in Appendix B [11]. This code is valid for polynomials with real or complex coefficients. When the coefficients of the step-size polynomial are real-valued [cf. (6)–(7)], further simplifications may still be introduced to this simple routine.

Concerning the nature of the roots, only two options are possible: either all three roots are real, or one is real and the other two form a complex conjugate pair. In the first case, one needs to check which of the tree real roots provide the lowest value of $J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g})$. In the second case, the real root typically provides the lowest value of the cost function. In scenarios composed of real-valued signals

and filters, one of the complex roots sometimes provides the lowest J_{CM}. However, an increase in the equalizer output mean square error (MSE) is also observed, which invalidates that option; the real root should always be preferred.

D. Convergence Analysis

Although a rigorous theoretical analysis of the OS-CMA convergence characteristics is beyond the scope of this paper, a brief preliminary comment can already be made in this section.

By design of steepest-descent methods, gradient vectors at consecutive iterations are orthogonal, which, depending on the initialization and the shape of the cost-function surface, may slow down convergence [11]. In the OS-CMA, gradient orthogonality is expressed as $\mathbb{R}\mathrm{e}(\mathbf{g}^{\mathrm{H}}\mathbf{g}') = 0$, with $\mathbf{g}' = \nabla J_{\mathrm{CM}}(\mathbf{f}')$. In our experiments, the OS-CMA always converged in less iterations than its constant step-size counterpart. In addition, the frequency of misconvergence to local extrema is notably diminished with the use of the optimal step-size strategy. These claims will be empirically demonstrated in Section IV.

E. Computational Complexity

The computational load of the OS-CMA is mainly due to the calculation of the polynomial coefficients (6) or (7). In practice, mathematical expectation is approximated by sample averaging across the observed signal burst. The computational cost of these averages in (6) is of order $\mathrm{O}(ND)$ per iteration, for data blocks composed of N sensor vectors $\tilde{\mathbf{x}}_n$. The cost per iteration of the alternative procedure (7) is approximately of order $\mathrm{O}(D^4)$. Consequently, this latter method should be preferred over the former when $N > D^4$. However, the second procedure needs to compute in advance the sensor-output 4th-order statistics, $\mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}$, $1 \leqslant i, j, k, l \leqslant D$, incurring in an additional cost of $\mathrm{O}(ND^4)$ operations. This initial load may render the second method more costly even when $N > D^4$, especially if convergence is achieved in few iterations.

Table I provides more precise figures for the OS-CMA computational cost in terms of the number of real floating point operations or *flops* (a flop represents a multiplication followed by an addition; multiplies and divisions are counted as flops as well). Also shown are the values for other CM-based algorithms such as the SG-CMA, the (block-gradient) CMA, and the RLS-CMA. Real-valued signals and filters are assumed; analogous values can similarly be obtained for the complex-valued scenario. The cost of extracting the roots of the step-size polynomial does not depend on the relevant equalization parameters (D, N) and can be considered negligible (see Section III-C). The derivation of the flop counts is tedious but straightforward, and is hence omitted in this paper.

IV. Experimental Results

Some computer experiments demonstrate the faster convergence rate of the OS-CMA relative to the constant step-size block CMA and the RLS-CMA, and its ability to escape the attraction basin of undesired equilibria in the CM cost surface. Bursts of $N_d = 200$ baud periods are observed at the output of a $T/2$-spaced channel ($P = 2$) excited by a BPSK source ($\gamma = 1$) and corrupted by AWGN with 10-dB SNR. Iterations are stopped when $\|\mathbf{f}' - \mathbf{f}\|/\|\mathbf{f}\| < 0.1\mu/\sqrt{N}$, where $\|\cdot\|$ denotes the Euclidean norm, and μ is the constant step size chosen for the conventional CMA in each experiment. A higher bound of 1000 iterations is also set. For the sake of a meaningful comparison, the same signal bursts and termination test are used in all methods.

Experiment 1. The first channel has an impulse response $\mathbf{h}_a = [0.2,\, 0.5,\, 1,\, -0.1]^{\mathrm{T}}$ (see [5, Section 2.4, pp. 82–83]). For $L = 1$, the minimum mean square error (MMSE) equalizers for delays zero and one are $\mathbf{f}^{(0)}_{\mathrm{MMSE}} = [0.18,\, 1.54]^{\mathrm{T}}$ and $\mathbf{f}^{(1)}_{\mathrm{MMSE}} = [0.91,\, -0.31]^{\mathrm{T}}$, with theoretical MSE of -7.11 and -11.92 dB, respectively. Fig. 1a plots the contour lines of the CM cost function (2) in the equalizer parameter space for a received burst. Also shown are the trajectories of the constant step-size CMA (3) for different arbitrary initializations of the equalizer tap vector, with $\mu = 10^{-2}$. Over the 16 initial points, the average number of iterations and computational cost (total number of flops) necessary for convergence are shown in Table II. Convergence could be accelerated with a larger value of μ, but at the expense of compromising the algorithm's stability. Fig. 1b–c display the CM cost and output MSE history, respectively, for the initial points labelled A to D in Fig. 1a. Points A–B converge to the CM local minima (close to the suboptimal-delay MMSE equalizer), whereas points C–D do so to the CM global minima (near the optimal-delay MMSE equalizer). The equalizer output signal after convergence from point A is displayed in Fig. 1d. The suboptimal-delay equalizer is not able to sufficiently open the eye of the output signal for a successful symbol detection or transfer to decision-directed (DD) operation.

