

Optimal Step-Size Constant Modulus Algorithm

Vicente Zarzoso, Pierre Comon

▶ **To cite this version:**

Vicente Zarzoso, Pierre Comon. Optimal Step-Size Constant Modulus Algorithm. IEEE Transactions on Communications, 2008, 56 (1), pp.10-13. 10.1109/TCOMM.2008.050484 . hal-00327594

HAL Id: hal-00327594
https://hal.science/hal-00327594

Submitted on 8 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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
(now with the Laboratoire I3S, Université de Nice – Sophia Antipolis, France)
zarzoso@i3s.unice.fr

[2] Laboratoire I3S, CNRS/UNSA, 2000 route des Lucioles, BP 121, 06903 Sophia Antipolis Cedex, France
pcomon@i3s.unice.fr

Abstract

The step size leading to the absolute minimum of the constant modulus (CM) criterion along the search direction can be obtained algebraically at each iteration among the roots of a third-degree polynomial. The resulting optimal step-size CMA (OS-CMA) is compared with other CM-based iterative techniques in terms of performance-versus-complexity trade-off.

Index Terms

Adaption coefficient, blind equalization, CMA, exact line search, SISO and SIMO channels.

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 drops the expectation operator and approximates the gradient of the criterion by a one-sample estimate, as in LMS-based algorithms. This rough approximation generally leads to slow convergence and poor misadjustment, even if the step size is carefully chosen.

† Submitted to *IEEE Transactions on Communications* on Nov. 10, 2004 (manuscript number TCOM 04-0639); Nov. 2, 2005 (TCOM-05-0484); revised Nov. 3, 2006.

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

As opposed to recursive (or sample-by-sample) algorithms, 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. Moreover, sample-by-sample versions are easily obtained from block implementations by considering signal blocks of one data vector and iterating over consecutive received vectors. 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. Asymptotically (for sufficient block size), the least-squares CMA (LSCMA) [7] guarantees global convergence to a cost function stationary point, for any initial weight setting, with a cost per iteration similar to CMA's. This is achieved at the expense of an increased computational overhead due to the calculation of the data matrix pseudoinverse or its QR factorization, needed to solve the LS step at each iteration. In the QR-CMA method of [6], data prewhitening through the QR decomposition of the sensor-output matrix simplifies the block-CMA iteration, so that bounds on its step size can be found to ensure monotonic convergence. The recently proposed recursive least squares CMA (RLS-CMA) [8], which operates on a sample-by-sample basis, also proves notably faster and more robust than the classical 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). Interestingly, the RLS-CMA turns out to be equivalent to the recursive CMA (RCMA), put forward over a decade earlier in [9]; it also bears close resemblance to the orthogonalized CMA (O-CMA) of [10].

Analytical solutions to the minimization of the CM criterion are developed in [11], [12]. After solving a linearized LS problem, these methods require to recover the right structure of the solution space when multiple equalization solutions exist. In the general case, this can be achieved through a costly QZ matrix iteration. In addition, special modifications are required for input signals with a one-dimensional (i.e., binary) alphabet [11]–[13]. More importantly, these analytic methods aim at exact solutions to the CM criterion, which may yield suboptimal equalizers in the presence of noise.

A judicious alternative to existing techniques consists of performing consecutive one-dimensional absolute minimizations of the CM cost function. This technique, known as exact line search, is generally considered computationally inefficient [14]. However, it was first observed in [15] 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. 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 present Letter carries out a more detailed (yet concise) theoretical development and

experimental evaluation of the optimal step-size CMA (OS-CMA) derived from this idea, which was briefly presented in [16] under a different name.

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. The channel order is M baud periods. Assuming perfect synchronization and carrier-residual elimination, fractionally-spaced sampling by a factor of P yields the discrete-time channel output

$$\mathbf{x}_n = \sum_{k=0}^{M} \mathbf{h}_k s_{n-k} + \mathbf{v}_n \tag{1}$$

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 end, 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 $\mathbf{f} = [\mathbf{f}_1^{\mathrm{T}},\, \mathbf{f}_2^{\mathrm{T}},\, \ldots,\, \mathbf{f}_L^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^D$, $D = PL$, $\mathbf{f}_l = [f_{l,1},\, f_{l,2},\, \ldots,\, f_{l,P}]^{\mathrm{T}} \in \mathbb{C}^P$, $l = 1, \ldots, L$. This filter produces the output signal $y_n = \mathbf{f}^{\mathrm{H}} \tilde{\mathbf{x}}_n$, where $\tilde{\mathbf{x}}_n = [\mathbf{x}_n^{\mathrm{T}},\, \mathbf{x}_{n-1}^{\mathrm{T}},\, \ldots,\, \mathbf{x}_{n-L+1}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^D$. In these conditions, the channel effects can be represented by a block Toeplitz convolution matrix with dimensions $D \times (L+M)$ [3], [17].

