Averaging Analysis of Local Stability of a Real Constant Modulus Algorithm Adaptive Filter

C. RICHARD JOHNSON, JR., SENIOR MEMBER, IEEE, SOURA DASGUPTA, AND WILLIAM A. SETHARES

Abstract—**The constant modulus algorithm (CMA) for adaptive filtering was developed for usage when the modulus of the desired signal is known, but its specific value at every sample instant is not known. A real arithmetic version of CMA was recently proposed and studied via simulations by Treichler and Larimore. This paper proves local stability of their real arithmetic version of CMA in two applications: i) channel equalization for a transmitted sequence of plus and minus ones, and ii) the separation of a sinusoidal signal from its sum with a number of sinusoidal interferers at separate frequencies. The proofs utilize dynamic system stability theorems from averaging theory, which is a technique currently being exploited in the stability analysis of a variety of adaptive systems.**

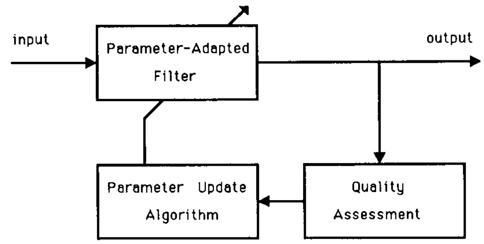


Fig. 1. Generic parameter-adaptive filter structure.

I. Introduction

A generic description of an adaptive filter appears in Fig. 1. The three basic components of this generic description are: i) a filter structure within which the parameters are adjusted, ii) a mechanism that quantifies the quality of the adaptive filter output, and iii) an algorithm that converts this quality assessment into an "improved" filter parameterization. For the traditional application of the ubiquitous LMS adaptive filter, i) is a direct form finite-impulse-response (FIR) structure, ii) is the difference between the adaptive filter output and the desired output, and iii) is based on a gradient descent procedure.

In a variety of adaptive signal processing tasks, a desired signal is not available and an alternative quality assessment mechanism is needed. Channel equalization in the absence of a training signal is one such example, and separation of tones of unknown frequency from their received sum is another. The recently developed constant modulus algorithm (CMA) [1]–[3] is one adaptive algorithm candidate that utilizes a novel quality assessment mechanism to solve certain adaptive filtering problems that lack a desired signal. Although CMA has been proposed, simulated, fabricated, and successfully applied, formal description of its behavioral properties is far from complete. As with all convergence and stability theory for adaptive filters, succinct and generic proofs of desirable behavior validate the applicability of a candidate algorithm and provide various guidelines for enhancing its exploitation. This paper considers particular instances of the channel equalization and tone separation problems cited above, to which a real arithmetic version of CMA [2] is applicable. Using averaging theory, the desired solutions to idealized versions of these problems are proven to be locally exponentially stable. This establishes that, once real CMA drives the adaptive filter parameterization into the vicinity of the desired filter parameterization, the adapted parameterization will remain close to this desired parameterization despite modest, practically unavoidable nonidealities. This paper provides the first proof of such stability and robustness properties of the real CMA of [2], which validates their observation in simulation and practice.

The assumption used in developing the constant modulus algorithm (CMA) is that the adaptive filter output has a constant modulus when the filter structure is appropriately parameterized. A positive function of the difference between some measure of the modulus of the adaptive filter output and the corresponding measure of the desired modulus offers one possibility for quantifying this quality assessment. For example, with $\hat{y}(k)$ the output of the adaptive filter, consider the performance function

$$J(k) = \tfrac{1}{4}(\hat{y}^2(k) - 1)^2, \tag{1.1}$$

which effectively assumes that the desired value of $|\hat{y}|$ is unity. A gradient descent based FIR filter parameterization adjustment can then be developed, as

$$\hat{\theta}(k+1) = \hat{\theta}(k) - \mu \frac{\partial J(k)}{\partial \hat{\theta}(k)}, \tag{1.2}$$

Manuscript received February 14, 1987; revised December 10, 1987. The work of Johnson and Sethares was supported by the National Science Foundation under Grant MIP-8608787. The work of Dasgupta was supported by the National Science Foundation under Grant ECS-8618240.

C. R. Johnson, Jr., is with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853.

S. Dasgupta is with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242.

W. A. Sethares was with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853. He is now with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706.

IEEE Log Number 8820767.

where μ is small and positive,

$$\hat{y}(k) = X^T(k)\,\hat{\theta}(k), \tag{1.3}$$

$$X(k) = \left[x(k) \;\; x(k-1) \;\; \cdots \;\; x(k-N)\right]^T, \tag{1.4}$$

$$\hat{\theta}(k) = \left[\hat{b}_0(k) \;\; \hat{b}_1(k) \;\; \cdots \;\; \hat{b}_N(k)\right]^T, \tag{1.5}$$

and

$$\frac{\partial J(k)}{\partial \hat{\theta}(k)} = (\hat{y}^2(k) - 1)\hat{y}(k)\,X(k). \tag{1.6}$$

The only conceptual distinction between the development in (1.1)–(1.6) and that for traditional LMS, in terms of the description in Fig. 1, is in the quality assessment mechanism, i.e., (1.1). However, this change leads to a quite different algorithm

$$\hat{\theta}(k+1) = \hat{\theta}(k) - \mu\hat{y}(k)(\hat{y}^2(k) - 1)X(k), \tag{1.7}$$

which is obtained by substituting (1.6) into (1.2). Recall that for LMS, the update term is $+\mu[y(k) - \hat{y}(k)]X(k)$, where $y(k)$ is the desired output, based on gradient descent of $(1/2)[y(k) - \hat{y}(k)]^2$ instead of the J in (1.1). Note that the desired signal $y(k)$ is absent, as intended, from (1.7).

The objective of (1.1) can be interpreted as an attempt to restore the unity modulus property to $\hat{y}(k)$ with an appropriate $\hat{\theta}$. This property restoral concept for adaptive filter quality assessment was introduced in [1] for the correction of FM signal corruption. In [1], the signal available for processing was assumed to be in digital quadrature form, which is a complex signal requiring a complex arithmetic filter. Since quadrature FM should have a constant amplitude, the magnitude of the received complex signal should be a constant, but due to channel distortion is not. The CMA adaptive filter of [1] seeks a parameterization that restores this constant modulus property and, in so doing, also equalizes the phase distortion suffered by the transmitted FM signal. In [2], this complex signal/arithmetic CMA adaptive filter was converted to the real signal/arithmetic version of (1.7). In [3], the CMA adaptive filter was applied to the problem of separating two sinusoids where one represented the signal of interest and the other an interferer. In this task, the constant modulus quality assessment is satisfied with the passage (and appropriate scaling) of either complex sinusoid and the suppression of the other. Thus, the function on which gradient descent is performed for adaptive filter parameter adjustment is multimodal. The principal objective of [3] was the analysis of the regions of attraction in parameter space to the lock (i.e., desired signal passage and interferer suppression) and capture (i.e., interferer passage and desired tone suppression) conditions.

The pioneering work of [1]–[3] does not formally prove the convergence/stability of the CMA adaptive filter. However, [2] and [3] provide a variety of insights useful in such a task. For example, in [2], the convergent parameterization of an FIR filter adapted by real CMA for a particular frequency sinusoidal input is observed via simulations to be a gain of $2/\sqrt{3}$ at that frequency. This observation was predicted analytically in [2] by showing (under mild assumptions) that the average of the update term of (1.6) used in (1.7) is zero for this gain, i.e., with $N = 1$, $X(k) = A \sin(\omega k)$ for any A, and $\hat{y}(k) = (2/\sqrt{3}) \sin(\omega k + \phi)$, $(1/2\pi)\int_0^{2\pi} \hat{y}(k)(\hat{y}^2(k) - 1) X(k)\, dk = 0$ is satisfied. Note, in this case, that $|\hat{y}(k)|$ is not intended to have a constant modulus at all k. Rather, zeroing the *average* of the update term in (1.6), which corresponds to minimizing the *average* of (1.1), provides an acceptable $\hat{y}$. Appropriately, the analysis of dynamic behavior in [3] is based upon the expectation of the update equation. This analysis is facilitated with assumptions of independence that are standard approximations in adaptive systems analysis given a small stepsize.