Under identical system conditions and signal realization, Fig. 2a shows the trajectories of the OS-CMA equalizer for the same set of initial points. The OS-CMA successfully avoids the local minima, converging in all tested cases near the optimum-delay MMSE equalizer. The algorithm's fast convergence (requiring only an average of 12 iterations, compared to the 245 iterations needed by the CMA) can be clearly appreciated in the CM-cost and output-MSE history plots of Fig. 2b–c. It is interesting to remark the trajectory from point A: after falling right next to a local minimum, the algorithm is able to traverse it, finally converging to an absolute minimum and providing the final equalizer output shown in Fig. 2d. The eye is now sufficiently open for DD-mode transfer, or even for direct symbol detection with an appropriate threshold.

Fig. 3 and the final column in Table II summarize the results by the RLS-CMA with the typical forgetting factor $\lambda = 0.99$ and inverse covariance matrix initialized at the identity ($\delta = 1$). The observed signal block is reused as many times as required. The method presents the same drawbacks as the conventional CMA, converging to the global solution only if properly initialized. Due to its sample-by-sample operation, the RLS-CMA shows the slowest convergence rate in terms of iterations; however, its low cost per iteration is able to maintain a satisfactory overall complexity.

Experiment 2. The performance of the OS-CMA in a higher-dimensional CM surface is illustrated with the channel $\mathbf{h}_b = [0.7571,\ -0.2175,\ 0.1010,\ 0.4185,\ 0.4038,\ 0.1762]^{\mathrm{T}}$ and a 4-tap equalizer ($L = 2$) with double-spike initialization $\mathbf{f}_0 = [1,\ 1,\ 0,\ 0]^{\mathrm{T}}/\sqrt{2}$ (second example of [5, Section 2.4, pp. 82–83]). This system presents the theoretical output MMSE vs. equalization delay profile of Fig. 4: delay 1 provides the best MMSE performance, delay 0 following close behind; the worst performance is obtained at delay 3. The CM-cost and output MSE evolution of the constant step-size CMA ($\mu = 0.5$), OS-CMA, and RLS-CMA ($\lambda = 0.99$) averaged over 1000 independent signal blocks are displayed in Figs. 5a–b. The normalized histogram of the equalization delay obtained by the three methods is shown in Fig. 5c, whereas Table III summarizes their computational cost. As already observed in [4], [5], CMA's convergence is slowed down by the presence of two saddle points. Moreover, the CMA often ends in a suboptimum equalization delay. By contrast, the OS-CMA seems to avoid the saddle areas and converges with higher probability near the optimum-delay MMSE equalizer in a few iterations. The RLS-CMA gets trapped in the worst equalization delay more often than the two other methods. The subsequent poor average MSE after convergence is not offset by the method's low complexity in this experiment. The CMA, OS-CMA and RLS-CMA converged to one of the best two equalization delays (0 and 1) with probability of 76%, 96.9% and 81.2%, respectively. The OS-CMA obtains the best performance at an affordable total cost which, due to its fast convergence, always lies well below that of the conventional CMA. Fig. 5d further illustrates the superiority of the OS-CMA for a particular signal burst.

Experiment 3. To test the methods' robustness to equalizer initialization in the previous setting, the equalizer filter taps are randomly drawn from a zero-mean unit-variance Gaussian distribution before processing each of the 1000 signal blocks. The same initialization is used for all methods. An adaption coefficient $\mu = 0.025$ is chosen to prevent divergence of the conventional block CMA. The results in Fig. 6 and Table IV indicate that the OS-CMA converges in about an order of magnitude fewer iterations and with higher probability (86.6%) near the optimal MMSE equalization delay than the CMA (67.8%) and the RLS-CMA (73.1%). Also, its computational cost is close to the latter's and notably improves the former's.

V. Conclusions

Global line minimization of the CM cost function can be carried out algebraically by finding the roots of a 3rd-degree polynomial with real coefficients. The closed-form expressions of these coefficients have been provided. Compared to the classic constant-step size CMA and the RLS-CMA, the optimum step-size CMA (OS-CMA) converges to the optimum MMSE delay equalizer with higher probability, thus yielding the best equalization performance among the tested methods. In addition, the OS-CMA shows the fastest convergence, though its overall complexity can sometimes exceed that of the RLS-CMA. Taking into account the computational power featured by current DSP hardware, this complexity increase is relatively modest. In consequence, the optimal step-size strategy arises as a promising practical approach to improving the performance of blind equalizers in burst-mode transmission systems. Although this paper has focused on the SIMO signal case, totally analogous results hold in the basic SISO scenario.

Future lines of inquiry include the incorporation of the optimum step-size scheme in alternative blind and semi-blind equalization criteria and its theoretical performance analysis.

