STEADY STATE ANALYSIS OF AN OUTPUT SIGNAL BASED COMBINATION OF TWO NLMS ADAPTIVE FILTERS

Tõnu Trump

Department of Radio and Telecommunication Engineering
Tallinn University of Technology
Tallinn, Estonia
Email: ttrump@lr.ttu.ee

ABSTRACT

This paper studies an affine combination of two NLMS adaptive filters, which is an interesting way of improving the performance of adaptive algorithms. The structure consists of two adaptive filters that adapt on the same input signal, one with a large and the other one with a small step size. The outputs of the individual filters are combined together with help of a parameter λ. Such a combination is capable of achieving fast initial convergence and small steady state error at the same time. In this paper we propose to compute the combination parameter λ from output signals of the individual filters and investigate the steady state performance of the resulting combined algorithm.

1. INTRODUCTION

When designing an adaptive algorithm, one faces a trade–off between the initial convergence speed and the mean–square error in steady state. In case of algorithms belonging to the Least Mean Square (LMS) family this trade–off is controlled by the step-size parameter. Large step size leads to a fast initial convergence but the algorithm also exhibits a large mean–square error in the steady state and in contrary, small step size slows down the convergence but results in a small steady state error [1, 2].

Recently there has been an interest in a combination scheme that is able to optimize this trade–off [3]. The scheme consists of two adaptive filters that are simultaneously applied to the same inputs as depicted in Figure 1. One of the filters has a large step size allowing fast convergence and the other one has a small step size for small steady state error. The outputs of the filters are combined through a combination parameter λ . The performance of this scheme has been studied for some parameter update schemes [4, 5, 6, 7]. The references [4] and [5] use the ratio of time averages of the instantaneous errors of the filters. The error function of the ratio is then computed to obtain λ. The references [6] and [7] take another approach. The parameter λ is there found using an LMS type adaptive scheme and computing the sigmoidal function of the result.

In this paper we propose to compute the combination parameter λ from output signals of the individual filters and investigate the steady state performance of the resulting algorithm. The proposed way of calculating the combination parameter is optimal in the sense that it results from minimization of the mean-squared error of the combined filter.

We will assume throughout the paper that the signals are real–valued and that the combination scheme uses two normalized LMS adaptive filters. The italic, bold face lower case and bold face upper case letters will be used for scalars, column vectors and matrices respectively. The superscript T denotes transposition of a matrix. The operator $E[\cdot]$ stands for mathematical expectation.

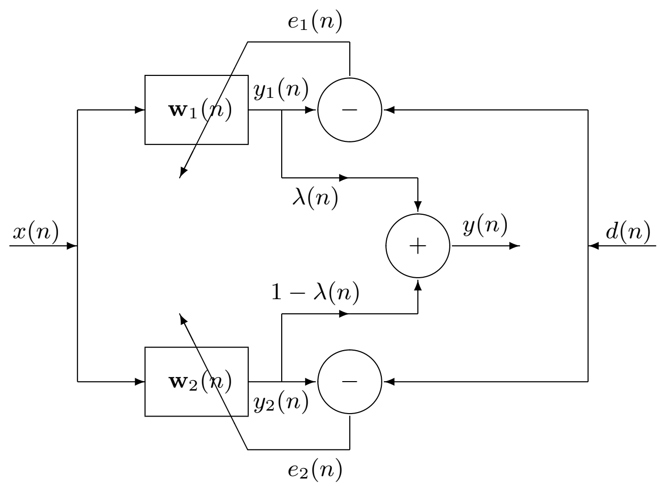

Figure 1: The combined adaptive filter.

2. DERIVATION

Let us consider two adaptive filters, as shown in Figure 1, each of them updated using the NLMS adaptation rule.

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \frac{\mu_i}{\mathbf{x}^T(n)\mathbf{x}(n)} e_i(n)\mathbf{x}(n), \quad i = 1, 2 \quad (1)$$

$$e_i(n) = d(n) - \mathbf{w}_i^T(n-1)\mathbf{x}(n), \quad (2)$$

$$d(n) = \mathbf{w}_o^T \mathbf{x}(n) + v(n). \quad (3)$$

In the above $\mathbf{w}_i(n)$ is the N vector of coefficients of the i-th adaptive filter, with $i = 1, 2$. The vector $\mathbf{w}_o$ is the true weight vector we aim to identify with our adaptive scheme. $\mathbf{x}(n)$ is the N input vector, common for both of the adaptive filters. The input process is assumed to be wide sense stationary. The desired signal $d(n)$ is a sum of the output of the filter to be identified and the measurement noise. The measurement noise is denoted by $v(n)$ and we assume this signal to be Gaussian, zero mean i.i.d. sequence, statistically independent of all the other signals. μ_i is the step size of i–th adaptive filter. We assume without loss of generality that $\mu_1 \geq \mu_2$.

