ON THE TRACKING PERFORMANCE OF COMBINATIONS OF LEAST MEAN SQUARES AND RECURSIVE LEAST SQUARES ADAPTIVE FILTERS

Vítor H. Nascimento[†], *Magno T. M. Silva*[†], *Luiz A. Azpicueta-Ruiz*[‡], *and Jerónimo Arenas-García*[‡]

[†] University of São Paulo, Brazil
{vitor,magno}@lps.usp.br

[‡] Univ. Carlos III de Madrid, Spain
{lazpicueta, jarenas}@tsc.uc3m.es

ABSTRACT

Combinations of adaptive filters have attracted attention as a simple solution to improve filter performance, including tracking properties. In this paper, we consider combinations of LMS and RLS filters, and study their performance for tracking time-varying solutions. We show that a combination of two filters from the same family (i.e., two LMS or two RLS filters) cannot improve the performance over that of a single filter of the same type with optimal selection of the step size (or forgetting factor). However, combining LMS and RLS filters it is possible to simultaneously outperform the optimum LMS and RLS filters. In other words, combination schemes can achieve smaller errors than optimally adjusted individual filters. Experimental work in a plant identification setup corroborates the validity of our results.

Index Terms— Adaptive filters, convex combination, steady-state analysis, tracking performance, RLS algorithm, LMS algorithm.

1. INTRODUCTION

Combinations of adaptive filters have gained considerable attention lately, since they decrease the sensitivity of the filter to choices of parameters such as the step size, forgetting factor or filter length (see, e.g., [1–5]). Using a combination of two filters with different step sizes, for example, one can obtain fast convergence and low steady-state misadjustment, or use the combination to find the optimum step size in a nonstationary environment [1]. In general, this combination approach is more robust than variable step-size schemes [5].

In tracking of time-varying scenarios, combination schemes offer improved tracking capabilities with respect to the component filters [4]. However, it has been noticed in simulations that the excess mean-square error (EMSE) obtained by the combination of two filters of the same family [e.g., two least mean-squares (LMS) filters with different step sizes, or two recursive least-squares (RLS) filters with different forgetting factors] will never be better than the performance of a single filter employing the optimum step size (or optimum forgetting factor) for a given nonstationary condition [1, 6].

More recently, the combination of filters from different families (one LMS and one RLS) was proposed as a way to take advantage of the different tracking properties of LMS and RLS [5]. In fact, despite the fast initial convergence provided by RLS, it was shown in [7] that LMS may outperform RLS depending on how the optimum solution changes with time. Although a theoretical analysis and several simulations were provided in [5], it was not noticed that the combination may actually obtain a smaller EMSE than its component filters even when the optimum step size is used for LMS and the optimum forgetting factor is used for RLS. In other words, the combination of filters of different families may obtain a performance that would not be possible with a combination of two filters of the same family, or with an algorithm that chooses the optimum step size for LMS (or the optimum forgetting factor for RLS). In this paper we illustrate these facts both through theoretical analysis and simulations.

The work of Nascimento and Silva was partly supported by CNPq under Grants 303361/2004-2 and 302633/2008-1; and FAPESP under Grants 2008/04828-5 and 2008/00773-1. The work of Azpicueta-Ruiz and Arenas-García was partly supported by MEC project TEC2008-02473 and CAM project S-0505/TIC/0223.

Table 1. Parameters of the considered algorithms.

Alg.	ρ_i	$\boldsymbol{M}_i^{-1}(n)$
LMS	μ_i	$\boldsymbol{I}$
RLS	1	$\widehat{\boldsymbol{R}}_i(n) = \sum_{l=1}^{n} \lambda_i^{n-l} \boldsymbol{u}(l)\boldsymbol{u}^T(l)$

The paper is organized as follows. In the next section, we present the data model and introduce the notation that will be used throughout the paper, reviewing also some results for the tracking performance of LMS and RLS filters. Then, in Sec. 3 we recall some theoretical results regarding the performance of convex and affine combination of adaptive filters in nonstationary environments. We also prove that combinations of filters of the same family cannot improve the performance over that of a single filter with optimum selection of the parameters (i.e., step size or forgetting factor), and that this limitation can be overcome when filters of different families are combined. Several examples that validate the analysis are provided in Sec. 4. Finally, Sec. 5 presents the main conclusions of our work.