Appendix A: Coefficients of Step-size Polynomial

Method 1:

Let $\mathbf{f}' = \mathbf{f} - \mu\mathbf{g}$. Then $J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{(|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}_n|^2 - \gamma)^2\}$. Calling $y_n = \mathbf{f}^{\mathrm{H}}\tilde{\mathbf{x}}_n$ and $g_n = \mathbf{g}^{\mathrm{H}}\tilde{\mathbf{x}}_n$, we have $|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}_n|^2 = \mu^2|g_n|^2 - 2\mu\mathbb{R}\mathrm{e}(y_n g_n^*) + |y_n|^2$. Hence, $J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{(a_n\mu^2 + b_n\mu + c_n)^2\}$, with $a_n = |g_n|^2$, $b_n = -2\mathbb{R}\mathrm{e}(y_n g_n^*)$ and $c_n = (|y_n|^2 - \gamma)$. Expanding the square results in $J_{\mathrm{CM}}(\mathbf{f}') = \mu^4\mathrm{E}\{a_n^2\} + 2\mu^3\mathrm{E}\{a_n b_n\} + \mu^2\mathrm{E}\{b_n^2 + 2a_n c_n\} + 2\mu\mathrm{E}\{b_n c_n\} + \mathrm{E}\{c_n^2\}$. Taking the derivative with respect to μ and eliminating common constant factors, we finally arrive at the polynomial with the coefficients shown in (6).

Method 2:

$J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{\left(|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}|^2 - \gamma\right)^2\} = \mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}|^4\} - 2\gamma\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}|^2\} + \gamma^2$. In the first place, $\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}|^2\} = \mathrm{E}\{\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\mathbf{f}'\} = \mu^2 C_{\mathbf{gg}} - 2\mu\mathbb{R}\mathrm{e}(C_{\mathbf{fg}}) + C_{\mathbf{ff}}$, where $C_{\mathbf{ab}} = \mathbf{a}^{\mathrm{H}}\mathbf{R}_{\tilde{x}}\mathbf{b}$, $\mathbf{R}_{\tilde{x}} = \mathrm{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\}$, $\mathbf{a}, \mathbf{b} \in \mathbb{C}^D$. Similarly, let us denote

$$C_{\mathbf{abcd}} = \mathrm{E}\{\mathbf{a}^{\mathrm{H}}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\mathbf{b}\mathbf{c}^{\mathrm{H}}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}\mathbf{d}\} = \sum_{i,j,k,l=1}^{D} \mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}a_i^* b_j c_k^* d_l, \qquad \text{with } \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{C}^D$$

which shows the symmetry properties $C_{\mathbf{abcd}} = C_{\mathbf{cdab}} = C_{\mathbf{cbad}} = C_{\mathbf{adcb}} = C_{\mathbf{badc}}^*$. Then, after some algebraic simplifications, we can express

$$\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\tilde{\mathbf{x}}|^4\} = \mu^4 C_{\mathbf{gggg}} - 4\mu^3\mathbb{R}\mathrm{e}(C_{\mathbf{gggf}}) + 2\mu^2[2C_{\mathbf{ffgg}} + \mathbb{R}\mathrm{e}(C_{\mathbf{fgfg}})] - 4\mu\mathbb{R}\mathrm{e}(C_{\mathbf{fffg}}) + C_{\mathbf{ffff}}.$$

Combining the previous expressions, taking the derivative with respect to variable μ and eliminating common constant factors, one arrives at the polynomial with the coefficients given in (7).

APPENDIX B: MATLAB CODE FOR 3RD-DEGREE POLYNOMIAL ROOT EXTRACTION

```
function r = pol3roots(pol);
%Roots of a 3rd-degree polynomial with real or complex coefficients.
%SYNTAX: r = pol3roots(pol);
%        r = [r1; r2; r3] : roots
%        pol = [a3, a2, a1, a0] : coefficients of polynomial
%                                 a3.x^3 + a2.x^2 + a1.x + a0.
pol = pol/pol(1);  %transform into monic polynomial
a = pol(2); b = pol(3); c = pol(4);
Q = (a^2 - 3*b)/9; R = (2*a^3 - 9*a*b + 27*c)/54;
Q3 = Q^3; R2 = R^2;
if isreal(Q) & isreal(R) & ( R2 < Q3 )
    th = acos(R/sqrt(Q3));
    r = -2*sqrt(Q)*cos((th + 2*pi*[-1:1]')/3) - a/3;
else
   sq = sqrt(R2 - Q3);
   A = -(R + sign(real(conj(R)*sq))*sq)^(1/3);
   if abs(A) < eps, B = 0; else B = Q/A; end
   i = sqrt(-1); ApB = A + B; AmB = i*sqrt(3)*(A - B);
   r = [ApB; (-ApB + AmB)/2; -(ApB + AmB)/2] - a/3;
end
```

REFERENCES

[1] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[2] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 2, pp. 459–472, Apr. 1983.

[3] Z. Ding, C. R. Johnson, and R. A. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2425–2432, Oct. 1992.

[4] S. Lambotharan, J. Chambers, and C. R. Johnson, "Attraction of saddles and slow convergence in CMA adaptation," *Signal Processing*, vol. 59, no. 3, pp. 335–340, June 1997.

[5] C. R. Johnson, P. Schniter, I. Fijalkow, L. Tong, *et al.*, "The core of FSE-CMA behavior theory," in *Unsupervised Adaptive Filtering, Vol. II: Blind Deconvolution*, S. S. Haykin, Ed. New York: John Wiley & Sons, 2000, ch. 2, pp. 13–112.

