BLIND CHANNEL EQUALIZATION WITH ALGEBRAIC OPTIMAL STEP SIZE

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 arguably the most widespread iterative method for blind equalization of digital communication channels. The present contribution studies a recently proposed technique aiming at avoiding the shortcomings of conventional gradient-descent implementations. This technique is based on the computation of the step size leading to the absolute minimum of the CM criterion along the search direction. For the CM as well as other equalization criteria, this optimal step size can be calculated algebraically at each iteration by finding the roots of a low-degree polynomial. After developing the resulting optimal step-size CMA (OS-CMA), the algorithm is compared to its conventional constant step-size counterpart and more recent alternative CM-based methods. The optimal step size seems to improve the conditioning of the equalization problem as in prewhitening (e.g., via a prior QR decomposition of the data matrix), although it becomes more costly for long equalizers. The additional exploitation of the i.i.d. assumption through prewhitening can further improve performance, an outcome that had not been clearly interpreted in former works.

1. 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 probably the most widespread blind channel equalization principle [1]. The CM criterion generally presents local extrema — often associated with different equalization delays — in the equalizer parameter space [2]. 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 [3, 4]. Also, 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.

Block (or fixed-window) methods obtain a more precise gradient estimate from a batch of channel output samples, improving convergence speed and accuracy [5]. 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 in block implementations. The block CMA method of [5] is based on a preliminary QR decomposition of the data matrix, followed by power iterations on an equivalent kurtosis minimization criterion. An appropriate choice of the step size ensures the monotonic convergence of this algorithm (referred to as QR-CMA herein), although global convergence is not guaranteed. The recursive least squares CMA (RLS-CMA) [6], 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 explicitly addressed by the RLS approach. Another attempt to improve convergence is based on an adaptive control tuner that adjusts the second derivative of the equalizer tap estimates [7]. This accelerating adaptive filtering CMA (AAF-CMA) presents enhanced convergence rate and tracking capabilities relative to the SG-CMA, and is able to avoid shallow local extrema.

A recently proposed methodology to avoid the shortcomings derived from the multimodality of the CM criterion 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 [8]. However, it was first observed in [9] 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 contribution carries out the theoretical development and experimental evaluation of the optimal step-size CMA (OS-CMA) derived from this idea, which was briefly presented in [10] under a different name. The OS-CMA is then compared to other CM-based implementations such as the CMA, the QR-CMA, the RLS-CMA and the AAF-CMA.

2. 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, baud-spaced sampling yields the discrete-time channel output

$$x_n = \sum_k h_k s_{n-k} + v_n \tag{1}$$

in which $x_n = x(nT)$, $x(t)$ denoting the continuous-time baseband received signal. Similar definitions hold for h_k

---
* Royal Academy of Engineering Research Fellow.

To appear in *EUSIPCO-2005*, Antalya, Turkey, Sept. 4-8, 2005.

and the additive noise v_n. Eqn. (1) represents the so-called single-input single-output (SISO) signal model. This model applies to scenarios where diversity in the form of time oversampling or multiple receive sensors is not available. The interest in the SISO model lies in its 'hardness': in general, FIR channels cannot be perfectly equalized using FIR filters. By contrast, in multichannel configurations, giving rise to multiple-output models (SIMO, MIMO), FIR channels accept zero-forcing FIR equalizers under relatively mild length-and-zero conditions [11]. The results presented in this paper are easily transposable to multichannel models [10,13].

To recover the original data symbols from the received signal, a linear equalizer is employed with finite impulse response spanning L baud periods $\mathbf{f} = [f_1, f_2, \ldots, f_L]^{\mathrm{T}} \in \mathbb{C}^L$, This filter produces the output signal $y_n = \mathbf{f}^{\mathrm{H}}\mathbf{x}_n$, where $\mathbf{x}_n = [x_n, x_{n-1}, \ldots, x_{n-L+1}]^{\mathrm{T}} \in \mathbb{C}^L$. The equalizer vector can be blindly estimated by minimizing the CM cost function [1]:

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

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 $\mathbf{g} \stackrel{\mathrm{def}}{=} \nabla J_{\mathrm{CM}}(\mathbf{f}) = 4\mathrm{E}\big\{(|y_n|^2 - 1)y_n^*\mathbf{x}_n\big\}$ 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 x_n is observed at the channel output, from which $N = (N_d - L + 1)$ vectors $\mathbf{x}_n$ can be constructed.

