ON COMBINATIONS OF CMA EQUALIZERS

Renato Candido, Magno T. M. Silva, and Vítor H. Nascimento

Electronic Systems Engineering Department, Escola Politécnica, University of São Paulo
E-mails: {renatocan, magno, vitor}@lps.usp.br

ABSTRACT

We extend the affine combination of one fast and one slow least mean-square (LMS) filter to blind equalization, considering the combination of two constant modulus algorithms (CMA). We analyze the proposed combination in stationary and nonstationary environments verifying that there are situations where the absence of the restriction on the mixing parameter can be advantageous for the combination. Furthermore, we propose a combination of two CMAs with different initializations. Preliminary simulations show that this scheme can avoid local minima and eventually can present a faster convergence rate than that of its components.

***Index Terms*—** Adaptive filters, adaptive equalizers, blind equalization, unsupervised learning, constant modulus algorithm.

1. INTRODUCTION

Convex combinations of two fixed step-size adaptive filters have received attention due to their relative simplicity and the proof that they are universal, i.e., the combined estimate is at least as good as the best of the component filters in steady-state, for stationary inputs [1]. This scheme was proposed to improve the fundamental tradeoff between convergence rate and steady-state excess mean-square error (EMSE) in adaptive filters. It has also been exploited in nonstationary environments to improve tracking performance, considering, e.g, the algorithm proposed in [1] or the combination of algorithms with different tracking capabilities of [2].

Recently, an affine combination of two least mean-square (LMS) filters was proposed in [3]. Differently from the convex combination, the mixing parameter is not restricted to the interval $[0,\ 1]$. Thus, this method is a generalization of the convex combination.

This paper has two main contributions. First, the affine combination of [3] is extended to the combination of one fast and one slow constant modulus algorithm (CMA) [4], the most used algorithm for blind equalization. We analyze this scheme in stationary and nonstationary environments verifying that there are situations where the use of the affine combination can be advantageous, and compare this scheme to the convex combination of two CMAs proposed in [5]. Second, we show by simulations that the combination of two CMAs with different initializations can avoid local minima and may present a faster convergence rate than that of its components. In order to simplify the arguments, we assume that all quantities are real.

2. PROBLEM FORMULATION

A simplified communications system with a $T/2$-fractionally-spaced equalizer (FSE) is shown in Fig. 1. The transmitted signal $a(n)$ is assumed independent, identically distributed (i.i.d.), and non Gaussian. The unknown channel is modeled by a transfer function $H(z)$ and additive white Gaussian noise. We assume an M-tap finite impulse response (FIR) equalizer, with input vector $\mathbf{u}(n)$ and output $y(n) = \mathbf{u}^T(n)\mathbf{w}(n)$, where $\mathbf{w}(n)$ is the equalizer weight vector, and $(\cdot)^T$ indicates transposition. The equalizer must mitigate the channel effects and recover the signal $a(n)$ for some delay τ_d, obtaining at the output of the decision device the estimate $\hat{a}(n-\tau_d)$. It is well known that FSEs may achieve the zero-forcing solution in the absence of noise [6]. In this case, there exists an optimum vector $\mathbf{w}_\mathrm{o}(n)$ such that $\mathbf{u}^T(n)\mathbf{w}_\mathrm{o}(n) \approx a(n-\tau_d)$.

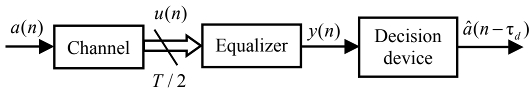

Fig. 1. Communications system with a $T/2$ FSE.

An adaptive combination of two equalizers may obtain a better compromise between convergence rate and EMSE. As depicted in Fig. 2, the outputs of the equalizers $i=1$ and $i=2$ are combined to obtain the overall output

$$y(n) = \lambda(n)y_1(n) + [1-\lambda(n)]y_2(n), \tag{1}$$

where $y_i(n) = \mathbf{u}^T(n)\mathbf{w}_i(n),\ i=1,2$ and $\lambda(n)$ is the mixing parameter. The coefficients are updated with CMA using different step-sizes, i.e.,

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu_i e_i(n)\mathbf{u}(n), \quad i=1,2, \tag{2}$$

in which $e_i(n) = \left[r - y_i^2(n)\right] y_i(n)$, $r = \mathrm{E}\{a^4(n)\}/\mathrm{E}\{a^2(n)\}$, and $\mathrm{E}\{\cdot\}$ is the expectation operator [4]. The overall "error" is defined as $e(n) = [r - y^2(n)]y(n)$.