[6] P. A. Regalia, "A finite-interval constant modulus algorithm," in *Proc. ICASSP-2002, 27th International Conference on Acoustics, Speech and Signal Processing*, vol. III, Orlando, FL, May 13–17, 2002, pp. 2285–2288.

[7] Y. Chen, T. Le-Ngoc, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1452–1456, May 2005.

[8] A.-J. van der Veen and A. Paulraj, “An analytical constant modulus algorithm,” *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[9] K. Doğançay and R. A. Kennedy, “Least squares approach to blind channel equalization,” *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 1678–1687, Nov. 1999.

[10] A.-J. van der Veen, “Analytical method for blind binary signal separation,” *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 1078–1082, Apr. 1997.

[11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed. Cambridge, UK: Cambridge University Press, 1992.

[12] P. Comon, “Contrasts, independent component analysis, and blind deconvolution,” *International Journal of Adaptive Control and Signal Processing (Special Issue on Blind Signal Separation)*, vol. 18, no. 3, pp. 225–243, Apr. 2004.

[13] E. Durand, *Solutions numériques des équations algébriques*. Paris: Masson, 1960, vol. I.

[14] C. Lanczos, *Applied Analysis*. New York: Dover, 1988.

TABLE I

COMPUTATIONAL COST IN NUMBER OF FLOPS FOR SEVERAL CM-BASED ALGORITHMS. D: NUMBER OF TAPS IN EQUALIZER VECTOR; N: NUMBER OF DATA VECTORS IN OBSERVED SIGNAL BURST.

	SG-CMA	CMA	OS-CMA		RLS-CMA
FLOPS			Method 1	Method 2	
initialization	—	—	—	$N\left[\binom{D+3}{4}+\binom{D+1}{2}\right]$	—
per iteration	$2(D+1)$	$2N(D+1)$	$N(6D+15)+2D$	$6D^4+3D^2+2D+9$	$D(4D+7)$

TABLE II

AVERAGE COMPUTATIONAL COST FOR CONVERGENCE OF THE CM-BASED METHODS IN EXPERIMENT 1 (FIGS. 1–3).

	CMA	OS-CMA		RLS-CMA
COST		Method 1	Method 2	
iterations	245	12		426
total flops ($\times 10^3$)	294	64.8	3.1	12.8

TABLE III

AVERAGE COMPUTATIONAL COST FOR CONVERGENCE OF THE CM-BASED METHODS IN EXPERIMENT 2 (FIGS. 4–5).

	CMA	OS-CMA		RLS-CMA
COST		Method 1	Method 2	
iterations	127	24		40
total flops ($\times 10^3$)	252.7	186.5	47.4	3.7

TABLE IV

AVERAGE COMPUTATIONAL COST FOR CONVERGENCE OF THE CM-BASED METHODS IN EXPERIMENT 3 (FIGS. 4 AND 6).

	CMA	OS-CMA		RLS-CMA
COST		Method 1	Method 2	
iterations	565	38		286
total flops ($\times 10^3$)	1124.4	295.2	69.8	26.3