The convergence of a different version of real CMA was analyzed in [4]. The setting in [4] was the equalization of channel-induced distortion of a fixed-frequency sinusoid with a time-varying phase. The convergence analysis in [4] relies on verifying that i) the stationary points of the algorithm cause the performance measure to be zero *at each sample instant*, ii) the gradient of the performance measure evaluated at the stationary point in the adaptive filter parameter space is also zero *at each sample instant*, and iii) the second derivative of the performance measure with respect to the adaptive filter parameters evaluated at the stationary point is positive *at each sample instant*. The version of real CMA analyzed in [4] uses an averaged modulus error, for which i)–iii) are satisfied. Use of this averaged modulus error, rather than its instantaneous version as in (1.1), for gradient descent update development results in the use of the precise period of the transmitted sinusoid in the real CMA of [4]. The instantaneous modulus error $\hat{y}^2(k) - 1$ used in the real CMA in [2] and (1.7) (cited as case 3 and labeled "simplified real" CMA in [4]) is not limited to applications in which the period, and thus frequency, of the transmitted sinusoid to be recovered is known (or accurately estimable) *a priori*. In any event, the version of real CMA in (1.7) from [2] is excluded from the analysis of [4] for its equalization problem setting due to dissatisfaction of i) for the stationary point candidate of the inverse channel model, i.e., the performance measure is not zero for this stationary point candidate *at every sample instant*. However, the real CMA of [2] can exhibit stable average convergence. Recall the simple sinusoidal input result from [2] cited above for the real CMA of (1.7), where $\hat{y}(k) = (2/\sqrt{3}) \sin(\omega k + \phi)$ is the achieved solution. From this case, it is apparent that i) is not satisfied *at every sample instant*. In fact, ii) and iii) are not satisfied at every sample in this case either. This excludes (1.7) from analysis by the approach taken in [4] when the achieved/desired output is sinusoidal, as it will be in our subsequent tone separation task.

The claim is made in [4] that (1.7) can be studied via appropriate extensions of the procedure in [4]. Suggested for consideration is the stability analysis technique in [5] based on evaluation of the ordinary differential equation associated with the expectation of the parameter update equation with vanishingly small stepsize. Although the averaging theory [6] approach taken in the present paper does not rely on a vanishing stepsize nor on a stochastic description of the constituent signals, as does the most popular form of the machinations in [5], the basic concept underlying the averaging analysis in [6] is the same as that used by the ODE analysis in [5]. This concept recognizes that the requirements of [4] summarized above in i)–iii) need not be true at every sample instant, but only on *average*. Translating ii) and iii) to "averaged" conditions yields

$$\operatorname{avg}\left[\frac{\partial J(k)}{\partial \hat{\theta}(k)}\right]\Bigg|_{\hat{\theta}(k)=\theta^*} = 0 \qquad (1.8)$$

$$\operatorname{avg}\left[\frac{\partial^2 J(k)}{\partial \hat{\theta}^2(k)}\right]\Bigg|_{\hat{\theta}(k)=\theta^*} > 0, \qquad (1.9)$$

where "avg" indicates a standard sample average operation and θ^* indicates an "average" stationary point. For example, with periodic inputs, the average could be taken over one signal period. The intuitive appeal of (1.8) and (1.9) is that they are *averaged* versions of the typical test for a local minimum of an instantaneous cost function J, such as in (1.1). Since the averaging and partial differentiation operations can be interchanged, satisfaction of (1.8) and (1.9) can be viewed as tests for the local minimum of an averaged cost function, such as the average of the value of J in (1.1) over one period of $\hat{y}$. For an "instantaneous gradient descent" based scheme such as (1.2), (1.8) and (1.9) indicate local stability of the average behavior of (1.2). Formally extrapolating locally stable behavior of the "averaged algorithm" to verification of locally stable behavior of the actual algorithm is, although nontrivial, one byproduct of averaging theory [6], [7].

An alternative interpretation of (1.8) and (1.9) stems from rewriting (1.2) as $\partial J(k)/\partial\hat{\theta}(k) = -(1/\mu)[\hat{\theta}(k+1) - \hat{\theta}(k)]$. Thus, (1.8) can be rewritten as

$$\operatorname{avg}\left[\hat{y}(k)(\hat{y}^2(k) - 1)X(k)\right]\Big|_{\hat{\theta}(k)=\theta^*} = \operatorname{avg}\left[X(k)\, X^T(k)\, \theta^*\left(\left[X^T(k)\theta^*\right]^2 - 1\right)\right] = 0, \qquad (1.10)$$

which is an implicit definition of an average stationary point θ^*, about which $\{\hat{\theta}(k)\}$ may wiggle. Essentially (1.10) could be derived via the observation that (1.7) ceases updating on *average* when (1.10) is true. The update term of (1.7) restores $\hat{\theta}$ back toward θ^*, on *average*, if the change in the algorithm correction term $-\mu\hat{y}(\hat{y}^2 - 1)X$ due to a change in $\hat{\theta}$ away from θ^* is negative, i.e., if

$$\operatorname{avg}\left[\frac{\partial\left\{-\hat{y}(k)(\hat{y}^2(k) - 1)\, X(k)\right\}}{\partial\hat{\theta}(k)}\right]\Bigg|_{\hat{\theta}(k)=\theta^*} < 0. \qquad (1.11)$$

Note that the left side of (1.11) also corresponds to the average of $-\partial^2 J(k)/\partial\hat{\theta}^2(k)$, which connects (1.11) to (1.9). Given (1.3), (1.11) can be evaluated as

$$\operatorname{avg}\left[(3\hat{y}^2(k) - 1)X(k)\, X^T(k)\right]\Big|_{\hat{\theta}(k)=\theta^*} > 0. \qquad (1.12)$$

Our objective in this paper is to present a formal justification of the crucial role of (1.10) and (1.12) for proof of the local stability of real CMA in (1.7) about θ^*. Succinctly, if (1.12) is fulfilled, then the real CMA algorithm is locally stable about the stationary point of (1.10).

These insights from [2]–[4] suggest the appropriateness, induced by the use of a small but nonvanishing stepsize, of averaging theory in analyzing the convergence/stability of the disarmingly simple version of real CMA in (1.7) introduced in [2]. In fact, a secondary objective of this paper is to proselytize the use of (appropriate) averaging theory for analysis of CMA (and other "new" adaptive filtering schemes associated with smooth, differentiable cost functions). Insights from average behavior analysis of adaptive systems are well documented in [5] and [6], albeit for adaptive parameter estimation schemes using a desired signal in a more familiar prediction error quality assessment, and therefore seemingly quite different from real CMA.

To illustrate the usefulness of averaging analysis, in general, and the importance of (1.12), in particular, in assessing the local stability properties of real CMA in (1.7), two applications are considered in the next two sections. In Section II, (1.7) is applied for channel equalization for a transmitted sequence of plus and minus ones, which is an apt abstraction of baud synchronous binary phase shift keying. In Section III, (1.7) is applied to achieve separation of one sinusoid from a sum of sinusoids, which is similar to the two sinusoid complex CMA tone capture problem studied in [3]. In the equalization problem of Section II, when the adaptive filter is appropriately parameterized, the cost function in (1.1) *is* minimized at every sample instant, and not just on average, as in the tone separation problem of Section III. Thus, the need for averaging in analyzing adaptive system behavior is diminished in the equalization problem. Still we begin with this equalization problem as a more transparent means than the tone separation problem of introducing the application of an appropriate averaging theorem to the real CMA of [2].

In both of the following "application" sections, we first present the (idealized) problem formulation and note the applicability of an appropriately dimensioned real CMA adaptive filter solution. This problem/solution description is followed by the statement of an applicable averaging theorem. In Section II, we are able to extract such a theorem rather directly from [7]; while in Section III, we need a "new" result. In this latter case, the proof of the theorem is provided in the Appendix. The resulting locally stable behavior of each real CMA application is stated as a corollary to the appropriate theorem. The proofs of these corollaries are included in the body of the paper since it is this formal translation of applicable averaging theorems to real CMA behavior description that is the heart of our paper. Remember throughout the reading of this paper that practically desirable robustness (e.g., to noise and mismodeling) is implied by such proof of local stability based on averaging theory.