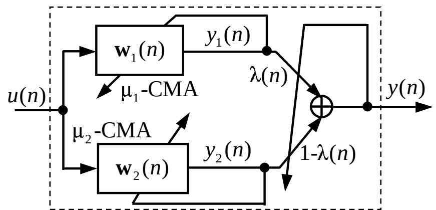

Fig. 2. Adaptive combination of two blind equalizers.

If $\lambda(n)$ is restricted to the interval $[0,\ 1]$, we have a convex combination [1, 5]. Otherwise, we have an affine combination [3]. In the convex combination of two CMAs of [5], $\lambda(n)$ is updated via a sigmoidal function and the auxiliary variable $\alpha(n)$, as shown in Table 1, where μ_α is a step-size. Theses equations were obtained in [5], using a stochastic gradient rule to minimize the instantaneous constant modulus cost $\hat{J}(n) = \left[r - y^2(n)\right]^2$. The variable $\alpha(n)$ is used to keep $\lambda(n)$ in the interval $[0,\ 1]$. This variable is restricted (by saturation) to lie inside an interval $[-\alpha^+,\ \alpha^+]$, which ensures that it does not stop updating whenever $\lambda(n)$ is close to 0 or 1 [1, 5].

In the affine combination, the auxiliary variable $\alpha(n)$ and the sigmoidal function are not used to keep the mixing parameter in $[0,\ 1]$. Considering the combination of two CMAs, the updating of $\lambda(n)$ based on the minimization of $\hat{J}(n)$ does not always ensure the desirable universal behavior of the combination. Thus, we propose

This work was supported by CNPq under grants 136050/2008-5 and 303.361/2004-2, and by FAPESP under 2008/00773-1 and 2008/04828-5.

a stochastic gradient algorithm to minimize the instantaneous square decision error $\hat{J}_d(n) = [\hat{a}(n-\tau_d) - y(n)]^2$, as shown in Table 1. If the step-size μ_λ is correctly chosen, the decision-error-based adaptation can ensure that the affine combination is nearly universal even in presence of noise. Similar to the affine combination of two LMS filters of [3], $\lambda(n)$ is constrained to be less than or equal to 2 for all n to obtain a tradeoff between stability and algorithm's tracking capability in the initial phase of adaptation.

Table 1. Adaptations of the mixing parameter.

Combination	Mixing parameter adaptation
Convex	$\lambda(n) = \mathrm{sgm}(\alpha(n)) = \{1 + \exp[-\alpha(n)]\}^{-1}$ $e_\alpha(n) = \left[r - y^2(n)\right] y(n)[y_1(n) - y_2(n)]$ $\alpha(n+1) = \alpha(n) + \mu_\alpha e_\alpha(n)\lambda(n)[1-\lambda(n)]$
Affine	$e_d(n) = \hat{a}(n-\tau_d) - y(n)$ $\lambda(n+1) = \lambda(n) + \mu_\lambda e_d(n)[y_1(n) - y_2(n)]$

3. STEADY-STATE ANALYSIS

In the tracking analysis of CMA [6,7], the optimum solution $\mathbf{w}_\mathrm{o}$ was assumed to vary according to a random-walk model, i.e., $\mathbf{w}_\mathrm{o}(n+1) = \mathbf{w}_\mathrm{o}(n) + \mathbf{q}(n)$, where $\mathbf{q}(n)$ is an i.i.d. random zero-mean vector with covariance matrix $\mathbf{Q} = \mathrm{E}\{\mathbf{q}(n)\mathbf{q}^T(n)\}$, and independent of $\mathbf{u}(m)$ for all $m \le n$ and of the initial conditions $\mathbf{w}_\mathrm{o}(0), \mathbf{w}_i(0), \lambda(0)$ [1,8].