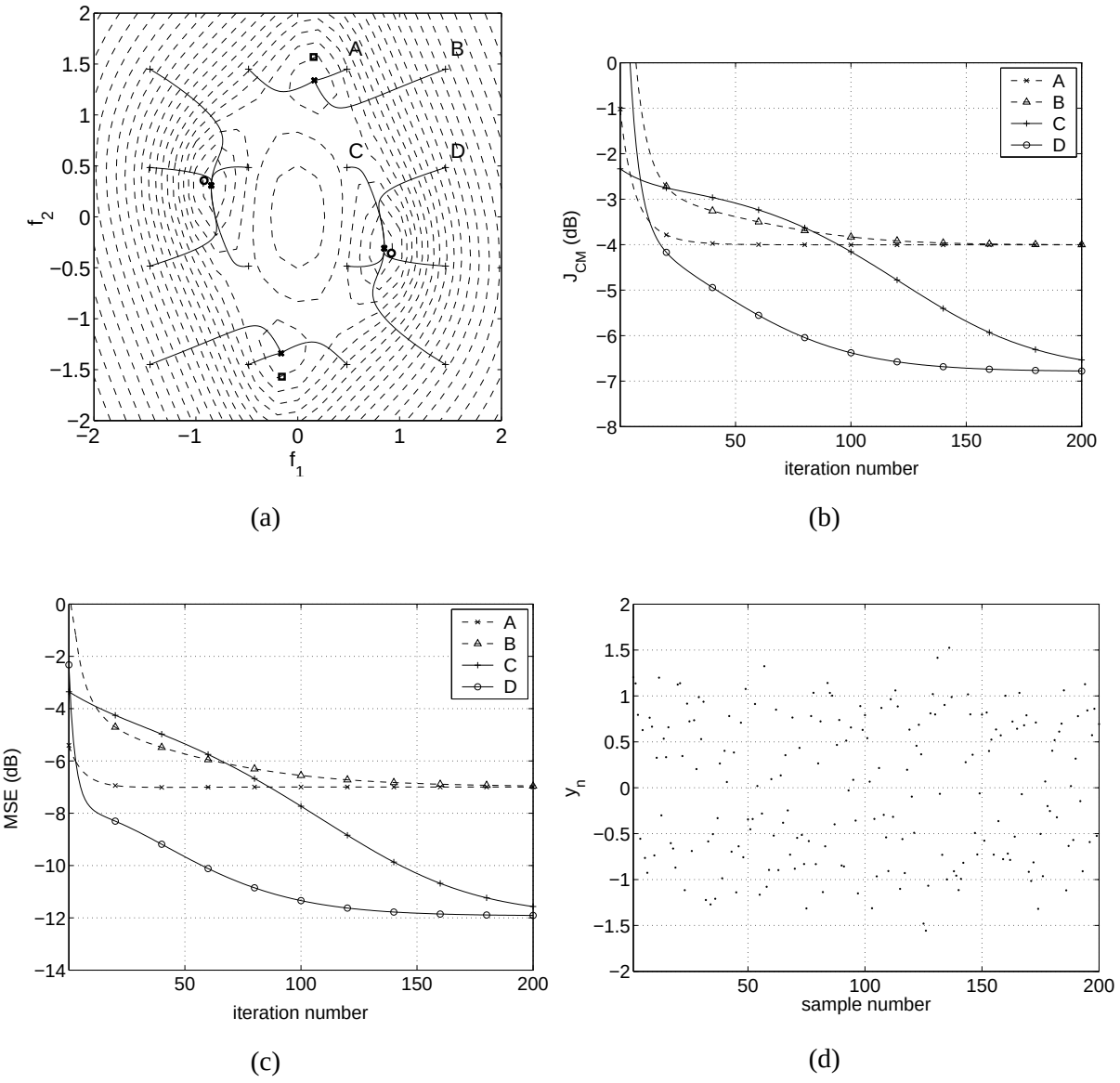

Fig. 1. Performance of the constant step-size CMA for channel $\mathbf{h}_a$, with $L = 1$ and $\mu = 10^{-2}$. (a) Contour lines and algorithm's trajectories for various equalizer initializations; '$+$': initial point; '$\times$': final point; '$\circ$': optimal-delay MMSE solution; '$\square$': suboptimal-delay MMSE solution. (b) CM-cost evolution for initial points A–D. (c) Equalizer output MSE evolution for initializations A–D. (d) Equalizer output signal after convergence from point A.