3. OPTIMAL STEP-SIZE CMA

3.1 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) [8]:

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

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, it was originally observed in [9] that the CM cost $J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g})$ is a low-degree 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 simple polynomial in μ. Exact line minimization of function (2) can thus be performed at relatively low complexity.

3.2 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

$$\begin{array}{ll} d_3 = 2\mathrm{E}\{a_n^2\}, & d_2 = 3\mathrm{E}\{a_n b_n\} \\ d_1 = \mathrm{E}\{2a_n c_n + b_n^2\}, & d_0 = \mathrm{E}\{b_n c_n\} \end{array} \tag{6}$$

where $a_n = |g_n|^2$, $b_n = -2\mathrm{Re}(y_n g_n^*)$, and $c_n = (|y_n|^2 - \gamma)$, with $g_n = \mathbf{g}^{\mathrm{H}}\mathbf{x}_n$.

Alternatively, the coefficients of the OS-CMA polynomial can be obtained as a function of the sensor-output statistics as:

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

where

$$C_{\mathbf{abcd}} \stackrel{\mathrm{def}}{=} \mathrm{E}\{\mathbf{a}^{\mathrm{H}}\mathbf{x}\mathbf{x}^{\mathrm{H}}\mathbf{b}\mathbf{c}^{\mathrm{H}}\mathbf{x}\mathbf{x}^{\mathrm{H}}\mathbf{d}\} = \sum_{ijkl} \mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}a_i^*b_jc_k^*d_l$$

and $C_{\mathbf{ab}} \stackrel{\mathrm{def}}{=} \mathbf{a}^{\mathrm{H}}\mathbf{R}_{\tilde{x}}\mathbf{b}$, with $\mathbf{R}_{\tilde{x}} = \mathrm{E}\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\}$ denoting the sensor-output covariance matrix. This second procedure needs to compute in advance the sensor-output covariance matrix $\mathbf{R}_{\tilde{x}}$ and 4th-order moments $\mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}$, $1 \leqslant i, j, k, l \leqslant L$. Coefficients (6)–(7) are derived in the Appendix.

Having obtained its coefficients through any of the above equivalent procedures, the roots of polynomial (5) can be extracted as explained in Sec. 3.3. 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)*. Specifically, we call OS-CMA-1 the method resulting from coefficient computation (6), and OS-CMA-2 that obtained from (7). Note that both methods are equivalent in equalization performance and convergence rate measured in terms of iterations. The only difference lies in their computational cost in number of operations (Sec. 3.5).

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

3.3 Root Extraction

Standard analytical procedures such as Cardano's formula, or more efficient iterative methods [12], are readily available for obtaining the roots of 3rd-degree polynomial (5); an efficient MATLAB implementation, valid for polynomials with real or complex coefficients, is given in [13]. Concerning the nature of the roots, two options are possible: either all three roots are real, or one is real and the other two form a complex conjugate pair. If all three roots are real valued, we check which of the three real roots provides the lowest value of $J_{\mathrm{CM}}(\mathbf{f} - \mu\mathbf{g})$. In our experiments, when one root was real and the other two formed a complex conjugate pair, the real root typically provided the lowest value of the cost function. Even when the real root did not yield the lowest J_{CM}, it generally produced better output mean square error (MSE) than the complex roots. Hence, the real root should be preferred.

3.4 Preliminary Convergence Analysis

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 [8]. In the OS-CMA, gradient orthogonality is mathematically represented by relation $\mathrm{Re}(\mathbf{g}^{\mathrm{H}}\mathbf{g}') = 0$, 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 [13]. Likewise, fast convergece and improved stability have been independently reported in [10]. In addition, the frequency of misconvergence to local extrema is diminished with the use of the optimal step-size strategy, as empirically demonstrated in [13] and briefly in Section 4.

3.5 Computational Complexity

The computational load of the OS-CMA is mainly due to the calculation of the polynomial coefficients (6) or (7). Mathe-

Table 1: Computational cost in number of flops for several CM-based algorithms (single-input case). L: number of taps in equalizer vector; N: number of data vectors in observed signal burst.