II. Real CMA for Channel Equalization for Baud Synchronous Binary Phase Shift Keying

Consider the channel equalization problem diagrammed in Fig. 2. The transmitted source y is a periodic sequence of plus and minus ones. The received signal x is y passed through a channel with stable transfer function $1/\{\sum_{i=0}^{N} b_i z^{-i}\}$, so that

$$y(k) = \sum_{i=0}^{N} b_i x(k-i) = X^T(k)\bar{\theta}, \tag{2.1}$$

where X is defined in (1.4) and $\bar{\theta} = [b_0, b_1, b_2, \cdots, b_N]$. The hope is that the parameters of the adaptive filter in (1.3)–(1.5) are adjusted such that $\hat{b}_i \to b_i$ and thus $\hat{y} \to y$. Since the modulus of y is one, i.e., $|y(k)| = 1 \ \forall k$, the cost function of (1.1) is appropriate. Thus, real CMA of (1.7) is an attractive candidate for updating the adaptive filter parameters given the absence of the desired signal y at the receiver.

Observe that there are three stationary point candidates, i.e., $\bar{\theta}$, 0, and $-\bar{\theta}$, which satisfy (1.10). Of these, $\theta^* = 0$ cannot possibly satisfy (1.12), and is thus locally unstable. For the other two, however, $\hat{y}^2(k)|_{\hat{\theta}=\bar{\theta}} = \hat{y}^2(k)|_{\hat{\theta}=-\bar{\theta}} = [X^T(k)\theta^*]^2 = 1$, for all k. After substituting $\hat{y}^2 = 1$ into (1.12), one finds that both of these candidates are locally stable provided the traditional persistent excitation condition associated with the LMS algorithm, i.e.,

$$\infty > \beta I > \text{avg}\left[X(k)\,X^T(k)\right] > \alpha I > 0, \tag{2.2}$$

is satisfied.

The following averaging theorem provides sufficient conditions for local stability about $\bar{\theta}$ of real CMA in this equalization problem. Its proof, being a modest modification of a result in [7], is omitted.

Theorem 2.1: Consider the difference equation

$$\theta(k+1) = \theta(k) - \mu f(\theta(k), k) \tag{2.3}$$

where $\theta(\cdot) \in R^{n+1}$ for all θ and for all k, $f(\theta, k+T) = f(\theta, k)$, and

$$f_{av}(\theta) \equiv \frac{1}{T}\sum_{k=S+1}^{S+T} f(\theta, k) \qquad \forall S \in Z_+. \tag{2.4}$$

Assume that f and f_{av} satisfy the following assumptions:

i) $\theta = 0$ is a stationary point of f

$$f(0, k) = 0 \qquad \forall k, \tag{2.5}$$

ii) f is Lipschitz continuous in its first argument

$$\left\| f(\theta_1, k) - f(\theta_2, k) \right\| \le M_1 \left\| \theta_1 - \theta_2 \right\|, \text{ and} \tag{2.6}$$

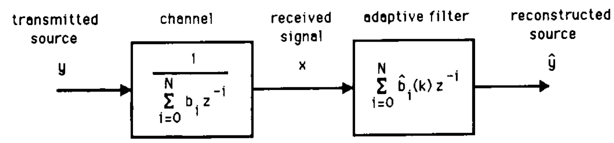


Fig. 2. Channel equalization.

iii) for some $\bar{\mu}$ and all $0 < \mu < \bar{\mu}$, the origin $\theta = 0$ is a locally exponentially stable equilibrium of the averaged equation

$$\theta_{av}(k+1) = \theta_{av}(k) - \mu f_{av}(\theta_{av}(k)). \tag{2.7}$$

Then there exists $\mu^* > 0$ such that for all $0 < \mu < \mu^*$, the origin $\theta = 0$ is also a locally exponentially stable equilibrium of (2.3). ▽▽▽

Remark:

2.a) For a system like (2.3), local exponential stability means there exist $0 \le \lambda_1 < 1$, $\infty > \lambda_i > 0$ for $i = 2$ and 3, such that $\|\theta(k)\| \le \lambda_2 \|\theta_0\| \lambda_1^k$ for all k and for all $\|\theta_0\| < \lambda_3$.

We now apply the above theorem to prove the local stability of the baud synchronous binary phase shift keying application of real CMA described at the beginning of this section.

Corollary 2.1: Consider the real CMA of (1.7) with $\hat{y}(k)$, $X(k)$, and $\hat{\theta}(k)$ as defined in (1.3), (1.4), and (1.5), respectively. Suppose, as suggested by (2.1), that the received signal $x(k)$ is an autoregressively filtered version of the transmitted source $y(k)$ such that

$$x(k) = \frac{1}{\sum_{i=0}^{N} b_i q^{-i}}\, y(k), \tag{2.8}$$

where q^{-1} is the delay operator (i.e., $q^{-1}x(k) = x(k-1)$) and $1/(\sum_{i=0}^{N} b_i q^{-1})$ is asymptotically stable. If $y(k)$ is a periodic sequence of 1's and -1's with period T, the nonperiodic component of $x(k)$ due to initial condition effects is exponentially decaying, since (2.8) is asymptotically stable. Denote the steady-state, periodic component of x as x_p, with $X_p(k)$ defined as the periodic vector of present and past x_p just as X is defined in (1.4) as a vector of present and past x. Now suppose that X_p fulfills the persistent excitation condition of LMS in (2.2), i.e., there exists $\alpha > 0$ such that for every S

$$\sum_{k=S+1}^{S+T} X_p(k)\, X_p^T(k) > \alpha I > 0. \tag{2.9}$$

Here T is the period of $x_p(k)$ and thus $X_p(k)$. Then there exists $\mu^* > 0$, such that for any $\mu < \mu^*$,

$$\theta^* = \bar{\theta} = [b_0, b_1, \cdots, b_N]^T \tag{2.10}$$

is a locally exponentially stable equilibrium of (1.7).

Proof: Consider the equation

$$\theta_p(k+1) = \theta_p(k) - \mu(\hat{y}_p^2(k) - 1)\,\hat{y}_p(k)\,X_p(k) \tag{2.11}$$

where $\theta_p(k) = \hat{\theta}_p(k) - \theta^*$ and $\hat{y}_p(k) = X_p^T(k)\,\hat{\theta}_p(k)$.

Then with $\theta_p(k) = 0$, $\hat{y}_p^2(k) = y^2(k) = 1$. With $f(k, \theta(k))$ of (2.3) identified with the update term in (2.11), (2.5) is satisfied. The stability of (2.8) ensures the satisfaction of (2.6). The averaged equation for (2.11) corresponding to (2.7) is locally exponentially stable around $\theta_p = 0$, if for some $\alpha_1 > 0$ [9]

$$\left.\frac{\partial f_{av}(\theta_p(k))}{\partial \theta_p(k)}\right|_{\theta(k)=0} > \alpha_1 I. \tag{2.12}$$

From (2.11)

$$\left.\frac{\partial f_{av}(\theta_p(k))}{\partial \theta_p(k)}\right|_{\theta(k)=0} = \frac{1}{T}\sum_{k=S+1}^{S+T} \left(3\hat{y}_p^2(k) - 1\right)X_p(k)\, X_p^T(k)\big|_{\theta_p=0} = \frac{2}{T}\sum_{k=S+1}^{S+T} X_p(k)\, X_p^T(k) > \alpha_1 I, \tag{2.13}$$

which holds due to (2.9) with $\alpha_1 = 2\alpha$. Thus, $\exists \mu^*$, such that for any $\mu < \mu^*$, (2.11) is locally exponentially stable around the origin. Since $X_p(k)$ is bounded and approaches $X(k)$ exponentially, this implies the result. ▽▽▽

More Remarks:

2.b) The satisfaction of (2.9) is guaranteed if, with $Y(k) \equiv [y(k), y(k-1), \cdots, y(k-N)]^T$, the matrix $\Sigma_{k=0}^{T-1} Y(k)\, Y^T(k)$ is positive definite [10].