2. PROBLEM FORMULATION

In this paper, we consider combinations of two LMS, two RLS or one RLS and one LMS filters. The update laws for LMS and RLS may be written as

$$\boldsymbol{w}_i(n) = \boldsymbol{w}_i(n-1) + \rho_i \boldsymbol{M}_i(n)\boldsymbol{u}(n)e_i(n), \qquad (1)$$

where the index $i = 1, 2$ refers to each filter in the combination, $\boldsymbol{w}_i(n) \in \mathbb{R}^M$ is the coefficient vector of each filter at time n, ρ_i is a step-size parameter, $\boldsymbol{u}(n) \in \mathbb{R}^M$ is the input regressor vector, and $e_i(n)$ is the estimation error. For RLS, $\boldsymbol{M}_i(n) \in \mathbb{R}^{M\times M}$ is an estimate of the inverse of the regressor autocovariance matrix, $\boldsymbol{R} = \mathrm{E}\{\boldsymbol{u}(n)\boldsymbol{u}^T(n)\}$, and can be computed in an efficient way using the matrix inversion lemma or, if lattice algorithms are used, its explicit evaluation may be avoided [8]. Table 1 lists the values of the different parameters in (1) for each class of the considered filters.

The estimation error is given by

$$e_i(n) = d(n) - y_i(n), \qquad (2)$$

Table 2. Analytical expressions for the steady-state EMSEs of LMS and RLS.

Alg.	ζ
LMS	$\dfrac{\mu_i\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\}+\mu_i^{-1}\,\mathrm{Tr}\{\boldsymbol{Q}\}}{2}$
RLS	$\dfrac{\nu_i\sigma_v^2 M+\nu_i^{-1}\,\mathrm{Tr}\{\boldsymbol{QR}\}}{2}$

Table 3. Optimum tracking parameters (μ_{o} and ν_{o}) and steady-state EMSEs (ζ_{o}) for LMS and RLS.

Alg.	$\mu_{\mathrm{o}},\nu_{\mathrm{o}}$	ζ_{o}
LMS	$\sqrt{\dfrac{\mathrm{Tr}\{\boldsymbol{Q}\}}{\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\}}}$	$\sqrt{\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\}\,\mathrm{Tr}\{\boldsymbol{Q}\}}$
RLS	$\sqrt{\dfrac{\mathrm{Tr}\{\boldsymbol{QR}\}}{\sigma_v^2 M}}$	$\sqrt{\sigma_v^2 M\,\mathrm{Tr}\{\boldsymbol{QR}\}}$

where $y_i(n)=\boldsymbol{w}_i(n-1)^T\boldsymbol{u}(n)$, $i=1,2$ are the filter outputs, and $d(n)$ is the desired response. As it is well-known, there exists a linear regression model relating $d(n)$ and $\boldsymbol{u}(n)$, such that

$$d(n)=\boldsymbol{w}_{\mathrm{o}}^T\boldsymbol{u}(n)+v(n), \tag{3}$$

where $v(n)$ is white and zero-mean measurement noise with variance σ_v^2, uncorrelated with $\boldsymbol{u}(n)$, and $\boldsymbol{w}_{\mathrm{o}}$ provides the optimum linear least mean-squares estimate of $d(n)$ given $\boldsymbol{u}(n)$ [8]. We assume in this paper, as it is common in studies of adaptive filters, that $\boldsymbol{u}(n)$ and $v(n)$ are stationary zero-mean processes, and moreover that $v(n)$ is independent of $\boldsymbol{u}(m)$ for all m,n.

Define also the *a priori* errors $e_{a,i}(n)=[\boldsymbol{w}_{\mathrm{o}}-\boldsymbol{w}_i(n-1)]^T\boldsymbol{u}(n)$. Under the stated assumptions, the mean-square error $\mathrm{E}\{e_i^2(n)\}$ may be shown to be [8]

$$\mathrm{E}\{e_i^2(n)\}=\zeta_i(n)+\sigma_v^2,$$

where $\zeta_i(n)=\mathrm{E}\{e_{a,i}^2(n)\}$ is the so-called excess mean-square error (EMSE) of each filter.