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

matical expectations are in practice approximated by sample averaging across the observed signal burst. The computational cost of these averages in (6) is of order $\mathrm{O}(NL)$ per iteration, for data blocks composed of N sensor vectors $\mathbf{x}_n$. The cost per iteration of the alternative procedure (7) is approximately of order $\mathrm{O}(L^4)$. However, the second procedure needs to compute in advance the sensor-output 4th-order statistics, $\mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}$, $1 \leqslant i, j, k, l \leqslant L$, incurring in an additional cost of $\mathrm{O}(NL^4)$ operations. Depending on the number of iterations for convergence and the relative values of N and L, this initial load may render the second method more costly.

Table 1 provides the figures for the OS-CMA computational cost in terms of the number of real floating point operations or *flops* (a flop represents a multiplication or a division followed by an addition or a subtraction). Also shown are the values for other CM-based methods. Only dominant terms in the relevant parameters (L, N) are retained in the flop-count calculations. Real-valued signals and filters are assumed, although 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 (L, N) and can thus be considered negligible (see Section 3.3).

3.6 Variants

The algebraic optimal step-size technique can also be applied to other blind equalization criteria. The kurtosis maximization (KM, also known as Shalvi-Weinstein) criterion [14] can be globally minimized along a given direction by rooting a polynomial degree 5 in μ (details are omitted due to the lack of space). This would give rise to the OS-KMA, with a computational cost per iteration similar to that of the OS-CMA. The optimal step-size technique remains applicable if the received data are prewhitened, e.g., using a QR decomposition of the data matrix, as in the QR-CMA method of [5]. Accordingly, we refer to the optimal step-size KM algorithm with prewhitening as OS-QR-KMA. Prewhitening improves conditioning and may lead to faster convergence under the i.i.d. input assumption.

4. EXPERIMENTAL RESULTS

The following experiments evaluate the comparative performance of the OS-CMA. Bursts of $N_d = 200$ baud periods are observed at the output of a baud-spaced order-4 channel $(M = 4)$ excited by an i.i.d. BPSK source $(\gamma = 1)$ and corrupted by AWGN with 10-dB SNR. To test robustness to the channel configuration, the channel impulse response coefficients are randomly drawn from a zero-mean unit-variance real-valued Gaussian distribution before processing each of 500 independent signal bursts. The typical center-tap filter serves as equalizer tap vector initialization. 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. To limit complexity, a higher bound of $500L$ iterations is set. The final equalizer vector is scaled to provide the lowest MSE value among all possible extraction delays. The same signal bursts, channel impulse response, and termination test are used for all methods under study. Regarding the methods' parameters, an adaption coefficient $\mu = 10^{-4}$ is chosen in a bid to prevent divergence of the conventional block CMA. The QR-CMA operates with the optimal value of [5, Secs. 4–5] $(\alpha = 2/3)$. The RLS-CMA is run with the typical forgetting factor $\lambda = 0.99$ and inverse covariance matrix initialized at the identity $(\delta = 1)$. The values $m_1 = 0.15$, $\kappa = 100$, $\mu = 0.5$ are used for the AAF-CMA, as suggested in [7]. In the latter two methods, which operate on a sample-by-sample basis, the received signal block is re-used as many times as required.

The average output MSE after convergence as a function of the equalizer length L is shown in Fig. 1, where the same 500 signal bursts are used at each value of L. Also plotted as a reference is the performance of the minimum MSE (MMSE) equalizer with optimum delay. Since the optimum-delay MMSE equalizer typically lies close to the CM-cost global extrema [4], the distance to the MMSE-bound curve provides an indication of global convergence. The average overall computational complexity (flops) for convergence in the same experiment appears in Fig. 2. The complexity of the OS-QR-KMA is very close to that of the OS-CMA (with a small extra cost due to prewhitening) and has not been plotted for the sake of clarity.