Assuming that $\mathbf{w}_i(0)$, $i = 1, 2$ is close enough to $\mathbf{w}_\mathrm{o}(0)$ and that $\mathbf{u}^T(n)\mathbf{w}_\mathrm{o}(n) \approx a(n-\tau_d)$, the output of the equalizer i can be approximated by $y_i(n) \approx a(n-\tau_d) - e_{a,i}(n)$, where we defined the *a priori* error $e_{a,i}(n) = \mathbf{u}^T(n)[\mathbf{w}_\mathrm{o}(n) - \mathbf{w}_i(n)]^T$. Using the above assumptions and considering that the constellation used to generate $a(n)$ has circular symmetry, we can rewrite the CMA "error" using the model

$$e_i(n) \approx \gamma(n)e_{a,i}(n) + \beta(n), \qquad (3)$$

where $\gamma(n) = 3a^2(n-\tau_d) - r$ and $\beta(n) = r\,a(n-\tau_d) - a^3(n-\tau_d)$. The variable $\beta(n)$ is identically zero for constant-modulus constellations, so the variability in the modulus of $a(n)$ (as measured by $\beta(n)$) plays the role of measurement noise for CMA [2,7,9].

One measure of the equalizer performance is given by the EMSE defined as $\zeta_{ii} \triangleq \lim_{n\to\infty} \mathrm{E}\{e^2_{a,i}(n)\},\ i = 1, 2$. In the steady-state analysis of the combination of two CMA equalizers, we also have to estimate the cross-EMSE given by $\zeta_{12} \triangleq \lim_{n\to\infty} \mathrm{E}\{e_{a,1}(n)e_{a,2}(n)\}$ [1,2]. Using (3) in conjunction with the energy conservation relations of [8], ζ_{ii} and ζ_{12} can be approximated by [2,7]

$$\zeta_{ij} \approx \frac{\mu_i\mu_j\sigma^2_\beta \mathrm{Tr}(\mathbf{R}) + \mathrm{Tr}(\mathbf{Q})}{\bar{\gamma}(\mu_i + \mu_j) - \mu_i\mu_j\mathrm{Tr}(\mathbf{R})\xi},\ \ i, j = 1, 2 \qquad (4)$$

where $\sigma^2_\beta = \mathrm{E}\{a^6(n) - r^2a^2(n)\}$, $\mathbf{R} = \mathrm{E}\{\mathbf{u}(n)\mathbf{u}^T(n)\}$, $\bar{\gamma} = 3\mathrm{E}\{a^2(n)\} - r$, $\xi = r(3\mathrm{E}\{a^2(n)\} + r)$, and $\mathrm{Tr}(\cdot)$ stands for the trace of a matrix.

An analytical expression for the optimum mixing parameter in steady-state $\bar{\lambda}_\mathrm{o}(\infty)$ can be obtained by equating to zero the derivative of $\mathrm{E}\{\hat{J}_d(n)\}$ with respect to $\lambda(n)$. Using (1) and assuming that $y_i(n) \approx a(n-\tau_d) - e_{a,i}(n)$ when $n \to \infty$ [2,6], we arrive at

$$\mathrm{E}\{e_d(n)[y_1(n) - y_2(n)]\} = \mathrm{E}\{e_d(n)[e_{a,2}(n) - e_{a,1}(n)]\} = 0.$$

Noting that in steady-state the overall *a priori* error is a combination of the *a priori* errors of the component filters, i.e.,

$$e_a(n) = \lambda(n)e_{a,1}(n) + [1-\lambda(n)]e_{a,2}(n), \qquad (5)$$

using (3), and assuming that $\lambda_\mathrm{o}(n)$ is independent of $e_{a,i}(n)$, after some algebraic manipulations we arrive at

$$\bar{\lambda}_\mathrm{o}(\infty) = \frac{\Delta\zeta_2}{\Delta\zeta_1 + \Delta\zeta_2}, \qquad (6)$$

where $\Delta\zeta_i = \zeta_{ii} - \zeta_{12}$, $i = 1, 2$. A similar expression was also obtained in [1, Eq.(29)] for the convex combination of two LMS filters. The difference is that in the convex combination, $\lambda(n)$ and consequently $\bar{\lambda}_\mathrm{o}(\infty)$ are restricted to the interval $[0,\ 1]$.