The optimum weight vector $\boldsymbol{w}_{\mathrm{o}}$ is usually not constant in practice. A common approach to model its variations is through the brownian motion model below, which allows a tractable analysis

$$\boldsymbol{w}_{\mathrm{o}}(n)=\boldsymbol{w}_{\mathrm{o}}(n-1)+\boldsymbol{q}(n), \tag{4}$$

where $\{\boldsymbol{q}(n)\}$ is a sequence of i.i.d. vectors with zero mean and autocorrelation matrix $\boldsymbol{Q}=\mathrm{E}\{\boldsymbol{q}(n)\boldsymbol{q}^T(n)\}$.

Using this model and (1)–(3), and assuming sufficiently small step size and forgetting factor sufficiently close to 1, it can be shown that the steady-state EMSE [i.e., $\zeta_i=\lim_{n\to\infty}\zeta_i(n)$] of LMS and RLS are as given in Table 2 [5,7,8]. For convenience, we will often use $\nu_i=1-\lambda_i$. Since ν_i plays in the expressions for RLS a similar role to that of μ_i for LMS, we will also refer to ν_i as a "step size". Differentiating the expressions in Table 2 with respect to either μ_i or ν_i, one can compute the optimum step sizes μ_{o} and ν_{o} for a given environment (i.e., values of $\boldsymbol{Q}$, $\boldsymbol{R}$ and noise variance σ_v^2). These optimum values, along with the resulting optimum EMSEs for each filter, are given in Table 3 [7,8].

Through these expressions, it was shown in [7] that, despite its slow initial convergence, LMS may present better tracking performance than RLS, depending on the values of $\boldsymbol{Q}$ and $\boldsymbol{R}$. In particular, if $\boldsymbol{Q}$ is proportional to $\boldsymbol{R}$, the optimum steady-state EMSE of LMS will be smaller than that of RLS. On the contrary, if $\boldsymbol{Q}$ is proportional to $\boldsymbol{R}^{-1}$, RLS will present better performance. When $\boldsymbol{Q}$ is proportional to $\boldsymbol{I}$ (the identity matrix), both algorithms will present similar behavior [7].

In the next section we introduce combinations of adaptive filters, and prove that the tracking performance of the combination of two LMS or two RLS filters is lower bounded by the values of Table 3, but the combination of one RLS with one LMS algorithm may achieve a better performance.

3. CONVEX COMBINATIONS AND OPTIMAL TRACKING

One promising way of increasing the performance of adaptive filters is to run two or more filters in parallel, and combine their outputs constructing an overall output given by

$$y(n)=\eta(n-1)y_1(n)+[1-\eta(n-1)]y_2(n).$$

Up to now, good results have been obtained with both a convex combination model, in which the mixing parameter $\eta(n)$ is constrained to remain in the interval $[0,1]$ [1], and an affine combination model, in which $\eta(n)$ may be any real number [3]. In the first case, the mixing parameter is usually updated using an auxiliary variable $a(n)$, according to

$$a(n)=a(n-1)+\tilde{\mu}_a(n)e(n)\Delta e(n)\,\eta(n-1)\,[1-\eta(n-1)],$$
$$\eta(n)=\frac{1}{1+\exp\left[-a(n)\right]},$$

in which $\tilde{\mu}_a(n)$ is a (possibly normalized) step size, $e(n)=d(n)-y(n)$ is the overall estimation error, and $\Delta e(n)=e_2(n)-e_1(n)$. In general, to avoid slow adaptation close to $\eta=1$ or $\eta=0$, $a(n)$ is constrained (by simple saturation) to the interval $[-a_+,a_+]$. A common choice for a_+ is 4.

For affine combinations, one possible method for updating the mixing parameter is through the recursion

$$\eta(n)=\eta(n-1)+\mu_a(n)e(n)\Delta e(n).$$

In both cases, it is convenient to choose a normalized step size $\tilde{\mu}_a(n)$ using an estimate $p(n)$ of $\mathrm{E}\{\Delta e^2(n)\}$, such that

$$p(n)=\lambda_p p(n-1)+(1-\lambda_p)\Delta e^2(n),$$

where λ_p is a forgetting factor, and $\tilde{\mu}_a(n)=\mu_a/[\delta+p(n)]$, δ being a small regularization term [6,9].