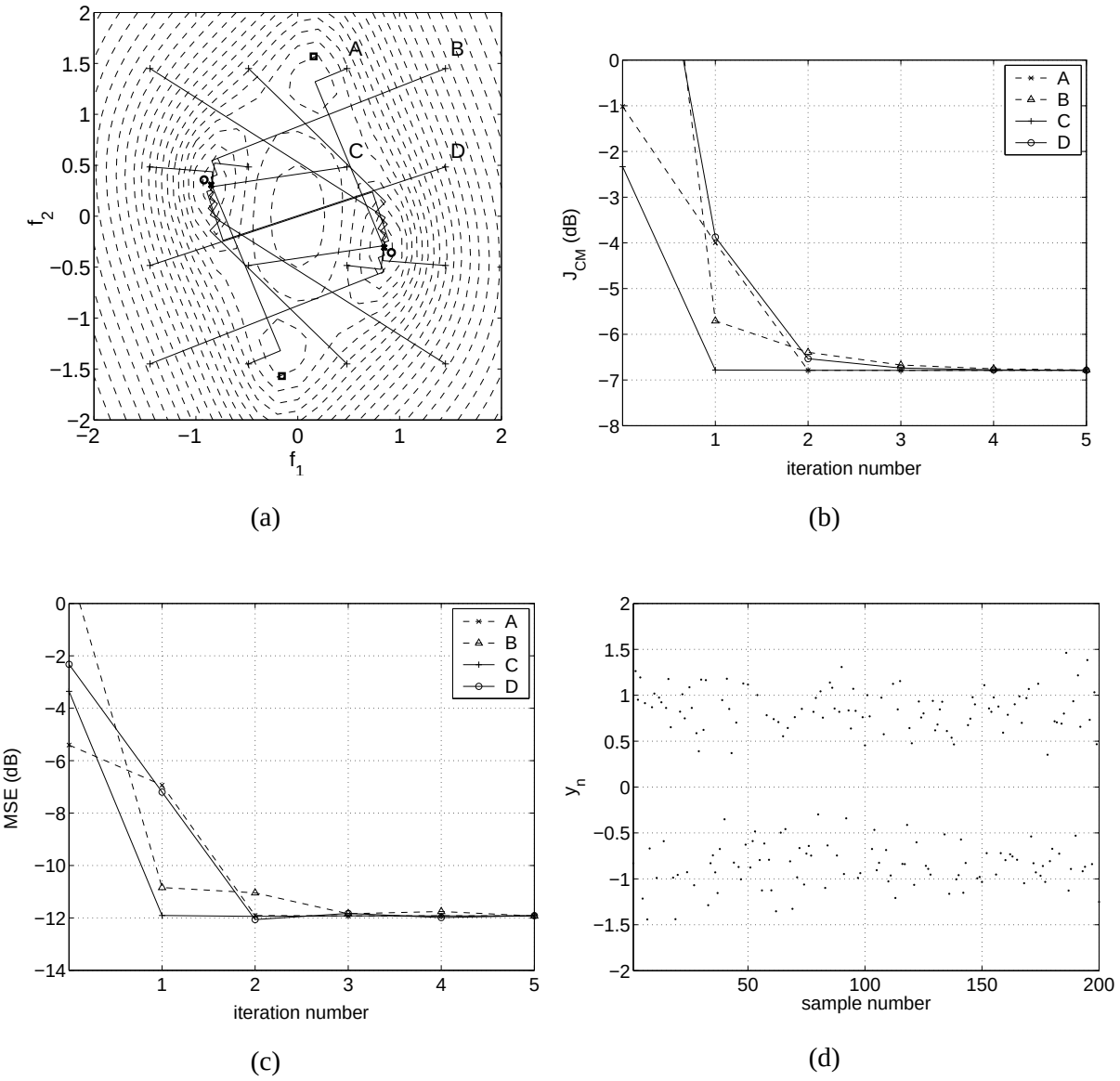

Fig. 2. Performance of the OS-CMA for channel $\mathbf{h}_a$, with $L = 1$. (a) Contour lines and algorithm's trajectories for various equalizer initializations; '$+$': initial point; '$\times$': final point; '$\circ$': optimal-delay MMSE solution; '$\square$': suboptimal-delay MMSE solution. (b) CM-cost evolution for initial points A–D. (c) Equalizer output MSE evolution for initializations A–D. (d) Equalizer output signal after convergence from point A.

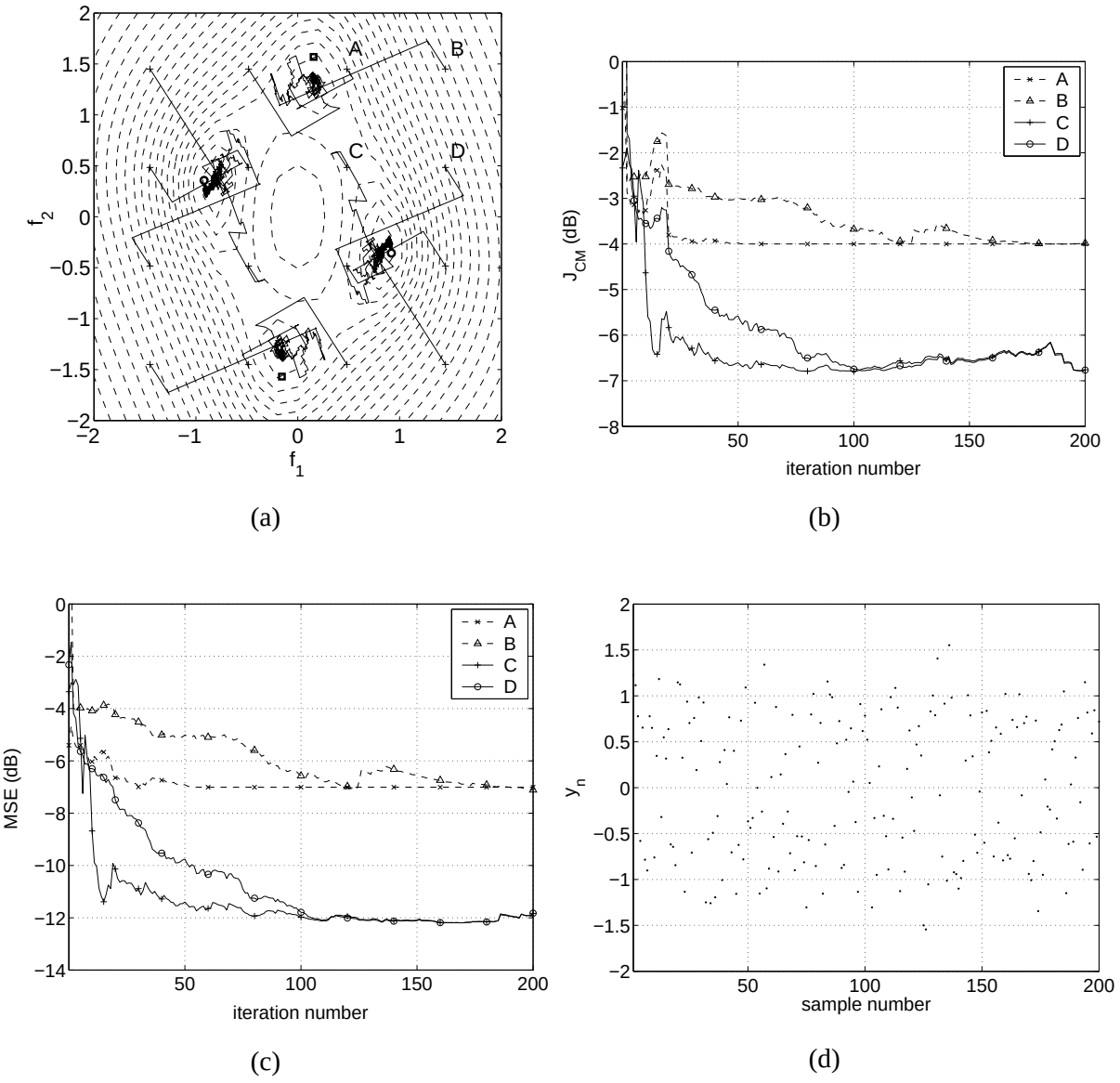

Fig. 3. Performance of the RLS-CMA for channel $\mathbf{h}_a$, with $L = 1$ and $\lambda = 0.99$. (a) Contour lines and algorithm's trajectories for various equalizer initializations; '$+$': initial point; '$\times$': final point; '$\circ$': optimal-delay MMSE solution; '$\square$': suboptimal-delay MMSE solution. (b) CM-cost evolution for initial points A–D. (c) Equalizer output MSE evolution for initializations A–D. (d) Equalizer output signal after convergence from point A.