By squaring and taking the expectations of both sides of (5) with $\lambda(n) = \lambda_\mathrm{o}(n)$ and assuming that the variance of $\lambda_\mathrm{o}(n)$ is sufficiently small such that $\lim_{n\to\infty} \mathrm{E}\{\lambda^2_\mathrm{o}(n)\} \approx \bar{\lambda}_\mathrm{o}^{\,2}(\infty)$, we obtain the following expression for the steady-state EMSE of the combination

$$\zeta \approx \zeta_{12} + \bar{\lambda}(\infty)\Delta\zeta_1 = \zeta_{12} + \frac{\Delta\zeta_1\Delta\zeta_2}{\Delta\zeta_1 + \Delta\zeta_2}. \qquad (7)$$

This expression was also obtained in [1, Eq. (33)] for the convex combination of two LMS filters. Note that when $\bar{\lambda}_\mathrm{o}(\infty) \approx 1$, $\zeta \approx \zeta_{11}$ and when $\bar{\lambda}_\mathrm{o}(\infty) \approx 0$, $\zeta \approx \zeta_{22}$. On the other hand, since $\lambda(n)$ is restricted to $[0,\ 1]$ in the convex combination, only for $0 < \bar{\lambda}_\mathrm{o}(\infty) < 1$ the convex combination can outperform the component equalizers. We show next that in some situations the absence of the restriction on $\lambda(n)$ can be advantageous for the combination.

3.1. Stationary performance

Replacing the model (4) with $\mathbf{Q} = \mathbf{0}$ in (6) and (7), we obtain for the stationary case

$$\bar{\lambda}_\mathrm{o}(\infty) \approx \frac{\delta\left[2 - \mu_1\mathrm{Tr}(\mathbf{R})\xi\,\bar{\gamma}^{-1}\right]}{2\,(\delta - 1)} \quad \text{and} \qquad (8)$$

$$\zeta \approx \frac{1}{2}\frac{\mu_2\sigma^2_\beta\mathrm{Tr}(\mathbf{R})}{(1+\delta)\bar{\gamma} - \mu_2\mathrm{Tr}(\mathbf{R})\xi}, \qquad (9)$$

where we have defined $\delta \triangleq \mu_2/\mu_1$, with $0 < \delta < 1$.

To ensure the stability of μ_1-CMA, $\mu_1 < 2\bar{\gamma}/(3\mathrm{Tr}(\mathbf{R})\xi)$ must be satisfied [9, Eq. (14)]. Hence, $\bar{\lambda}_\mathrm{o}(\infty)$ is always negative, which does not occur in the convex combination due to the restriction on $\lambda(n)$. This behavior was observed in [3] for the affine combination of two LMS filters. Comparing ζ to ζ_{22}, we conclude that the affine combination can outperform both components in steady-state. Specially, when $\delta \to 1$, i.e., when the components have approximately the same step-size, $\bar{\lambda}_\mathrm{o}(\infty) \to -\infty$, $\zeta \to \zeta_{22}/2$, and a 3 dB reduction occurs. On the other hand, if $\mu_1 \approx \mu_2$, the convex combination performs close to one of its components, and an EMSE reduction does not occur.

In order to explain the behavior of the affine combination when $\mu_1 \approx \mu_2$, using (1) and the model (3), the overall steady-state error is written as

$$e(n) = \underbrace{e_2(n)}_{d(n)} + \lambda(n)\underbrace{\gamma(n)[\mathbf{w}_2(n) - \mathbf{w}_1(n)]^T\mathbf{u}(n)}_{-x(n)}. \qquad (10)$$

From the point of view of the computation of $\lambda(n)$, $d(n)$ represents the signal which has to be estimated, and $x(n)$ plays the role of input signal. If $\mathbf{w}_i$ varies slowly compared to λ, the affine combination seeks the best weight vector in the line $\mathbf{w}_2 + \lambda(\mathbf{w}_1 - \mathbf{w}_2)$. In the case of $\mu_1 \approx \mu_2$, we also have $\mathbf{w}_1 \approx \mathbf{w}_2$, and λ has to assume a large value to take the combined vector as close as possible to $\mathbf{w}_\mathrm{o}$, since the input signal $x(n)$ depends on the difference between $\mathbf{w}_1$ and $\mathbf{w}_2$. Thus, if $\delta \to 0$, $(\mathbf{w}_1 - \mathbf{w}_2) \to 0$, and $|\lambda| \to \infty$.