It can be shown that the optimum mixing parameter in steady state is given by [1,3,6]

$$\eta_*=\frac{\zeta_2-\zeta_{12}}{\zeta_1-2\zeta_{12}+\zeta_2}, \tag{5}$$

where $\zeta_{12}=\lim_{n\to\infty}\mathrm{E}\{e_{a,1}(n)e_{a,2}(n)\}$ is the steady-state cross-EMSE between both filters in the combination, given in Table 4 for the different combination possibilities considered in this paper [5]. For convex combinations η_* is given by (5) only if the value falls in the interval $[0,1]$, otherwise $\eta_*=0$ (resp. 1), if (5) is negative (resp. larger than 1).

The EMSE of the combination, using the optimum η_*, is given by

$$\zeta_*=\frac{\zeta_1\zeta_2-\zeta_{12}^2}{\zeta_1-2\zeta_{12}+\zeta_2} \tag{6}$$

Table 4. Analytical expressions for the steady-state cross-EMSE of the considered combinations.

Combination	ζ_{12}
μ_1-LMS and μ_2-LMS	$\dfrac{\mu_1\mu_2\,\mathrm{Tr}\{\boldsymbol{R}\}\sigma_v^2 + \mathrm{Tr}\{\boldsymbol{Q}\}}{\mu_1+\mu_2}$
λ_1-RLS and λ_2-RLS	$\dfrac{\nu_1\nu_2 M\sigma_v^2 + \mathrm{Tr}\{\boldsymbol{QR}\}}{\nu_1+\nu_2}$
λ_1-RLS and μ_2-LMS	$\mu_2\nu_1\,\sigma_v^2\ \mathrm{Tr}\left\{\boldsymbol{\Sigma}\right\} + \mathrm{Tr}\left\{\boldsymbol{Q\Sigma}\right\}$, where $\boldsymbol{\Sigma} \triangleq (\nu_1 \boldsymbol{I} + \mu_2\boldsymbol{R})^{-1}\boldsymbol{R}$.

We remark that the optimality here is with respect to the choice of η only, which we denote by the subscript $*$.

We will now search for the optimum ζ_* with respect to the step sizes for the particular case of a combination of two filters of the same family. Consider first the combination of two LMS filters. Assuming that the optimal η_* is selected in steady state (which is usually a good approximation for the considered recursions [6, 9]), the steady-state EMSE of the combination will be given by (6) with ζ_1 and ζ_2 given by the first row in Table 2, and ζ_{12} given by the first row in Table 4. Differentiating ζ_* with respect to μ_2, and after some manipulations, we obtain

$$\frac{\mathrm{d}\,\zeta_*^{\mathrm{LMS}}}{\mathrm{d}\,\mu_2} = -\frac{1}{2}\frac{(\mathrm{Tr}\{\boldsymbol{Q}\} - \mu_1^2\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\})^2(\mathrm{Tr}\{\boldsymbol{Q}\} - \mu_2^2\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\})}{(\mu_1\mu_2\sigma_v^2\,\mathrm{Tr}\{\boldsymbol{R}\} + \mathrm{Tr}\{\boldsymbol{Q}\})^2(\mu_1+\mu_2)^2},$$

where we have used the superscript $^{\mathrm{LMS}}$ to emphasize that we are considering a combination of two LMS filters. From this expression one can see that $\mathrm{d}\,\zeta_*^{\mathrm{LMS}}/\,\mathrm{d}\,\mu_2 = 0$ if $\mu_2 = \mu_{\mathrm{o}}$ (notice the factor depending on μ_2 in the numerator), so the minimum value of ζ_*^{LMS} is attained when one of the component filters has optimum step size μ_{o}. Note that, due to the symmetry of the combination, we would reach the same conclusion if we had differentiated with respect to μ_1 (this is easily seen if one replaces η by $1-\eta$ in all expressions).