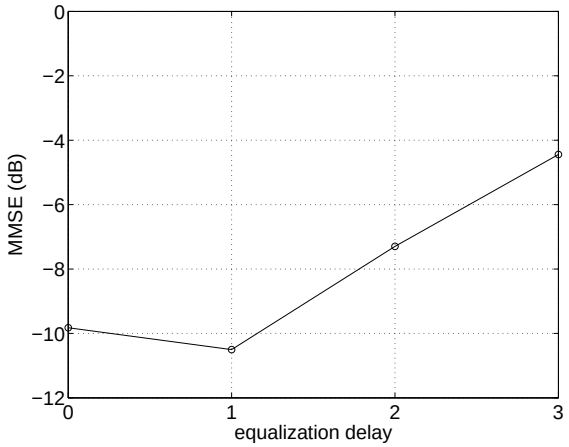

Fig. 4. Theoretical output MMSE as a function of the equalization delay for channel $\mathbf{h}_a$ with $L = 2$.

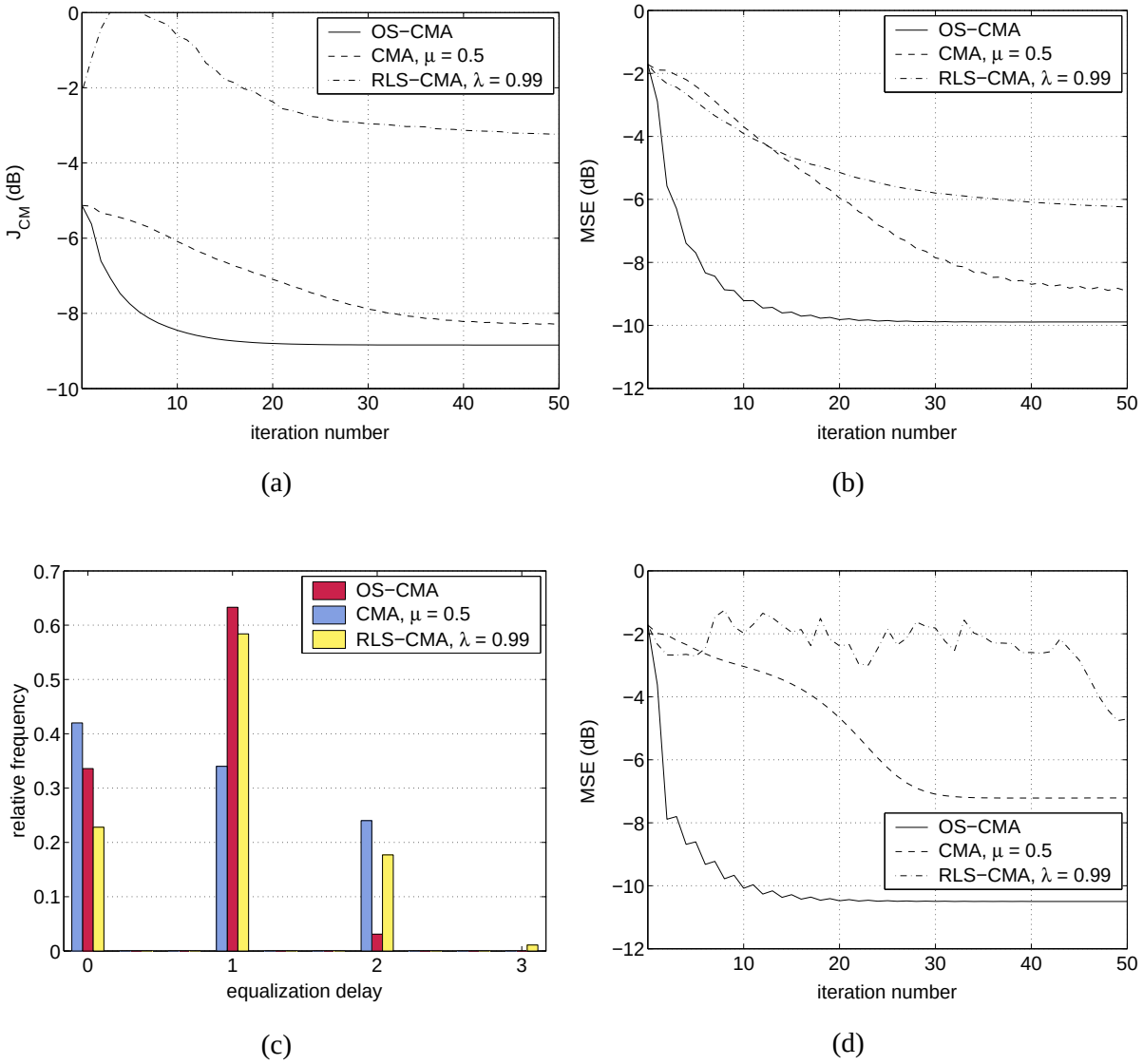

Fig. 5. Performance of the CMA ($\mu = 0.5$), OS-CMA and RLS-CMA ($\lambda = 0.99$) for channel $\mathbf{h}_a$, with $L = 2$ and double-spike first-tap initialization. (a) CM-cost evolution averaged over 1000 signal realizations. (b) Equalizer output MSE evolution averaged over 1000 signal realizations. (c) Normalized histogram