The properties of the affine combination in a stationary environment can be exploited to improve the EMSE reduction of the combination. To verify if a larger reduction can be achieved, we consider the scheme of Fig. 3, where the outputs of two affine combinations are combined with a mixing parameter to obtain the overall output. The first combination considers two CMA equalizers with step-sizes μ_1 and $\mu_2 = \delta_1\mu_1$ with $0 \ll \delta_1 < 1$. The second combines two CMAs with μ_3 and $\mu_4 = \delta_2\mu_3$ with $0 \ll \delta_2 < 1$. To obtain the largest EMSE reduction of the scheme, we assume that the four step-sizes are different but close to one another. The steady-state performance of the proposed scheme can be evaluated using (7). Besides $\zeta_{ii},\ i = 1, \ldots, 4$, ζ_{12}, and ζ_{34}, we have to estimate ζ_{13}, ζ_{14}, ζ_{23}, and ζ_{24}. Thus, after some algebraic manipulations, we conclude that the overall EMSE for close step-sizes is

$$\lim_{(\delta_1,\delta_2)\to(1,1)} \zeta \approx \frac{3}{8}\zeta_{11}, \qquad (11)$$

which represents an EMSE reduction of 4.26 dB.

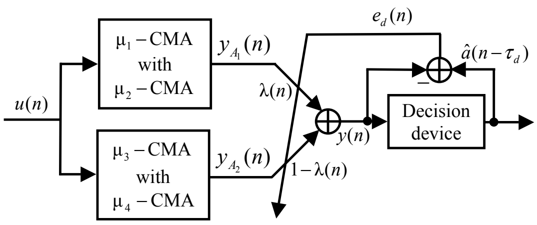

Fig. 3. A combination of two affine combinations to improve the EMSE reduction in a stationary environment.

3.2. Nonstationary performance

As in the stationary case, the largest EMSE reduction of the affine combination in relation to its components occurs when $\zeta_{11} \approx \zeta_{22}$. This can happen in two situations: (i) when $\mathrm{Tr}(\mathbf{Q}) \approx \mu_1\mu_2\sigma_\beta^2\mathrm{Tr}(\mathbf{R})$; or (ii) when $\mu_1 \approx \mu_2$. In case (i), replacing the model (4) under the small step-size approximation in (7), after some algebra, we arrive at

$$\frac{\zeta}{\zeta_{ii}} \approx \frac{1}{2} + \frac{2\delta}{(\delta+1)^2},\ i=1,2. \tag{12}$$

In case (ii), we obtain the following limit (recall that $\zeta_{11} \approx \zeta_{22}$)

$$\lim_{\delta\to 1} \zeta = \frac{\zeta_{22}}{2} + \frac{\sigma_\beta^2\mathrm{Tr}(\mathbf{R})\mathrm{Tr}(\mathbf{Q})}{2\bar{\gamma}^2\zeta_{22}}. \tag{13}$$

Note that the EMSE reduction in both cases is limited by 3 dB. A reduction close to 3 dB will occur when $\delta \to 0$ in (12) or when the second term of the r.h.s. of (13) can be disregarded in relation to $\zeta_{22}/2$. We should notice that case (i) also occurs for the convex combination. However, in case (ii), the convex combination performs as its best component, due to the $\lambda(n)$ restriction. Although there may exist an EMSE reduction, the minimum steady-state EMSE of both combinations is equal to the steady-state EMSE of a CMA equalizer with optimal step-size μ_o, which happens when $\mathrm{Tr}(\mathbf{Q}) \approx q_i,\quad i=1,2$, where $q_i \triangleq \mu_i^2\sigma_\beta^2\mathrm{Tr}(\mathbf{R})$ [10].

4. SIMULATIONS

In the first two simulations, we assume 4-PAM (pulse amplitude modulation) such that $r = 8.2$, $\sigma_\beta^2 = 28.8$, and $\bar{\gamma} = 6.8$, and an FIR channel with coefficients $[0.1,\ 0.3,\ 1,\ -0.1,\ 0.5,\ 0.2]$ in the absence of noise [6]. In the combinations, each component filter has $M = 4$ coefficients as a $T/2$-FSE and is initialized with only one non-null element in the second position.