Furthermore, if we substitute $\zeta_2 = \zeta_{\mathrm{o}}^{\mathrm{LMS}}$ and $\mu_2 = \mu_{\mathrm{o}}$ in the expressions for η_* from (5) and for ζ_{12} from Table 4, we conclude that $\eta_{*\mathrm{o}} = 0$. This means that

$$\zeta_{*\mathrm{o}}^{\mathrm{LMS}} = \zeta_{\mathrm{o}}^{\mathrm{LMS}},$$

where $\zeta_{*\mathrm{o}}^{\mathrm{LMS}}$ is the optimal EMSE of the combination, both with respect to the mixing parameters and the step sizes of the constituent filters. In other words, the smallest EMSE obtainable with a combination of two LMS algorithms is exactly equal to that obtained with a single LMS filter with optimum step size μ_{o}.

Just the same conclusion is obtained for a combination of two RLS filters. The only difference is that for two RLS filters, the derivative $\mathrm{d}\,\zeta_*^{\mathrm{RLS}}/\,\mathrm{d}\,\nu_2$ reads

$$\frac{\mathrm{d}\,\zeta_*^{\mathrm{RLS}}}{\mathrm{d}\,\nu_2} = -\frac{1}{2}\frac{(\mathrm{Tr}\{\boldsymbol{QR}\} - \nu_1^2\sigma_v^2 M)^2(\mathrm{Tr}\{\boldsymbol{QR}\} - \nu_2^2\sigma_v^2 M)}{(\nu_1\nu_2\sigma_v^2 M + \mathrm{Tr}\{\boldsymbol{QR}\})^2(\nu_1+\nu_2)^2},$$

and again we see that $\nu_2 = \nu_{\mathrm{o}}$ minimizes the overall steady-state error of the combination, and $\zeta_{*\mathrm{o}}^{\mathrm{RLS}} = \zeta_{\mathrm{o}}^{\mathrm{RLS}}$.

The above results allow us to conclude that, although a combination of two LMS (or two RLS) filters can improve the tracking performance when the degree of nonstationarity is not known *a priori* or time-varying, the steady-state EMSE of the combination is lower bounded by the optimal EMSE of an individual filter from the same family. In the next section, we will illustrate that this is not the case for a heterogeneous combination of one LMS and one RLS filters. Since for the combination of one RLS and one LMS filters the expressions become too complex for an analytical approach, we will proceed to show via examples that for this case it is possible to obtain an overall steady-state EMSE strictly smaller that the minimum of $\zeta_{\mathrm{o}}^{\mathrm{LMS}}$ and $\zeta_{\mathrm{o}}^{\mathrm{RLS}}$.

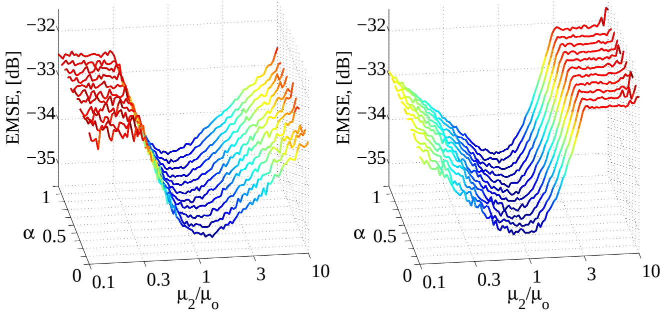

Fig. 1. Steady-state EMSE of a combination of two LMS filters for varying α and μ_2, and $\mu_1 = 0.3\mu_{\mathrm{o}}$ (left) and $\mu_1 = 3\mu_{\mathrm{o}}$ (right).

4. EXAMPLES

In this section we include several experiments for the identification of a time-varying system. Three sets of experiments have been considered: the first one consists of a combination of two LMS filters with different step sizes; two RLS filters with different forgetting factors are combined in the second group of simulations; and the last experiment implements a combination of one LMS and one RLS filters, both of them with optimal step sizes (i.e., μ_{o} and ν_{o}).