2.c) The stationary point candidate $-\bar{\theta}$ is also locally exponentially stable. The proof of this fact is essentially identical to that for Corollary 2.1.

2.d) For a discussion of the relaxation of the periodicity assumption, see remark 3.e).

III. Real CMA for Tone Capture and Multiple Tone Interference Suppression

Consider the received signal x that is a sum of a sinusoidal source y and a number of sinusoidal interferers

$$x(k) = y(k) + \sum_{i=1}^{N/2} A_i \sin(\omega_i k + \phi_i), \tag{3.1}$$

where

$$y(k) = A \sin(\bar{\omega} k), \tag{3.2}$$

$\bar{\omega} \neq \omega_i\ \forall i$, $\omega_i \neq \omega_j\ \forall i \neq j$, and N is even. With the interferer frequencies labeled such that $0 < \omega_1 < \omega_2 < \cdots < \omega_{N/2} < \pi$, the desired zero locations of the tone separator are illustrated in Fig. 3. With this constellation for the zeros of $\beta(z) = 1 + \Sigma_{i=1}^{N}\{b_i/b_0\}z^{-i}$, all of the interferers are cancelled. With b_0 selected such that $|b_0\beta(e^{j\bar{\omega}})| = 2/(A\sqrt{3})$, (1.10) is satisfied. Thus,

$$\bar{\theta} = [b_0, b_1, \cdots, b_N]^T \tag{3.3}$$

is a stationary point that may be locally stable. For this stationary point candidate $\theta^* = \bar{\theta}$,

$$\hat{y}(k)|_{\hat{\theta}=\theta^*=\bar{\theta}} \to y^*(k) = (2/\sqrt{3}) \sin(\bar{\omega} k + \phi), \tag{3.4}$$

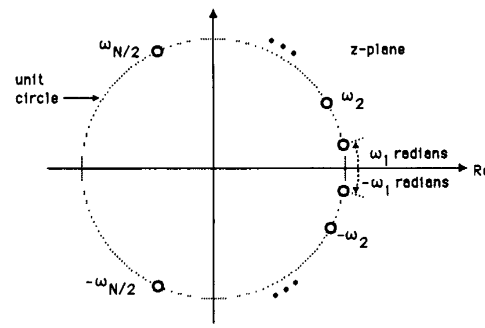


Fig. 3. Tone separator desired zero constellation.

with

$$\phi = \arg\left(b_0\beta(e^{j\bar{\omega}})\right). \tag{3.5}$$

Since $\hat{y}(k)|_{\hat{\theta}=\bar{\theta}}$ converges to a sinusoid of frequency $\bar{\omega}$ and $y(k) = A\sin(\bar{\omega}k)$ is a component of x in X in (1.4), the verification of (1.12) is nontrivial.

To provide sufficient conditions for the local stability about $\theta^* = \bar{\theta}$ in (3.3), we utilize the following averaging theorem. Its proof appears in the Appendix.

Theorem 3.1: Consider the difference equation

$$\theta(k+1) = \theta(k) - \mu f(\theta(k), X(k)) \tag{3.6}$$

with $\mu \in R_+$ and $\theta(k), X(k), f(\cdot, \cdot) \in R^{n+1}$. Suppose there exists $T > 0$ such that for all k

$$X(k+T) = X(k). \tag{3.7}$$

Define the average of the update kernel as

$$f_{av}(\theta) = \frac{1}{T}\sum_{i=0}^{T-1} f(\theta, X(k+i)) \quad \forall k > 0. \tag{3.8}$$

Suppose that

i) $\theta = 0$ is a stationary point of f_{av}

$$f_{av}(0) = 0, \tag{3.9}$$

ii) the origin $\theta = 0$ is a locally exponentially stable equilibrium of the averaged version of (3.6), i.e., there exists $\alpha > 0$ such that

$$\left.\frac{\partial f_{av}(\theta)}{\partial \theta}\right|_{\theta=0} > \alpha I, \text{ and} \tag{3.10}$$

iii) f is locally Lipschitz continuous in its first argument about $\theta = 0$, i.e., there exist ϵ_1 and M_1 such that $\forall \|\theta_1\|, \|\theta_2\| \leq \epsilon_1$

$$\left\| f(\theta_1, X(k) - f(\theta_2, X(k))\right\| \leq M_1 \|\theta_1 - \theta_2\|. \tag{3.11}$$

Then $\exists K, \mu^*, \epsilon^* > 0$, satisfying

$$K\mu^* < \epsilon^* < \min(\epsilon_1, R) \tag{3.12}$$

such that $\forall \|\theta(0)\| < \epsilon^*$ and $\forall \mu < \mu^*$

$$\lim_{i\to\infty} \|\theta(i)|^2 \to [0, K\mu]. \tag{3.13}$$

▽▽▽

Remarks:

3.a) Equation (3.13) is a notationally convenient way of stating that provided $\|\theta(0)\|$ is "small enough," for large enough i, $\|\theta(i)\|$ becomes smaller than $K\mu$. It does not mean that the limit itself exists.

3.b) The theorem differs from the available averaging theorems, e.g., in [6] and [7], in that it does not require that

$$f(0, X(k)) = 0 \qquad \forall k. \tag{3.14}$$

Rather, only the average of $f(0, X(k))$ needs to equal zero. Compare (3.9) to (2.5) of Theorem 2.1. In terms of (1.1), (1.6), and (1.7), (3.9) requires that $\theta = 0$ be a minimum of the average of the cost function.

In the remainder of this section, we demonstrate the application of Theorem 3.1 to the local stability of the real CMA of (1.7) with $x(k)$ as in (3.1). The result stated below as Corollary 3.1 is much as is expected from the introduction, i.e., (1.12) plays a critical role. However, the difference in requirements i)–iii) for Theorem 3.1 versus those of Theorem 2.1 leads to a more tortuous path to the verification of the importance of (1.12). Also, the increase in the number of local minima in this sinusoidal separation problem versus the two in the previous equalization problem for sequences of plus and minus ones causes added difficulty. In the previous equalization problem, only the true answer and its negative were locally stable points. In this sinusoidal separation problem, if the adaptive filter selects (with the appropriate gain) any sinusoid in the received sum and suppresses all the others, (1.10) is equally well satisfied. These local stability candidates can be arbitrarily close together depending on the various frequencies ω_i in the received sum. Thus, it is only reasonable to require that they be sufficiently well separated for them to be viable local stability candidates. Such an assumption will be made in the translation of Theorem 3.1 to Corollary 3.1. Another artifact of this sum of sinusoids case is that the averages in (1.10) and (1.12) will lead to integrals of cross-products and self-products of sinusoids. For very specific frequency combinations, (1.10) could be satisfied without the situation in Fig. 3 emerging. Similarly, terms that would normally be zero in the left side of (1.12) could be nonzero for specific frequency combinations. Assumptions will be made to avoid such degenerate situations.

The application of Theorem 3.1 to the tone separation problem of this section results in the following statement of local stability.

Corollary 3.1: Consider the real CMA of (1.7) with $\hat{y}(k)$ and $X(k)$ defined by (1.3)–(1.5) and (3.1)–(3.2). Suppose the following assumptions hold.

I) The frequencies ω_i and $\bar{\omega}$ are rationally related and distinct, i.e.,

$$\omega_i = \frac{2\pi N_i}{T} \tag{3.15}$$

and

$$\bar{\omega} = \frac{2\pi \bar{N}}{T} \tag{3.16}$$

for some positive integers N_i, $\bar{N}$, and T with

$$\left.\begin{aligned} &N_i \neq N_j \qquad \forall i \neq j, \\ &\bar{N} \neq N_i \qquad \forall i \\ \text{and} \\ &N_i, \bar{N} < 2T \end{aligned}\right\}, \tag{3.17}$$

which results in periodic $x(k)$ in (3.1) with period T.