To verify the behavior of the scheme of Fig. 3, we consider two affine combinations of CMAs with close step-sizes. Fig. 4 shows the EMSE and mixing parameter along the iterations estimated from the ensemble-average of 500 independent runs. To facilitate the visualization, the curves were filtered by a moving-average filter with 512 coefficients. The dashed lines in the figure show the predicted values of ζ for each algorithm and their combinations. Since the four component equalizers have close EMSEs, we show in Fig. 4 only the EMSEs of μ_1-CMA and μ_4-CMA. Although there is no exact agreement between analysis and simulation, the predicted values model the overall behavior of the algorithms and of their combinations. Note that a difference of a few dB is common in models for blind algorithms, due to the strong assumptions necessary for the analysis. We can observe from the figure that affine combinations of two CMAs with close step-sizes provide an EMSE reduction of approximately 3 dB as predicted by (9). An affine combination of the outputs of the combinations provides a reduction of approximately 4.26 dB in relation to each component equalizer, as predicted by (11). A drawback of this scheme is that the combinations converge slowly, since the convergence of the algorithm for the updating of the mixing parameters depends on the difference of the outputs of the components (see Table 1), which is very small in this case.

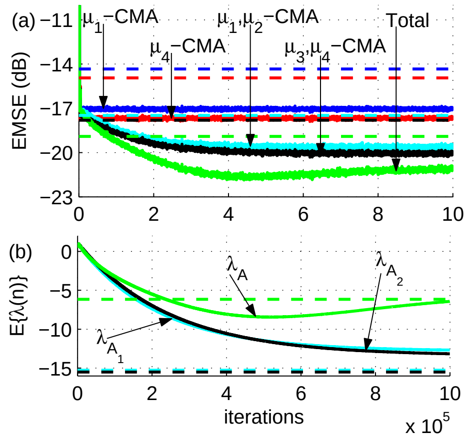

Fig. 4. (a) Theoretical and experimental EMSE for the scheme of Fig. 3 with $\mu_1 = 10^{-3}$, $\mu_3 = 9.4\times 10^{-4}$, $\mu_2 = \delta_1\mu_1$, $\mu_4 = \delta_2\mu_3$, $\delta_1 = \delta_2 = 0.95$, and $\mu_{\lambda_1} = \mu_{\lambda_2} = \mu_\lambda = 0.1$. (b) Ensemble-average of $\lambda(n)$, $\lambda_{A_1}(n)$, $\lambda_{A_2}(n)$ and theoretical value of $\bar{\lambda}_o(\infty)$; ensemble-average of 500 independent runs; the theoretical values are indicated by dashed lines; $\lambda_{A_i},\ i=1,2$ are the mixing parameters of the affine combinations.

As in [1], we also use in a nonstationary environment the normalized square deviation (NSD): $\mathrm{NSD}_i(\infty) = \zeta_{ii}/\zeta_o,\quad i=1,2$, $\mathrm{NSD}_{12}(\infty) = \zeta_{12}/\zeta_o$, $\mathrm{NSD}(\infty) = \zeta/\zeta_o$, where ζ_o is the optimum steady-state EMSE of a CMA equalizer [10]. Fig. 5 shows the theoretical and experimental NSD as a function of $\mathrm{Tr}(\mathbf{Q})$ for the affine and convex combinations of two CMAs. Since $\delta = 0.1$, the EMSE reduction provided by the affine combination outside the interval $[q_2,\ q_1]$ is almost imperceptible for the level of detail in the figure. In this example, both combinations present very similar performance, in spite of the restriction on $\lambda(n)$ of the convex combination. Note that the maximum EMSE reduction is of approximately 1.8 dB (as predicted by (12)) and occurs for both combinations at $\mathrm{Tr}(\mathbf{Q}) \approx \mu_1\mu_2\sigma_\beta^2\mathrm{Tr}(\mathbf{R}) \approx 4\times 10^{-7}$, where $\zeta_{11} \approx \zeta_{22}$. This reduction makes both combinations have an EMSE close to the optimum inside $[q_2,\ q_1]$. The theoretical models (4), (7), and (6) show good agreement with the experimental results.