The OS-CMA considerably improves its conventional constant-step counterpart and the AAF-CMA; also, it slightly outperforms the RLS-CMA over the whole equalizer-length range, and the QR-CMA for short equalizer lengths. Hence, the OS-CMA ability to escape local extrema [9, 13] seems more evident in lower-dimensional equalizer spaces. As expected, the OS-CMA-2 is more complex than the OS-CMA-1 for long equalizers, due to the extra complexity introduced by the computation of the sensor-output 4th-order moments before starting the iterations. The OS-CMA-1 complexity remains above that of the other non-conventional methods in this scenario. Nevertheless, the OS-CMA appears less complex than the conventional CMA, as it converges in over an order of magnitude fewer iterations. Just like the QR-CMA, the OS-QR-KMA takes advantage of both the constellation and the i.i.d. character of the input signal. With the incorporation of the algebraic optimal step-size, the OS-QR-KMA is able to outperform the QR-CMA, getting closer to the MMSE bound and requiring up to an order of magnitude less iterations, yet becoming more costly for longer equalizers.

5. 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 blind equalization algorithm, the OS-CMA, has been studied in this contribution, which expands the brief description of this technique independently developed in [10]. The OS-CMA clearly outperforms in equalization quality and cost the conventional constant step-size CMA; it is also able to improve other non-conventional methods for short equalizer lengths. The exploitation of the i.i.d. assumption through prewhitening (e.g., based on a QR decomposition of the data matrix) can further improve performance regardless of the criterion employed (CM, KM); this feature has not been clearly interpreted in previous works [5]. The optimal step size seems to have a conditioning effect similar to prewhitening, as both techniques yield very similar results, the former becoming less costly for short equalizer settings. The optimal step-size strategy, which is not exclusive to the CM criterion [15, 16], can also be easily implemented in semi-blind

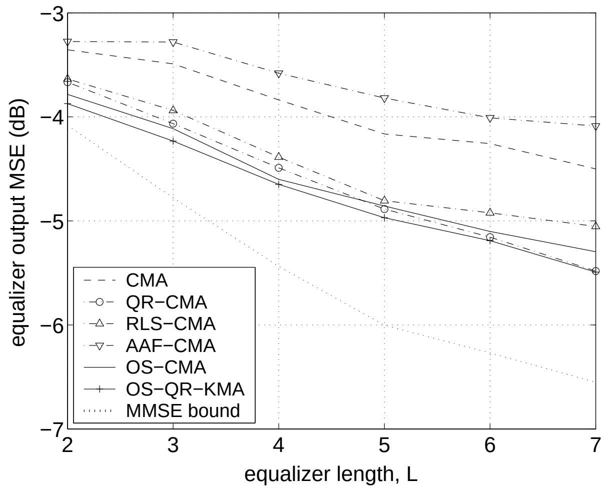

Figure 1: Equalizer output MSE after convergence.

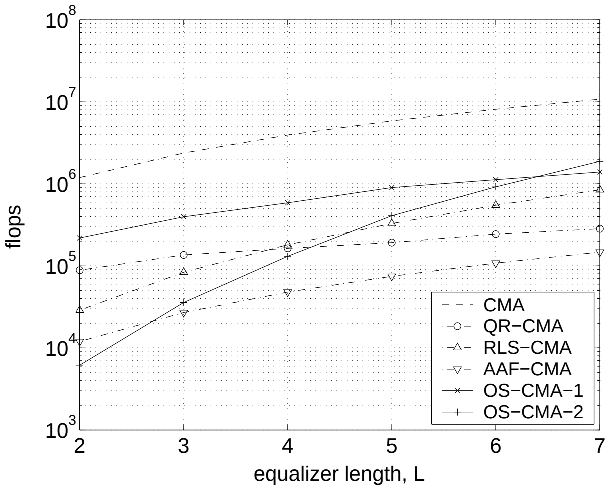

Figure 2: Computational cost for convergence.

operation [16, 17], and its extension to multichannel configurations (e.g., the SIMO model) is straightforward [13]. In consequence, this strategy arises as a promising approach to improving the performance of blind equalizers in burst-mode transmission systems. Further work should include a more comprehensive performance evaluation and comparison in a wider variety of equalization scenarios, and the search for new variants aiming at a reduced complexity in large equalizer spaces.