II) The frequencies $\bar{\omega}$ and ω_i are such that $\forall i$, and any integer λ

$$\omega_i \neq 2\pi\lambda \pm 3\bar{\omega} \tag{3.18}$$

and

$$\bar{\omega} \neq \pi/2, \tag{3.19}$$

which allows satisfaction of (1.10) to achieve the tone separator zero constellation in Fig. 3.

III) The frequencies ω_i and $\bar{\omega}$ are such that $\forall i, j$, and any integer λ

$$|\omega_i \pm \omega_j| \neq 2\pi\lambda \pm 2\bar{\omega}, \tag{3.20}$$

which allows satisfaction of (1.12) given the persistent excitation condition associated with LMS, as in (2.2), and achievement of the tone separator pattern in Fig. 3.

IV) The phase shift ϕ for the desired convergent $\hat{y}$ in (3.3) satisfies

$$\phi = m\pi + \delta \tag{3.21}$$

for some integer m with $|\delta| << \pi/2$, which implies that the interferers are "sufficiently" separated in frequency from the desired signal.

Then with $\theta^* = \bar{\theta}$ in (3.3) and the corresponding y^* from (3.4), there exist δ^*, μ^*, $\epsilon^* > 0$ such that $\forall \mu < \mu^*$, $|\delta| < \delta^*$, and $\|\hat{\theta}(0) - \theta^*\| < \epsilon^*$, $\lim_{k\to\infty} \|\hat{y}(k) - y^*(k)\| \to [0, O(\mu)]$.

Proof: With $\theta(k) \equiv \hat{\theta}(k) - \theta^*$, the real CMA of (1.7) can be written as

$$\theta(k+1) = \theta(k) - \mu\hat{y}(k)(\hat{y}^2(k) - 1)X(k). \tag{3.22}$$

The candidate for local stability is $\theta^* = \bar{\theta}$, which is the origin of (3.22). From (3.4) and (3.16),

$$y^*(k) = X^T(k)\theta^* = \frac{2}{\sqrt{3}}\sin\left(\frac{2\pi\bar{N}}{T}k + \phi\right) \tag{3.23}$$

for some ϕ between $-\pi$ and π. Thus, given (3.3) and the definition

$$B^*(z^{-1}) \equiv b_0^* + \sum_{i=1}^{N} b_i^* z^{-i}, \tag{3.24}$$

it follows that

$$B^*(z^{-1}) = a\prod_{i=1}^{N/2}(1 - e^{-j\omega_i}z^{-1})(1 - e^{j\omega_i}z^{-1}) \tag{3.25}$$

with

$$|B^*(e^{-j\omega^*})| = \frac{2}{A\sqrt{3}}. \tag{3.26}$$

In view of the dimension of θ^*, ϕ is thus fixed at

$$\phi = \arg\left\{B^*(e^{-j\omega^*})\right\}, \tag{3.27}$$

which is similar to (3.5). Our objective then is to show that for small stepsizes, with $\|\theta(0)\|$ sufficiently small, $\|\theta(k)\|$ for sufficiently large k will also be small. Due to the boundedness of $X(k)$, this in turn will imply that $\hat{y}(k)$ will be "close" to $y^*(k)$ in (3.23).

It is easily verified that by identifying $f(\theta(k), X(k))$ of (3.6) with $\hat{y}(k)(\hat{y}^2(k) - 1)X(k)$ of (3.22), condition (3.11) of Theorem 3.1 is satisfied for bounded X. Satisfaction of (3.7) requires $x(k)$ to be periodic with period T. Assumption I) assures such periodicity of $x(k)$. Thus, for local stability of $\theta(k)$ around the origin, conditions (3.9) and (3.10) will have to hold or, as stated in the Introduction, the equivalent conditions (1.10) and (1.12) must be satisfied. Their satisfaction is respectively demonstrated in Lemmas 3.1 and 3.2. We realize that it is unconventional to insert Lemmas and their proofs within the proof of a corollary. However, we have chosen this format to preserve the flow of the conceptual logic used in establishing the corollary. Recall that a major objective of this paper is to detail the translation of averaging theorems to corollaries describing real CMA behavior.

We now state and prove Lemma 3.1.

Lemma 3.1: Under Assumptions I) and II),

$$\frac{1}{T}\sum_{i=0}^{T-1} y^*(k+i)\left\{y^{*2}(k+i) - 1\right\} \cdot X(k+i) = 0, \quad \forall k. \tag{3.28}$$

Proof: By the periodicity of y^* and X are given (3.16) and (3.23),

$$\begin{aligned}\frac{1}{T}\sum_{i=0}^{T-1} & y^*(k+i)\left\{y^{*2}(k+i) - 1\right\} X(k+i)\\ &= \frac{1}{T}\sum_{k=0}^{T-1} y^*(k)\left\{y^{*2}(k) - 1\right\} X(k)\\ &= \frac{1}{T}\sum_{k=0}^{T-1} \frac{2}{\sqrt{3}}\left(\frac{4}{3}\sin^3(\bar{\omega}k + \phi) - \sin(\bar{\omega}k + \phi)\right)X(k)\\ &= \frac{2}{3\sqrt{3}}\frac{1}{T}\sum_{k=0}^{T-1} \sin(3\bar{\omega}k + 3\phi)X(k).\end{aligned} \tag{3.29}$$

The $(i+1)$th element of (3.29) given (1.4), (3.1), and (3.2) is

$$\frac{2}{3\sqrt{3}}\frac{1}{T}\sum_{k=0}^{T-1} \sin(3\bar{\omega}k + 3\phi)\left\{A\sin\left(\bar{\omega}(k-i)\right) + \sum_{\lambda=0}^{N/2} A_\lambda \sin\left(\omega_\lambda(k-i) + \phi_\lambda\right)\right\}. \tag{3.30}$$

We will now consider each term in the expansion of (3.30) separately. First,

$$\begin{aligned}\frac{1}{T}\sum_{k=0}^{T-1} & A_\lambda \sin(3\bar{\omega}k + 3\phi)\sin(\omega_\lambda k - \omega_\lambda i + \phi_\lambda)\\ &= \frac{A_\lambda}{2T}\sum_{k=0}^{T-1}\left\{\cos\left[(3\bar{\omega} - \omega_\lambda)k + 3\phi + \omega_\lambda i - \phi_\lambda\right] - \cos\left[(3\bar{\omega} + \omega_\lambda)k + 3\phi - \omega_\lambda i + \phi_\lambda\right]\right\}\\ &= 0,\end{aligned} \tag{3.31}$$

the last step following due to (3.15), (3.16), and (3.18). Likewise,

$$\begin{aligned}\frac{1}{T}\sum_{k=0}^{T-1} & \sin(3\bar{\omega}k + 3\phi)\sin\bar{\omega}(k-i)\\ &= \frac{1}{2T}\sum_{k=0}^{T-1}\left\{\cos(2\bar{\omega}k + 3\phi + \bar{\omega}i) - \cos(4\bar{\omega}k + 3\phi - \bar{\omega}i)\right\} = 0,\end{aligned}$$

as (3.16) holds and $2\bar{\omega} \neq 2\pi$ given (3.17) and $4\bar{\omega} \neq 2\pi$ given (3.19). Thus, (3.29) equals zero, which establishes (3.28). ▽

Having established conditions for the satisfaction of (1.10), we now turn to (1.12). Specifically, we need to show that there exists $\alpha > 0$, such that

$$\frac{1}{T}\sum_{k=0}^{T-1}(3y^{*2}(k) - 1)X(k)X^T(k) > \alpha I. \tag{3.32}$$

This condition is essentially a richness condition and sets a precise value for the dimension of θ. Notice that a necessary condition for its satisfaction is that

$$\frac{1}{T}\sum_{k=0}^{T-1} X(k)X^T(k) > 0, \tag{3.33}$$

as in (2.2). Due to the form of $X(k)$ as a sum of $(N/2) + 1$ sinusoids, it is well known, e.g., [10], that $\dim(\theta^*) \leq N + 2$. Moreover, the need for suppressing $N/2$ sinusoids, and having a gain of $2/(A\sqrt{3})$ at $\bar{\omega}$, requires $\dim(\theta^*) \geq N + 1$. Thus, it would seem that $\dim(\theta^*)$ could be $N + 1$ or $N + 2$. However, it must exactly equal $N + 1$. This is because (3.32) requires that θ^* be an isolated singularity of $\partial J(\hat{\theta})/\partial\hat{\theta}$. Fixing the dimension of θ^* at $N + 2$ would prevent this isolation from occurring. This can be heuristically justified by noting in this tone separation problem that the zero pattern of Fig. 3 plus a particular gain at $\bar{\omega}$ can be achieved with $\dim(\theta^*) = N + 1$ in establishing B^* of (3.23). There would be a continuum of acceptable solutions for $\dim(\theta^*) = N + 2$. However, the correct dimension by itself will not guarantee (3.32). Assumption III) is also required. Lemma 3.2, below, verifies (3.32) and adds an appreciable degree of isolation by using Assumption IV).