In order to avoid local minima, we combine two CMAs with the same step-size but with different initializations. To improve the convergence rate of the algorithms of Table 1, we consider a normalized adaptation scheme similar to that of [11]. Basically, the step-sizes μ_α and μ_λ are divided by $[b(n)+\epsilon]$, where $b(n) = \rho b(n-1) + (1-\rho)[y_1(n) - y_2(n)]^2$ with the forgetting factor $0 \ll \rho < 1$ and the small positive constant ϵ. Furthermore, we replace the overall CMA "error" $e(n)$ in the adaptation of $\alpha(n)$ of the convex combination by the decision error $e_d(n)$ to ensure its nearly universal behavior. We assume the transmission of binary signals with symbols $\{\pm 1\}$ through the channel $H(z) = [1+0.6z^{-1}]^{-1}$ with signal-to-noise ratio (SNR) of 25 dB, and an FIR equalizer with $M = 2$ coefficients working in the symbol rate. Fig. 6 shows an ensemble average of 10^3 independent runs for two different initialization sets. In situation 1 (Fig. 6-a), the combinations perform close to the best component in steady-state and reach the global minimum. The affine combination presents a faster convergence, since it seeks the best weight vector in the whole line $\mathbf{w}_2 + \lambda(\mathbf{w}_1 - \mathbf{w}_2)$. Note that in the convergence, the mixing parameter of the affine combination is negative (Fig. 6-d). The restriction on $\lambda(n)$ in the convex combination causes its slower convergence. In situation 2, μ_2-CMA may converges to two minima with EMSEs of -5 dB and -20 dB, as shown in Figs. 6-b and c, respectively. In Fig. 6-b, for 412 out of 10^3 realizations of the filters, the combinations performed similarly to situation 1. In the remaining 588 realizations (Fig. 6-c), the affine combination has an interesting behavior at the beginning of the convergence: it rapidly achieves -9 dB and gets close to the global minimum, but returns to the local minimum when μ_2-CMA converges to -5 dB. Note that, in steady-state both components get stuck at the same local minimum. The convex combination performs like μ_2-CMA and the affine combination makes an useless effort to reverse this situation since $\mathrm{E}\{\lambda(n)\} \rightarrow -200$.

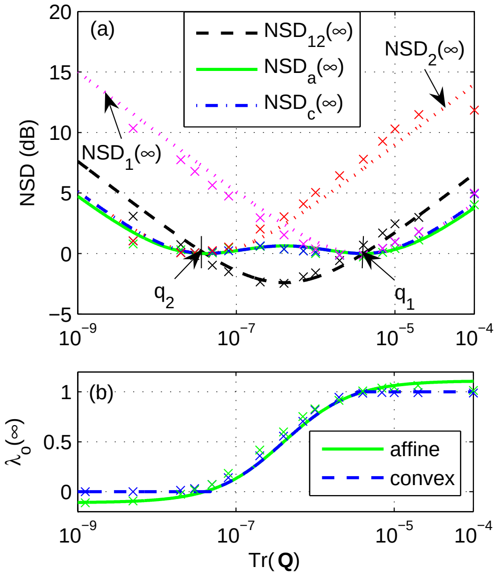

Fig. 5. (a) Theoretical and experimental NSD of two CMA equalizers ($\mu_1 = 10^{-4}$, $\mu_2 = 10^{-5}$), their cross-NSD, and NSD of their convex ($\mathrm{NSD_c}$, $\mu_\alpha = 0.075$, $\alpha^+ = 4$) and affine ($\mathrm{NSD_a}$, $\mu_\lambda = 0.0075$) combinations as a function of $\mathrm{Tr}(\mathbf{Q})$; (b) Theoretical and experimental steady-state optimum mixing parameter. The experimental points are indicated by $\times$ and were obtained by an ensemble-average of 50 independent runs.

5. CONCLUSION

We proposed and analyzed an affine combination of two CMA equalizers. Analytical expressions for the steady-state optimum mixing parameter and for the steady-state EMSE of the combination were obtained for stationary and nonstationary environments. Due to the absence of the restriction on the mixing parameter, when the component equalizers have close step-sizes, the affine combination can provide an EMSE reduction limited to 3 dB. However, in a nonstationary environment, the minimum steady-state EMSE of the convex or affine combinations is equal to the steady-state EMSE of a CMA equalizer with optimal step-size. Thus, depending on the step-sizes of the components, the affine combination has similar performance to that of the convex combination in a nonstationary environment. Additionally, To avoid local minima, we combined two CMAs with the same step-sizes but with different initializations. Through simulations, we observed that there may exist situations where the combined scheme avoids local minima. Comparing to the convex combination, the affine combination may present faster convergence and search a minimum more efficiently. Further work is necessary to extend this scheme to the combination of more than two filters with more coefficients, and to ensure convergence to a global minimum.

6. REFERENCES

[1] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, “Mean-square performance of a convex combination of two adaptive filters,” *IEEE Trans. Signal Process.*, vol. 54, pp. 1078–1090, Mar. 2006.