In all cases, the unknown plant $\boldsymbol{w}_{\mathrm{o}}$, of length $M = 7$, was initialized with random values from interval $[-1, 1]$, being $\boldsymbol{w}_{\mathrm{o}}(0) = [.9003, -.5377, .2137, -.028, .7826, .5242, -.0871]^T$. Then, the solution is changed at each iteration according to the random-walk model (3), with a covariance matrix of $\boldsymbol{q}(n)$ given by

$$\boldsymbol{Q} = \gamma\left[\alpha\frac{\boldsymbol{R}}{Tr(\boldsymbol{R})} + (1-\alpha)\frac{\boldsymbol{R}^{-1}}{Tr(\boldsymbol{R}^{-1})}\right], \qquad (7)$$

where constant γ has been selected to be $\gamma = 10^{-5}$, so that $\mathrm{Tr}(\boldsymbol{Q}) = \gamma$, and $\alpha \in [0, 1]$ is a control parameter that allows to tradeoff between a situation with $\boldsymbol{Q} \propto \boldsymbol{R}$ (for $\alpha = 1$), for which $\zeta_{\mathrm{o}}^{\mathrm{LMS}} < \zeta_{\mathrm{o}}^{\mathrm{RLS}}$, and $\boldsymbol{Q} \propto \boldsymbol{R}^{-1}$ ($\alpha = 0$), in which the reverse situation occurs.

The input signal is the output of a first-order AR model with transfer function $[1-a^2]/(1-az^{-1})$ using $a = 0.8$, fed with i.i.d. Gaussian noise with variance $\sigma_u^2 = \frac{1}{7}$, so that $\mathrm{Tr}(\boldsymbol{R}) = 1$. The output additive noise is i.i.d. Gaussian with zero-mean and variance $\sigma_v^2 = 10^{-2}$. Regarding the adjustment for the combinations, we have used convex combinations with fixed step size $\mu_a = 100$, while the step sizes of the constituent filters are selected as explained below.

All estimated steady-state EMSEs have been obtained by averaging 25000 runs of the algorithms once the filters have completely converged, and 100 independent runs.

To start with, we will consider the combination of two LMS filters. Note that in this case $\mu_{\mathrm{o}} = \sqrt{10^{-3}}$ independently of the value of α. Fig. 1 depicts the steady-state EMSE of the combination for different α and μ_2, and for two different values of the step size for the first component, $\mu_1 = 0.3\mu_{\mathrm{o}}$ and $\mu_1 = 3\mu_{\mathrm{o}}$. In both subfigures, we observe a flat region for which the combination inherits the performance of the filter with step size μ_1. As predicted by our analysis in the previous section,

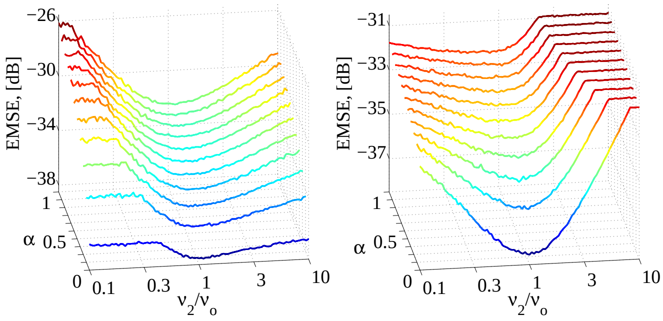

Fig. 2. Steady-state EMSE of a combination of two RLS filters for varying α and ν_2, and $\nu_1 = .001$ (left) and $\nu_1 = .0186$ (right).

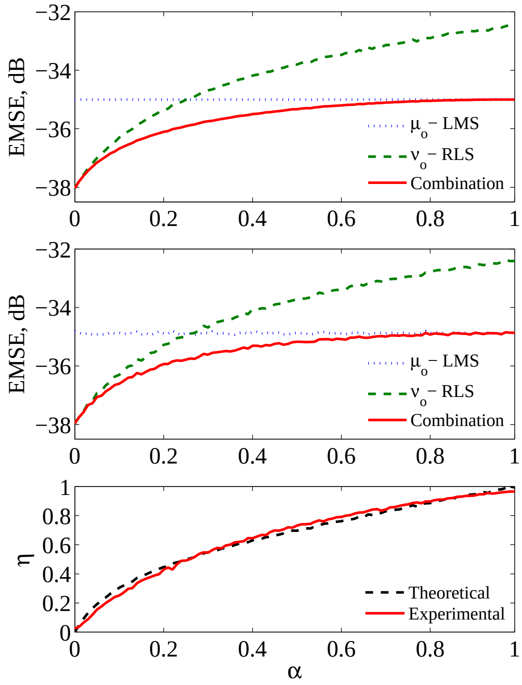

Fig. 3. Steady-state performance of an adaptive combination of one LMS and one RLS filters with optimal step sizes (μ_o and ν_o, respectively). The figure displays, from top to bottom, the theoretical EMSE of both constituent filters and of the combination for different values of α, the observed EMSEs obtained through simulation, and the theoretical and simulated steady-state values of $\eta(n)$.

the optimal behavior of the combination is observed when $\mu_2 = \mu_o$. Therefore, the combination performs in this situation similarly to an LMS filter with optimal step size.