Lemma 3.2: Under assumptions I)–IV), there exists a $\delta^* > 0$, such that (3.32) holds for all $|\delta| < \delta^*$ in (3.21).

Proof: With $y(k)$ as in (3.2), define

$$Y(k) = \left[y(k), y(k-1), \cdots, y(k-n)\right]^T, \tag{3.34}$$

and

$$\bar{X}(k) = X(k) - Y(k). \tag{3.35}$$

Note that $\bar{X}(k)$ would have been the regressor had $x(k)$ in (3.1) been the linear combination of the interferers alone.

Now, defining

$$F = \frac{1}{T}\sum_{k=0}^{T-1} (3y^{*2}(k) - 1)X(k)\, X^T(k)$$

and using (3.23) yields

$$F = \frac{1}{T}\sum_{k=0}^{T-1} \{4\sin^2(\bar{\omega}k + \phi) - 1\}X(k)\, X^T(k)$$
$$= \frac{1}{T}\sum_{k=0}^{T-1} \{1 - 2\cos(2\bar{\omega}k + 2\delta)\}X(k)\, X^T(k), \tag{3.36}$$

the last equality following from (3.21). Consider decomposing F as $F_1 - F_2$. Thus,

$$F_1 = \frac{1}{T}\sum_{k=0}^{T-1} X(k)\, X^T(k)$$
$$= \frac{1}{T}\sum_{k=0}^{T-1} \{\bar{X}(k) + Y(k)\}\{\bar{X}(k) + Y(k)\}^T.$$

In view of Assumption I),

$$F_1 = \frac{1}{T}\sum_{k=0}^{T-1} \bar{X}(k)\, \bar{X}^T(k) + \frac{1}{T}\sum_{k=0}^{T-1} Y(k)\, Y^T(k). \tag{3.37}$$

Consider now

$$F_2 = \frac{1}{T}\sum_{k=0}^{T-1} 2\cos(2\bar{\omega}k + 2\delta)[\bar{X}(k)\, \bar{X}^T(k) + \bar{X}(k)\, Y^T(k) + Y(k)\, \bar{X}^T(k) + Y(k)\, Y^T(k)]. \tag{3.38}$$

There are three types of contributions to the $(\lambda + 1, m + 1)$ element of F_2. Two of these are linear combinations of the following two forms.

Form I:

$$\frac{1}{T}\sum_{k=0}^{T-1} 2\cos(2\bar{\omega}k + 2\delta)\sin\bar{\omega}(k - \lambda) \cdot \sin\{\omega_i(k - m) + \phi_i\}$$
$$= \frac{1}{T}\sum_{k=0}^{T-1} \cos(2\bar{\omega}k + 2\delta)[-\cos\{(\bar{\omega} + \omega_i)k + \phi_i - (\bar{\omega}\lambda + \omega_i m)\} + \cos\{(\bar{\omega} - \omega_i)k - \phi_i - \bar{\omega}\lambda + \omega_i m\}],$$

which, due to (3.18) and Assumption I), must equal zero. Note that this form covers contributions from $\bar{X}Y^T$ and $Y\bar{X}^T$.

Form II:

$$\frac{1}{T}\sum_{k=0}^{T-1} 2\cos(2\bar{\omega}k + 2\delta)\sin(\omega_i(k - \lambda) + \phi_i) \cdot \sin(\omega_j(k - m) + \phi_j)$$
$$= \frac{1}{T}\sum_{k=0}^{T-1} \cos(2\bar{\omega}k + 2\delta)[\cos\{(\omega_j - \omega_i)k + \phi_j - \phi_i + \omega_i\lambda - \omega_j m\} - \cos\{(\omega_j + \omega_i)k + \phi_j + \phi_i - \omega_i\lambda - \omega_j m\}],$$

which, due to Assumptions I) and III), must equal zero. This form covers contributions from $\bar{X}\bar{X}^T$.

The remaining contribution to F_2 is due to YY^T. It is the only nonzero component. Thus, (3.38) reduces to

$$F_2 = \frac{1}{T}\sum_{k=0}^{T-1} 2\cos(2\bar{\omega}k + 2\delta)Y(k)\, Y^T(k).$$

Thus, given (3.36) and (3.37),

$$F = F_1 - F_2 = \frac{1}{T}\sum_{k=0}^{T-1} \bar{X}(k)\, \bar{X}^T(k) + \frac{1}{T}\sum_{k=0}^{T-1} \{1 - 2\cos(2\bar{\omega}k + 2\delta)\}Y(k)\, Y^T(k)$$
$$= \frac{1}{T}\sum_{k=0}^{T-1} \bar{X}(k)\, \bar{X}^T(k) + G, \tag{3.39}$$

where

$$G = \frac{1}{T}\sum_{k=0}^{T-1} \{1 - 2\cos(2\bar{\omega}k + 2\delta)\}Y(k)\, Y^T(k). \tag{3.40}$$

Now

$$G(\lambda + 1, m + 1) = \frac{1}{T}\sum_{k=0}^{T-1} \{1 - 2\cos(2\bar{\omega}k + 2\delta)\} \cdot \sin(\bar{\omega}(k - \lambda))\sin(\bar{\omega}(k - m))$$
$$= \frac{1}{2T}\sum_{k=0}^{T-1} \left(1 - 2\cos\left(\frac{4\pi\bar{N}k}{T} + 2\delta\right)\right) \cdot \left[\cos\frac{2\pi\bar{N}(\lambda - m)}{T} - \cos\frac{2\pi\bar{N}(2k - \lambda - m)}{T}\right]$$
$$= \frac{1}{2}\cos\frac{2\pi\bar{N}(\lambda - m)}{T} + \frac{1}{2T}\sum_{k=0}^{T-1} 2\cos\left(\frac{4\pi\bar{N}k}{T} + 2\delta\right) \cdot \cos\left(\frac{2\pi\bar{N}}{T}(2k - \lambda - m)\right) \tag{3.41}$$

$$= \frac{1}{2} \cos \frac{2\pi \bar{N}(\lambda - m)}{T}$$

$$+ \frac{1}{2T} \sum_{i=0}^{T-1} \left[\cos \left(\frac{8\pi \bar{N} k}{T} + 2\delta - \frac{2\pi \bar{N}}{T} (\lambda + m) \right) \right.$$

$$\left. + \cos \left(2\delta + \frac{2\pi \bar{N}}{T} (\lambda + m) \right) \right]$$

$$= \frac{1}{2} \left[\cos \frac{2\pi \bar{N}(\lambda - m)}{T} \right.$$

$$\left. + \cos \left(2\delta + \frac{2\pi \bar{N}}{T} (\lambda + m) \right) \right]. \quad (3.42)$$

In the foregoing, (3.41) follows due to (3.17), and (3.42) follows due to (3.19) and (3.17). Consider

$$\frac{1}{2} \left[\cos \left(\frac{2\pi \bar{N}(\lambda - m)}{T} \right) + \cos \left(\frac{2\pi \bar{N}}{T} (\lambda + m) \right) \right]$$

$$= \cos \left(\frac{2\pi \bar{N} \lambda}{T} \right) \cos \left(\frac{2\pi \bar{N} m}{T} \right).$$