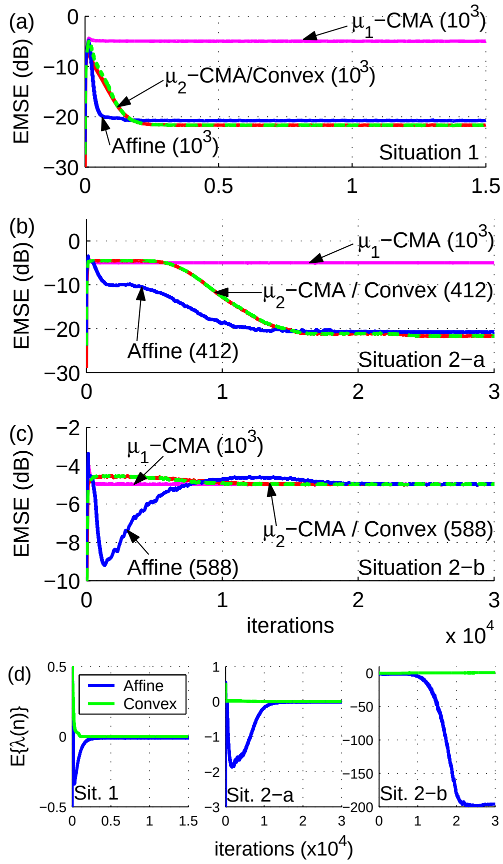

Fig. 6. EMSE for two CMAs and their convex and affine combinations (a) $\mathbf{w}_1^T(0) = [0.05,\ -0.4]$, $\mathbf{w}_2^T(0) = [1,\ 0]$, $\mu_\alpha = 0.2$; (b) and (c) $\mathbf{w}_1^T(0) = [0.05,\ 0.4]$, $\mathbf{w}_2^T(0) = [-0.6,\ 0.4]$, $\mu_\alpha = 3$; (d) mixing parameter; $\mu_1 = \mu_2 = 0.016$, $\alpha^+ = 4$, $\alpha(0) = 0$, $\mu_\lambda = 0.4$, $\rho = 0.9$, $\epsilon = 5 \times 10^{-5}$; (b) and (c) are related to the same initialization; in (b) we show an ensemble-average when μ_2-CMA achieves a steady-state EMSE of -20dB; in (c) the ensemble-average corresponds to the convergence to -5 dB. The number of independent runs for each situation is in parenthesis.

[2] M. T. M. Silva and V. H. Nascimento, “Improving the tracking capability of adaptive filters via convex combination,” *IEEE Trans. Signal Process.*, vol. 56, pp. 3137–3149, Jul. 2008.

[3] N. J. Bershad, J. C. M. Bermudez, and J-Y. Tourneret, “An affine combination of two LMS adaptive filters - transient mean-square analysis,” *IEEE Trans. Signal Process.*, vol. 56, pp. 1853–1864, May 2008.

[4] D. N. Godard, “Self-recovering equalization and carrier tracking in two dimensional data communication system,” *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.

[5] J. Arenas-García and A. R. Figueiras-Vidal, “Improved blind equalization via adaptive combination of constant modulus algorithms,” in *Proc. of ICASSP’06*. IEEE, 2006, vol. III, pp. 756–759.

[6] N. R. Yousef and A. H. Sayed, “A feedback analysis of the tracking performance of blind adaptive equalization algorithms,” in *Proc. of IEEE Conf. on Dec. and Control*, Dec. 1999, vol. 1, pp. 174–179.

[7] M. T. M. Silva and V. H. Nascimento, “Tracking analysis of the Constant Modulus Algorithm,” in *Proc. of ICASSP’08*, 2008, pp. 3561–3564.

[8] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, NJ, 2003.

[9] V. H. Nascimento and M. T. M. Silva, “Stochastic stability for the Constant Modulus Algorithm,” *IEEE Trans. Signal Process.*, vol. 56, pp. 4984–4989, Oct. 2008.

[10] M. T. M. Silva and M. D. Miranda, “Tracking issues of some blind equalization algorithms,” *IEEE Signal Process. Lett.*, vol. 11, pp. 760–763, Sept. 2004.

[11] L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, and J. Arenas-García, “A normalized adaptation scheme for the convex combination of two adaptive filters,” in *Proc. of ICASSP’2008*. IEEE, 2008, pp. 3301–3304.