Similar conclusions can be extracted for the convex combination of two RLS filters. Fig. 2 illustrates the behavior of such a combination scheme for different values of α and ν_2. In this case, ν_o changes with α. Therefore, in this situation we have selected $\nu_1 = 0.001$ (left panel of Fig. 2) and $\nu_1 = 0.0186$ (right panel), respectively smaller and larger than the optimal step sizes for $\alpha = 0$ and $\alpha = 1$. For each α, we explore values for the step size of the second filter in the range from $\nu_o/10$ to $10\nu_o$. Again, as predicted by the analysis, the best behavior results for $\nu_2 = \nu_o$, showing that $\zeta_{*o}^{\text{RLS}} = \zeta_o^{\text{RLS}}$.

We conclude the section considering the more interesting case of a combination including one LMS and one RLS filters. We will illustrate that it is possible for this heterogeneous combination to outperform the smallest EMSEs that could be achieved by any individual LMS or RLS filters. To this end, let us select for each α the optimal step sizes for the LMS and RLS components according to Table 3. The upper panel of Fig. 3 displays the theoretical steady-state EMSEs of both constituent filters and of their combination, and shows good agreement with the real curves obtained through simulation (intermediate panel). We can see that for all values of α other than $\alpha = 0$ and $\alpha = 1$, the combined LMS-RLS scheme reduces the individual EMSEs of both components, thus leading to the interesting result that $\zeta_{*o}^{\text{LMS-RLS}} < \min[\zeta_o^{\text{LMS}}, \zeta_o^{\text{RLS}}]$, i.e., this combined scheme is able to improve the tracking capabilities of optimal LMS and RLS filters.

It is also interesting to pay attention to the optimal values of the mixing parameter (bottom panel of Fig. 3). In first place, we see that for $\alpha \in [0, 1]$ the optimal mixing parameter lies in interval $[0, 1]$, i.e., affine combinations can be expected to work equally well –but not better– than convex combinations for the considered scenario. It is also important to remark that no gains over the tracking performance of optimal LMS or RLS can occur when $\boldsymbol{Q} \propto \boldsymbol{R}$ ($\alpha = 1$) or $\boldsymbol{Q} \propto \boldsymbol{R}^{-1}$ ($\alpha = 0$), since in these cases optimal selections of the mixing parameters are $\eta = 0$ and $\eta = 1$, respectively.

5. CONCLUSIONS

In this paper we have studied the tracking performance of combinations of LMS and RLS filters. We have provided theoretical and empirical evidence that the steady-state EMSE of a combination of two filters of the same family is lower bounded by the optimal EMSE of a filter of the same family. However, heterogeneous combinations of one LMS and one RLS filters have been shown to simultaneously reduce the EMSEs of both optimal LMS and RLS filters, thus providing a way to obtain filters with superior tracking capabilities.

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, no. 3, pp. 1078–1090, Mar. 2006.

[2] Y. Zhang and J. Chambers, "Convex combination of adaptive filters for a variable tap-length LMS algorithm," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 628–631, Oct. 2006.

[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, no. 5, pp. 1853–1864, May 2008.

[4] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, "Tracking properties of a convex combination of two adaptive filters," in *Proc. 2005 IEEE Workshop on Stat. Signal Process.*, Bourdeaux, France.

[5] M. T. M. Silva and V. H. Nascimento, "Improving the tracking capability of adaptive filters via convex combination," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3137–3149, July 2008.

[6] R. Candido, M. T. M. Silva, and V. H. Nascimento, "Affine combinations of adaptive filters," in *Conf. Rec. of the 42nd Asilomar Conf. on Sign., Syst. & Comp.*, 2008.

[7] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2937–2944, Nov. 1994.

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

[9] 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., ICASSP 2008*, Las Vegas, NV, USA, pp. 3301–3304.