Thus,

$$G = \Gamma \Gamma^T + \Delta G, \quad (3.43)$$

where

$$\Gamma = [1, \cos(\bar{\omega}), \cdots, \cos(N\bar{\omega})]^T \quad (3.44)$$

and for sufficiently small δ

$$\|\Delta G\| = O(\delta). \quad (3.45)$$

Thus, from (3.39),

$$F = \frac{1}{T} \sum_{k=0}^{T-1} \bar{X}(k) \bar{X}^T(k) + \Gamma \Gamma^T + \Delta G \quad (3.46)$$

$$= \bar{F} + \Delta G. \quad (3.47)$$

We shall now show that $\bar{F}$ is positive definite. It is clearly positive semidefinite. Thus, if it is not positive definite, $\exists$ nonzero $\xi \in R^{n+1}$, such that

$$\xi^T \bar{F} \xi = 0,$$

which implies that both

$$\xi^T \bar{X}(k) = 0, \quad (3.48)$$

for all k since $\bar{X}(k)$ is periodic, and

$$\xi^T \Gamma = 0. \quad (3.49)$$

By the definitions of $\bar{X}(k)$ and θ^*, (3.48) and (3.24) imply

$$\xi = \gamma \theta^* \quad (3.50)$$

for any scalar $\gamma \neq 0$. Thus, (3.49) implies

$$\theta^{*T} \Gamma = 0,$$

which implies that

$$\text{Re } B^*(e^{-j\omega^*}) = 0,$$

which is satisfied only if

$$|\delta| = \pi/2.$$

Thus, for $|\delta| < \pi/2$, $\bar{F}$ is positive definite. By (3.47) and (3.45), there exists δ^* such that $\forall |\delta| < \delta^*$, F is positive definite. ∇

Lemmas 3.1 and 3.2, together with Theorem 3.1 and the boundedness of $X(k)$, yield the corollary. $\nabla\nabla\nabla$

More Remarks:

3.c) Assumption IV) essentially requires that $\bar{\omega}$ be "sufficiently" separated from $\omega_1, \cdots, \omega_{N/2}$. To see this requirement, note that for the filter to block $\omega_1, \cdots, \omega_{N/2}$, the zeros of $\beta(z)$ must lie at $e^{\pm j\omega_i}$. Thus, if inf $|\omega_i - \bar{\omega}|$ is large, due to (3.5), ϕ must be close to being an integer multiple of π.

3.d) It must be stressed that for $x(k)$ as in (3.1), the cost function $(\hat{y}^2(k) - 1)^2$ has numerous minima. Under appropriate conditions, most of these could be locally stable. For example, for (3.1), each of $\hat{y}(k) = (2/\sqrt{3}) \sin(\omega_i k + \bar{\phi}_i)$ could correspond to locally stable points.

3.e) The algorithm is essentially robust to violations of some of the technical conditions in Corollary 3.1. Thus, even if $f_{av}(0) \neq 0$ for the candidate θ^*, as required for local stability by (3.9) and (1.10), but is small, this point could still be locally stable. We have found in simulations that with $x(k) = \sin(\pi/30)k + \sin(3\pi/30)k$, $\hat{y}(k)|_{\hat{\theta}=\theta^*}$ corresponding to both $(2/\sqrt{3}) \sin((\pi/30)k + \phi)$ and $(2/\sqrt{3}) \sin((3\pi/30)k + \phi)$ are locally stable. This is despite the fact that in the former case, (3.18) in Assumption II) is violated and $f_{av} \neq 0$. Moreover, in this example, the corresponding θ^* are relatively close to each other. Yet, with appropriate initialization of $\hat{\theta}(k)$, both in turn have been verified via simulation to be locally stable. This indicates that the ϵ^*, μ^*, and δ^* defined in Corollary 3.1 can be allowed to have reasonably large values, without impairing local attractivity. This robust behavior is also indicative of tolerance to modest noise and almost periodicity of $x(k)$, i.e., when the ω_i and $\bar{\omega}$ do not necessarily have rational ratios as required in Assumption I). To elaborate on robustness with respect to almost periodicity, we note the following. First, a review of the proof of Theorem 3.1 reveals that even if (3.9) does not exactly equal zero, but the α in (3.10) is sufficiently large, (3.13) will still hold. As stated in [11, p. 221], an almost periodic function satisfies the periodic property within an arbitrarily small error. Thus, with an almost periodic signal, while (3.9) will not exactly hold, over certain intervals T it would be close to zero, and if over such intervals the corresponding α are sufficiently large, (3.13) may still hold. For more formal discussion of averaging theorems and almost periodic signals, refer to [7] and [12]. By a similar argument, one can relax many of the assumptions, including the requirement that interferer and desired signal frequencies be commensurate and the idealization of a noise-free environment. Thus, in the presence of bounded noise, and small enough μ, the behavior of CMA is still expected to be acceptable. The above comments, with respect to almost periodicity and noisy environment robustness, also apply to the results of Section II.

3.f) To reinforce the need for Assumption III), observe that with $\bar{\omega} = 2\pi/20$, $\omega_1 = 4\pi/20$, $\omega_2 = 8\pi/20$, (3.32) is not satisfied despite the use of $\theta^* = \bar{\theta}$.

IV. Conclusion

It has long been popular in adaptive filter theory and applications to assess and design adaptive filters based on their average behavior. Given stochastic models of the signals impinging on an adaptive filter, this inclination for average behavior analysis quite naturally involves expectations of the actual adaptive system dynamics [5]. Independence and/or mixing assumptions are common for the analytical tractability of such problems. A conceptual cousin, called averaging theory [6], is an alternative based on deterministic, (almost) periodic signal models.

A major thrust of the use of this averaging theory, as emphasized in [6], is its proof of local exponential asymptotic stability (EAS) in ideal use where the adaptive filter quality measure can be zeroed, at least on average. As described in [6], such local EAS can be combined with the total stability theorem to prove a degree of robustness in nonideal use. The potential practical impact of such theoretical analyses is substantial.

As demonstrated in this paper, these powerful techniques of [6] (and, similarly, although not exploited here, of [5]) for average behavior analysis are not limited to the prediction error forms of adaptive parameter estimators, to which they were originally applied. Indeed, as we have done here, they can be extended to use in the analysis of "new" adaptive systems with novel quality assessment mechanisms, such as the class of property restoral adaptive filter algorithms, of which the real CMA of [2] is one member.

Appendix

Proof of Theorem 3.1: Due to (3.11), there exist R, $M > 0$, such that for all $\|\theta\| \le R$ and $k > 0$

$$\left\| f(\theta, X(k)) \right\| \le M. \tag{A.1}$$

Choose μ_1 such that

$$g(\mu_1) \equiv M \sum_{i=0}^{T-1} (\mu_1)^i < \min(R, \epsilon_1) \tag{A.2}$$

and ϵ_3 such that

$$0 < \epsilon_3 < \min(R, \epsilon_1) - g(\mu_1). \tag{A.3}$$

Suppose, for a particular k,

$$\left\| \theta(k) \right\| < \epsilon_3. \tag{A.4}$$

Then for all $\mu < \mu_1$, by (3.6) and (A.1)–(A.3)

$$\left\| \theta(k+i) \right\| \le \min(R, \epsilon_1) \quad \forall i \in \{0, 1, \cdots, T-1\}. \tag{A.5}$$

Now, for all $\mu < \mu_1$ above

$$\begin{aligned} \left\| \theta(k+T) \right\|^2 &= \left\| \theta(k) \right\|^2 - 2\mu \sum_{i=0}^{T-1} \theta^T(k+i) \cdot f(\theta(k+i), X(k+i)) \\ &\quad + \mu^2 \sum_{i=0}^{T-1} \left\| f(\theta(k+i), X(k+i)) \right\|^2 \\ &\le \left\| \theta_k \right\|^2 - 2\mu \sum_{i=0}^{T-1} \theta^T(k+i) f(\theta(k+i), X(k+i)) + \mu^2 T M^2 \end{aligned} \tag{A.6}$$

where the last inequality follows from (A.5) and (A.1).