Appendix: Coefficients of Step-Size Polynomial

Method 1: Let $\mathbf{f}' = \mathbf{f} - \mu\mathbf{g}$. Then $J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{(|\mathbf{f}'^{\mathrm{H}}\mathbf{x}_n|^2 - \gamma)^2\}$. Calling $y_n = \mathbf{f}^{\mathrm{H}}\mathbf{x}_n$ and $g_n = \mathbf{g}^{\mathrm{H}}\mathbf{x}_n$, we have $|\mathbf{f}'^{\mathrm{H}}\mathbf{x}_n|^2 = \mu^2|g_n|^2 - 2\mu\mathrm{Re}(y_n g_n^*) + |y_n|^2$. Hence, $J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{(a_n\mu^2 + b_n\mu + c_n)^2\}$, with $a_n = |g_n|^2$, $b_n = -2\mathrm{Re}(y_n g_n^*)$ 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: $J_{\mathrm{CM}}(\mathbf{f}') = \mathrm{E}\{(|\mathbf{f}'^{\mathrm{H}}\mathbf{x}|^2 - \gamma)^2\} = \mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\mathbf{x}|^4\} - 2\gamma\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\mathbf{x}|^2\} + \gamma^2$. In the first place, $\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\mathbf{x}|^2\} = \mathrm{E}\{\mathbf{f}'^{\mathrm{H}}\mathbf{x}\mathbf{x}^{\mathrm{H}}\mathbf{f}'\} = \mu^2 C_{\mathbf{gg}} - 2\mu\mathrm{Re}(C_{\mathbf{fg}}) + C_{\mathbf{ff}}$, where $C_{\mathbf{ab}} = \mathbf{a}^{\mathrm{H}}\mathbf{R}_{\tilde{x}}\mathbf{b}$, $\mathbf{R}_{\tilde{x}} = \mathrm{E}\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\}$, $\mathbf{a}$, $\mathbf{b} \in \mathbb{C}^L$. Similarly, let us denote

$$C_{\mathbf{abcd}} = \mathrm{E}\{\mathbf{a}^{\mathrm{H}}\mathbf{x}\mathbf{x}^{\mathrm{H}}\mathbf{b}\mathbf{c}^{\mathrm{H}}\mathbf{x}\mathbf{x}^{\mathrm{H}}\mathbf{d}\} = \sum_{ijkl=1}^{L} \mathrm{E}\{\tilde{x}_i\tilde{x}_j^*\tilde{x}_k\tilde{x}_l^*\}a_i^* b_j c_k^* d_l$$

with $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{C}^L$, 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

$$\begin{aligned}\mathrm{E}\{|\mathbf{f}'^{\mathrm{H}}\mathbf{x}|^4\} &= \mu^4 C_{\mathbf{gggg}} - 4\mu^3\mathrm{Re}(C_{\mathbf{gggf}}) \\ &\quad + 2\mu^2[2C_{\mathbf{ffgg}} + \mathrm{Re}(C_{\mathbf{fgfg}})] - 4\mu\mathrm{Re}(C_{\mathbf{fffg}}) + C_{\mathbf{ffff}}.\end{aligned}$$

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).

Acknowledgement. The authors would like to thank one of the reviewers for pointing out reference [10].

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] 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.

[3] 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.

[4] 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.

[5] P. A. Regalia, "A finite-interval constant modulus algorithm," in *Proc. ICASSP-2002*, vol. III, Orlando, FL, May 13–17, 2002, pp. 2285–2288.

[6] 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.

[7] M. T. M. Silva, M. Gerken, and M. D. Miranda, "An accelerated constant modulus algorithm for space-time blind equalization," in *Proc. EUSIPCO-2004*, Vienna, Austria, Sept. 6–10, 2004, pp. 1853–1856.

[8] 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.

[9] 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.

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

[11] 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.

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

[13] V. Zarzoso and P. Comon, "Optimal step-size constant modulus algorithm," *IEEE Transactions on Communications*, Oct. 2004, submitted.

[14] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. on Information Theory*, vol. 36, no. 2, pp. 312–321, Mar. 1990.

[15] L. Rota and P. Comon, "Blind equalizers based on polynomial criteria," in *Proc. ICASSP-2004*, vol. IV, Montreal, Canada, May 17–21, 2004, pp. 441–444.

[16] V. Zarzoso and P. Comon, "Blind and semi-blind equalization based on the constant power criterion," *IEEE Transactions on Signal Processing*, 2005, to appear in Oct.–Nov. 2005.

[17] ——, "Semi-blind constant modulus equalization with optimal step size," in *ICASSP-2005*, vol. III, Philadelphia, PA, Mar. 18–23, 2005, pp. 577–580.