of equalization delay after convergence. (d) Equalizer output MSE evolution for a particular signal realization.

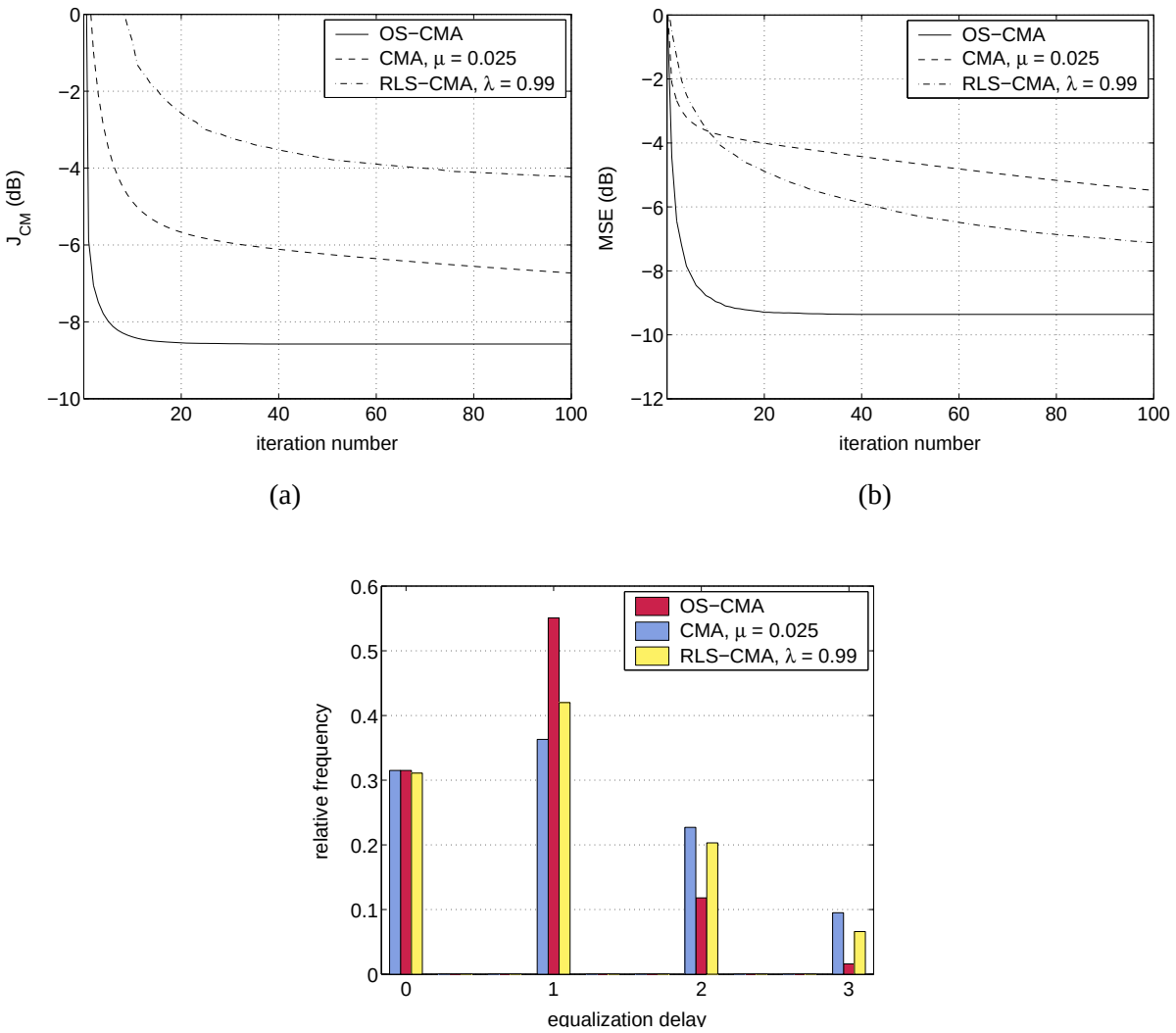

Fig. 6. Performance of the CMA ($\mu = 0.025$), OS-CMA and RLS-CMA ($\lambda = 0.99$) for channel $\mathbf{h}_a$, with $L = 2$ and random Gaussian equalizer initialization. (a) CM-cost evolution averaged over 1000 signal realizations. (b) Equalizer output MSE evolution averaged over 1000 signal realizations. (c) Normalized histogram of equalization delay after convergence.