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

$$\mathbf{f}(k+1) = \mathbf{f}(k) - \mu \mathbf{g}(k) \tag{3}$$

where $\mathbf{g} \stackrel{\mathrm{def}}{=} \nabla J_{\mathrm{CM}}(\mathbf{f}) = 4\mathrm{E}\big\{(|y_n|^2 - \gamma) y_n^* \tilde{\mathbf{x}}_n\big\}$ is the gradient vector at $\mathbf{f}$, symbol μ represents the step-size parameter and k denotes the iteration number. 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 $N = (N_d - L + 1)$ received data vectors $\tilde{\mathbf{x}}_n$ can be constructed.

III. Optimal Step-Size CMA

A. Exact Line Search

Exact line search consists of finding the absolute minimum of the cost function along the line defined by the search direction (typically the gradient) [14]:

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

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 [15] and later also remarked in [16], the CM cost $J_{\rm CM}(\mathbf{f} - \mu\mathbf{g})$ is a polynomial in the step size μ. Consequently, it is possible to find the optimal step size $\mu_{\rm opt}$ 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 $J_{\rm CM}(\mathbf{f} - \mu\mathbf{g})$ with respect to μ is the 3rd-degree polynomial with real-valued coefficients:

$$p(\mu) = d_3\mu^3 + d_2\mu^2 + d_1\mu + d_0$$

$$d_3 = 2\mathrm{E}\{a_n^2\}, \qquad d_2 = 3\mathrm{E}\{a_n b_n\}, \qquad d_1 = \mathrm{E}\{2a_n c_n + b_n^2\}, \qquad d_0 = \mathrm{E}\{b_n c_n\} \tag{5}$$

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 $g_n = \mathbf{g}^{\rm H}\tilde{\mathbf{x}}_n$. Alternatively, the coefficients of the OS-CMA polynomial can be obtained as a function of the sensor-output statistics, calculated before starting the iterative search. These two equivalent forms of the OS-CMA coefficients are derived in [18], [19].

Having obtained its coefficients, the roots of 3rd-degree polynomial (5) can be extracted with any standard algebraic procedures such as Cardano's formula, or more efficient iterative methods [20], [21].[1] The optimal step size corresponds to the root attaining the lowest value of the cost function, thus accomplishing the *global* minimization of $J_{\rm CM}$ in the gradient direction. When complex conjugate roots appear, the real root typically provides the lowest equalizer output mean square error (MSE). Once $\mu_{\rm opt}$ 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)*.

[1]The MATLAB code of a general algorithm for extracting the roots of a 3rd-degree polynomial is given in [18] (see also [14]).

To improve numerical conditioning in the determination of $\mu_{\rm opt}$, gradient vector $\mathbf{g}$ should be normalized. This normalization does not cause any adverse effects since the relevant parameter in the optimal step-size technique is the search direction $\tilde{\mathbf{g}} = \mathbf{g}/\|\mathbf{g}\|$.

C. Computational Complexity

Table I summarizes the OS-CMA's computational cost in terms of the number of real-valued 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 is the cost for other CM-based algorithms such as the CMA, the LSCMA [7], the QR-CMA [6] and the RLS-CMA [8], [9]. Complex-valued signals and filters are assumed; rough estimates of complexity for the real-valued signal scenario can be obtained by dividing the flop figures by 4. For typical values of (D, N), the OS-CMA is more costly per iteration over the observed signal block than the other CM-based algorithms except the RLS-CMA. The initial cost and the cost per iteration are of order $\mathrm{O}(D^4 N)$ and $\mathrm{O}(D^4)$, respectively, with the second form of the OS-CMA polynomial [18], [19].

IV. Experimental Results

We evaluate and compare the equalization quality as a function of computational cost (performance vs. complexity trade-off) achieved by the CM-based methods considered in this Letter. Bursts of $N_d = 200$ baud periods are observed at the output of a $T/2$-spaced channel ($P = 2$) excited by a QPSK source ($\gamma = 1$) and corrupted by complex circular additive white Gaussian noise with 20-dB SNR. For $L = 2$, these parameters result in an equalizer vector $\mathbf{f}$ composed of $D = 4$ taps. The channel impulse response coefficients are randomly drawn from a normalized complex Gaussian distribution. After a given number of iterations, performance is measured as the MSE between the equalizer output and the original channel input. Results are averaged over 1000 channel, source and noise realizations. For each plot in the figures, markers are placed at block iterations $[1, 2, 3, 5, 8, 14, 24, 41, 69, 118, 200]$. We set $\mu = 10^{-3}$ for the conventional fixed step-size CMA (a value found empirically to provide fastest performance while trying to avoid divergence in our simulation set-up), and the typical forgetting factor $\lambda = 0.99$ and inverse covariance matrix initialized at the identity for the RLS-CMA [8]. Double first-tap initializations are chosen for the equalizer vectors. Two scenarios are considered, depending on the linear invertibility of the channel matrix.