Define

$$\Delta f(k, i) = f(\theta(k+i), X(k+i)) - f(\theta(k), X(k+i)).$$

Then by (A.4), (A.5), (3.6), and (3.11),

$$\begin{aligned} \left\| \Delta f(k, i) \right\| &\le M_1 \left\| \theta(k+i) - \theta(k) \right\| \\ &\le \mu M_1 T M. \end{aligned} \tag{A.7}$$

Thus, from (A.5)–(A.7),

$$\begin{aligned} &\left\| \theta(k+T) \right\|^2 \\ &\le \left\| \theta(k) \right\|^2 - 2\mu \sum_{i=0}^{T-1} \theta^T(k+i) \cdot f(\theta(k), X(k+i)) \\ &\quad + 2\mu \sum_{i=0}^{T-1} \left\| \theta(k+i) \right\| \left\| \Delta f(k, i) \right\| + \mu^2 T M^2 \\ &\le \left\| \theta(k) \right\|^2 - 2\mu \sum_{i=0}^{T-1} \theta^T(k) \cdot f(\theta(k), X(k+i)) \\ &\quad + 2\mu \sum_{i=0}^{T-1} \left\| \theta(k+i) - \theta(k) \right\| \cdot \left\| f(\theta(k), X(k+i)) \right\| \\ &\quad + 2\mu^2 T^2 M^2 M_1 + \mu^2 T M^2. \end{aligned} \tag{A.8}$$

Furthermore, due to (3.11), there exist ϵ_2 and M_2 such that for all $\|\theta\| \le \epsilon_2$,

$$\left\| f_{av}(\theta) - \left. \frac{\partial f_{av}(\theta)}{\partial \theta} \right|_{\theta=0} \right\| \le M_2 \left\| \theta \right\|^2. \tag{A.9}$$

With $K_1 = 2T^2 M^2 M_1 + T M^2$, we have that

$$\begin{aligned} \left\| \theta(k+T) \right\|^2 &\le \left\| \theta(k) \right\|^2 - 2\mu T \theta^T(k) f_{av}(\theta(k)) + K_1 \mu^2 \\ &\le \left\| \theta(k) \right\|^2 - 2\mu T \left\| \theta(k) \right\|^2 \cdot \left[\alpha_1 - M_2 \left\| \theta(k) \right\| \right] + K_1 \mu^2. \end{aligned} \tag{A10}$$

The last inequality follows by obtaining the Taylor series expansion of $f_{av}(\theta(k))$ around $\theta = 0$, and applying (3.9), (3.10), and (A.9). Choose ϵ^* such that

$$\epsilon^* < \min\left[\epsilon_3, \epsilon_2, \frac{1}{2}\frac{\alpha_1}{M_2}\right]. \tag{A.11}$$

Then with $\|\theta(k)\| < \epsilon^*$ and $\mu < \mu_1$,

$$\|\theta(k+T)\|^2 \leq \|\theta(k)\|^2 - \mu T\|\theta(k)\|^2 + K_1\mu^2. \tag{A.12}$$

Choose K_2, $\mu_2 < \mu_1$, such that

$$K_2 = \frac{K_1}{T}, \qquad K_2\mu_2 < \epsilon^*. \tag{A.13}$$

In view of (A.2) and (A.3), the choices outlined in (A.11) and (A.13) are possible. Moreover, from (A.12), with these selections of the respective parameters with $\|\theta(0)\| \leq \epsilon^*$ and $\mu < \mu_2$,

$$\lim_{k\to\infty} \|\theta(kT)\|^2 \to [0, K_2\mu].$$

By (A.13), (3.6), and (A.1),

$$\lim_{k\to\infty} \|\hat{\theta}(k)\|^2 \to [0, K_2\mu + g(\mu)].$$

Then one can choose a K and small enough μ^* such that for $\mu < \mu^*$

$$K_2\mu + g(\mu) < K\mu < \epsilon^*,$$

whence the result follows. ▽▽▽

Acknowledgment

The authors thank J. Treichler and M. Larimore of Applied Signal Technology, Inc., Sunnyvale, CA, for their numerous helpful discussions of the constant modulus algorithm.

References

[1] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459–472, Apr. 1983.

[2] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 420–431, Apr. 1985.

[3] ——, "The tone capture properties of CMA-based interference suppressors," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 946–958, Aug. 1985.

[4] J. O. Smith and B. Friedlander, "Global convergence of the constant modulus algorithm," in *Proc. 1985 Int. Conf. Acoust., Speech, Signal Processing*, Tampa, FL, Mar. 1985, pp. 1161–1164.

[5] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* Cambridge, MA: M.I.T. Press, 1983.

[6] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging.* Cambridge, MA: M.I.T. Press, 1986.

[7] E.-W. Bai, L.-C. Fu, and S. Sastry, "Averaging analysis for discrete time and sampled data adaptive systems," Memo UCB/ERL M86/67, Univ. Calif., Berkeley, Sept. 1986.

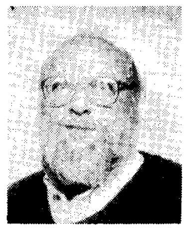
[8] E. W. Bai and S. S. Sastry, "Persistence of excitation, sufficient richness and parameter convergence in discrete time adaptive control," *Syst. Contr. Lett.*, vol. 6, pp. 153–163, Aug. 1985.

[9] J. L. Willems, *Stability Theory of Dynamical Systems.* Camden, NJ: Nelson, 1970.

[10] B. D. O. Anderson and C. R. Johnson, Jr., "Exponential convergence of adaptive identification and control algorithms," *Automatica*, vol. 18, pp. 1–13, Jan. 1982.

[11] T. L. Saaty and J. Bram, *Nonlinear Mathematics.* New York: Dover, 1964.

[12] R. L. Kosut, B. D. O. Anderson, and I. M. Y. Mareels, "Stability theory for the adaptive systems: Method of averaging and persistency of excitation," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 26–34, Jan. 1987.



C. Richard Johnson, Jr. (S'74–M'77–SM'82) was born in Macon, GA, in 1950. He received the Ph.D. degree in electrical engineering, with minors in engineering–economic systems and art history from Stanford University, Stanford, CA, in 1977.

He is currently a Professor of Electrical Engineering at Cornell University, Ithaca, NY. His research interests are in adaptive parameter estimation as applied to digital filtering, identification, and control. He is the author of *Lectures on Adaptive Parameter Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 1988).

Dr. Johnson was selected by Eta Kappa Nu as the Outstanding Young Electrical Engineer in 1982, and as the C. Holmes MacDonald Outstanding Teacher of 1983. He was a recipient of the 1982 Senior (Best Paper) Award of the Acoustics, Speech, and Signal Processing Society of the IEEE. He has served as an Editor of the *International Journal of Adaptive Control and Signal Processing*, and as an Associate Editor for the IEEE Transactions on Acoustics, Speech, and Signal Processing, the IEEE Transactions on Automatic Control, and *Automatica*.

Soura Dasgupta was born in Calcutta, India, in April 1959. He received the B.E. degree in electrical engineering from the University of Queensland, Australia, in 1980. From 1982 to 1984 he was with the Australian National University, Canberra, where he received the Ph.D. degree in systems engineering.

In 1981 he was a Research Fellow at the Electronics and Communication Sciences Unit in the Indian Statistical Institute, Calcutta. In the academic year of 1984–1985 he was a Visiting Assistant Professor in the Department of Electrical Engineering at the University of Notre Dame. He is currently with the Department of Electrical and Computer Engineering at the University of Iowa. His research interests are in adaptive systems theory, robust stability, and neural networks.

William A. Sethares received the B.A. degree in mathematics from Brandeis University, Waltham, MA, and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY.

He has worked at the Raytheon Company as a Systems Engineer, and has recently accepted a teaching position in the Department of Electrical and Computer Engineering at the University of Wisconsin in Madison. His current research interests include adaptive systems—identification, filtering, and control.