Scenario 1: linearly invertible channel. A channel order $M = 2$ yields an equivalent (4×4) channel convolution matrix that can be perfectly inverted in the absence of noise, thus guaranteeing the global convergence of the fractionally-spaced CMA [17]. Fig. 1(a) shows that the OS-CMA dramatically

outperforms the conventional fixed step-size CMA and slightly improves on the other CM-based methods at low complexity.

Scenario 2: lack of linear invertibility. A channel order $M = 4$ results in a (4×6) channel convolution matrix. Despite the lack of linear invertibility of the channel, a linear equalizer may still attempt to estimate the channel input at an extraction delay with reasonably low MSE. As shown in Fig. 1(b), the OS-CMA's quality-cost trade-off is only surpassed by the RLS-CMA's for sufficient complexity. In both scenarios, results at the reported 20-dB SNR level are quite representative of the methods' relative performance under the same fixed complexity over a wider $[0, 40]$-dB SNR range.

Optimal step-size trajectory. The average evolution of the OS-CMA's optimal step size in the above experiments is represented in Fig. 2. Depending on the cost function shape (which is determined by the actual channel, source and noise realizations), the optimal step size may take negative values at a given iteration. This fact may explain the peaks observed in the curves. Nevertheless, the optimal step size shows a monotonically decreasing trend.

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 resulting OS-CMA presents a performance versus complexity trade-off similar to the LSCMA [7], the QR-CMA [6] and the RLS-CMA [8], [9], slightly improving on those methods when perfect equalization conditions are not met. Due to space constraints, the numerical study presented in this Letter is of rather limited scope, and thus needs to be completed with a more thorough theoretical and experimental analysis of the OS-CMA technique evaluating its performance against a variety of system parameters such as block size, SNR, equalizer length, channel conditioning, etc. Indeed, additional experimental results reported in [18], [19] seem to point out that the optimal step-size strategy arises as a promising practical approach to improving the performance of blind equalizers in burst-mode transmission systems. The continuation of this work should also include the incorporation of the optimum step-size scheme in alternative blind and semi-blind criteria for equalization and beamforming. A first step in this direction has already been taken in [22], [23].

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] B. G. Agee, "The least-squares CMA: a new technique for rapid correction of constant modulus signals," in *Proc. ICASSP, 11th IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc.*, Tokyo, Japan, Apr. 7–11, 1986, pp. 953–956.

[8] 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 2004.

[9] R. Pickholtz and K. Elbarbary, "The recursive constant modulus algorithm: a new approach for real-time array processing," in *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, Nov. 1–3, 1993, pp. 627–632.

[10] R. Gooch and J. Lundell, "The CM array: an adaptive beamformer for constant modulus signals," in *Proc. ICASSP, 11th IEEE Intl. Conf. on Acoustics, Speech and Signal Processing*, Tokyo, Japan, Apr. 7–11, 1986, pp. 2523–2526.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] P. Comon, "Blind equalization with discrete inputs in the presence of carrier residual," in *Proc. ISSPIT, 2nd IEEE International Symposium on Signal Processing and Information Technology*, Marrakech, Morocco, Dec. 18–21, 2002.

[16] C. Xu and J. Li, "A batch processing constant modulus algorithm," *IEEE Communications Letters*, vol. 8, no. 9, pp. 582–584, Sept. 2004.

[17] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 818–826, Apr. 1996.

[18] V. Zarzoso and P. Comon, "Optimal step-size constant modulus algorithm," Laboratoire I3S, Sophia Antipolis, France, Tech. Rep. I3S/RR-2004-23-FR, Sept. 2004 (`http://www.i3s.unice.fr/~zarzoso/RR-04.23-V.ZARZOSO.pdf.gz`).

[19] ——, "Blind channel equalization with algebraic optimal step size," in *Proc. EUSIPCO, XIII European Signal Processing Conference*, Antalya, Turkey, Sept. 4–8, 2005.

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

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

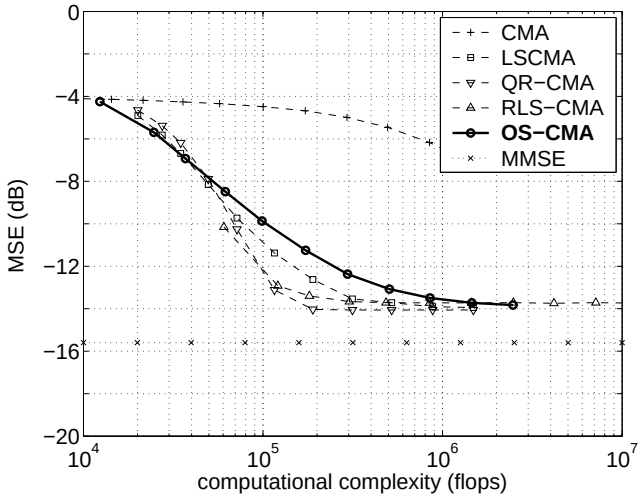
[22] V. Zarzoso and P. Comon, "Semi-blind constant modulus equalization with optimal step size," in *Proc. ICASSP, 30th Intl. Conf. on Acoustics, Speech and Signal Processing*, vol. III, Philadelphia, PA, Mar. 18–23, 2005, pp. 577–580.

[23] ——, "How fast is FastICA?" in *Proc. EUSIPCO, XIV European Signal Processing Conference*, Florence, Italy, Sept. 4–8, 2006.

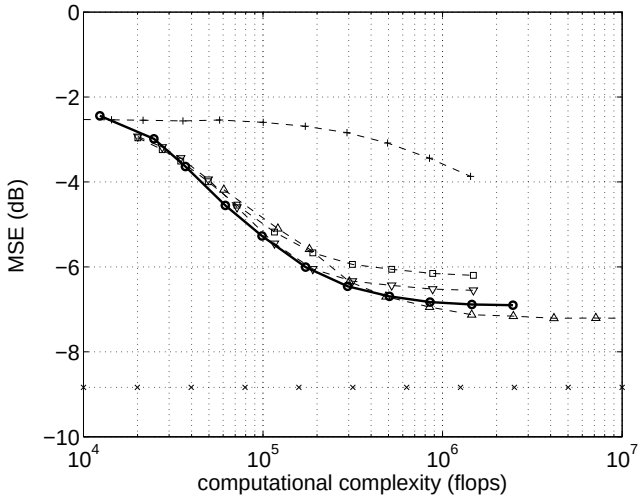
TABLE I

COMPUTATIONAL COST IN NUMBER OF REAL-VALUED FLOPS FOR SEVERAL CM-BASED ALGORITHMS. D: NUMBER OF TAPS IN EQUALIZER VECTOR; N: NUMBER OF DATA VECTORS IN OBSERVED SIGNAL BURST. THE BOTTOM HALF OF THE TABLE CORRESPONDS TO THE EXPERIMENTAL SET-UP OF SEC. IV AND FIGS. 1–2.

	Flops	CMA	LSCMA	QR-CMA	RLS-CMA	OS-CMA
(D, N)	initialization	—	$4D^2N$	$4D^2N$	—	—
	per block iteration	$4(2D+1)N$	$(8D+5)N$	$(8D+5)N$	$2D(7D+10)N$	$2(6D+7)N$
$(4, 199)$	initialization	0	12736	12736	0	0
	per block iteration	7164	7363	7363	60496	12338



(a)



(b)

Fig. 1. Performance vs. complexity trade-off of CM-based algorithms with QPSK source, signal bursts of $N_d = 200$ symbols, equalizer length $L = 2$ baud periods, oversampling factor $P = 2$, SNR = 20 dB, 1000 Monte Carlo runs. (a) Linearly invertible (4×4) channel convolution matrix (channel order $M = 2$). (b) Lack of linear invertibility of the channel, with a (4×6) channel convolution matrix (channel order $M = 4$).

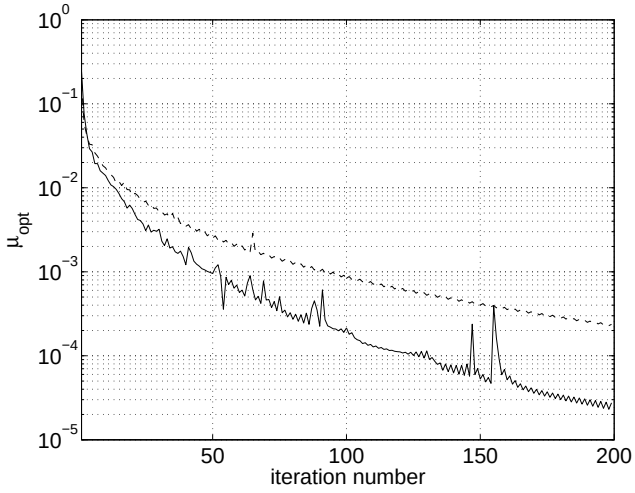


Fig. 2. Optimal step-size average trajectory in the simulation scenarios of Fig. 1(a) (dashed line) and Fig. 1(b) (solid line).