The outputs of the two adaptive filters are combined according to

$$y(n) = \lambda(n)y_1(n) + [1 - \lambda(n)]y_2(n), \quad (4)$$

where $y_i(n) = \mathbf{w}_i^T(n-1)\mathbf{x}(n)$ and λ can be any real number. The output combination is, thus, affine as in [4], not convex as in [6, 7].

We define the *a priori* system error signal as difference between the output signal of the true system $y_o(n) = \mathbf{w}_o^T\mathbf{x}(n) = d(n) - v(n)$, and the output signal of our adaptive scheme $y(n)$

$$e_a(n) = y_o(n) - y(n) = y_o(n) - \lambda(n)y_1(n) - (1-\lambda(n))y_2(n). \tag{5}$$

Let us now find $\lambda(n)$ by minimizing the mean square of the *a priori* system error. The derivative of $E[e_a^2(n)]$ with respect to $\lambda(n)$ reads

$$\begin{aligned}\frac{\partial E[e_a^2(n)]}{\partial\lambda(n)} &= 2E[(y_o(n) - \lambda(n)y_1(n) \\ &\quad -(1-\lambda(n))y_2(n))(-y_1(n)+y_2(n))] \\ &= 2E[(y_o(n)-y_2(n))(y_2(n)-y_1(n)) \\ &\quad +\lambda(n)(y_2(n)-y_1(n))^2].\end{aligned}$$

Setting the derivative to zero results in

$$\lambda(n) = \frac{E[(d(n)-y_2(n))(y_1(n)-y_2(n))]}{E[(y_1(n)-y_2(n))^2]}, \tag{6}$$

where we have replaced the true system output signal $y_o(n)$ by its observable noisy version $d(n)$. Note however, that because we have made the standard assumption that the input signal $\mathbf{x}(n)$ and measurement noise $v(n)$ are independent random processes, this can be done without introducing any error into our calculations.

3. STEADY STATE PERFORMANCE

In this section we are interested in finding expressions that characterize performance of the combined algorithm in steady state i.e. when $n \to \infty$. Before we can proceed we need to introduce some notations. First let us define the weight error vector of i–th filter as

$$\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o - \mathbf{w}_i(n). \tag{7}$$

We then define the equivalent weight error vector of the combined adaptive filter to be

$$\tilde{\mathbf{w}}(n) = \lambda\tilde{\mathbf{w}}_1(n) + (1-\lambda)\tilde{\mathbf{w}}_2(n). \tag{8}$$

The *a priori* and *a posteriori* estimation errors of the individual filters are defined as

$$e_{i,a}(n) = \mathbf{x}^T(n)\tilde{\mathbf{w}}_i(n-1) \tag{9}$$

and

$$e_{i,p}(n) = \mathbf{x}^T(n)\tilde{\mathbf{w}}_i(n). \tag{10}$$

In what follows we often drop the explicit time index n, if it is not necessary to avoid a confusion.

Noting that $y_i(n) = \mathbf{w}_i^T(n)\mathbf{x}(n)$, we can rewrite (6) as

$$\lambda(n) = \frac{E[\tilde{\mathbf{w}}_2^T\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2] - E[\tilde{\mathbf{w}}_2^T\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_1]}{E[\tilde{\mathbf{w}}_1^T\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_1] - 2E[\tilde{\mathbf{w}}_1^T\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2] + E[\tilde{\mathbf{w}}_2^T\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2]}. \tag{11}$$

The above expression consists of expectations of the type $E[\tilde{\mathbf{w}}_i^T(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\tilde{\mathbf{w}}_j(n-1)]$. Because of the assumed wide sense stationarity of the input process we can replcace this expectation with $E[\tilde{\mathbf{w}}_i^T(n-1)\mathbf{x}(n-1)\mathbf{x}^T(n-1)\tilde{\mathbf{w}}_j(n-1)]$ and continue with investigation of that. For $E[\tilde{\mathbf{w}}_i^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\tilde{\mathbf{w}}_i(n)]$ we use the result from [1] stating that for normalized LMS

$$\begin{aligned}&E\left[\tilde{\mathbf{w}}_i^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_i(n)\right] \\ &= E\left[\tilde{\mathbf{w}}_i^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_i(n-1)\right](1-\mu_i)^2 + \mu_i^2\sigma_v^2.\end{aligned} \tag{12}$$

To evaluate the cross term we use the following relation between weight error vectors, *a priori* and *a posteriori* errors [1]

$$\tilde{\mathbf{w}}_i(n) + \frac{\mathbf{x}}{\mathbf{x}^T\mathbf{x}}e_{i,a}(n) = \tilde{\mathbf{w}}_i(n-1) + \frac{\mathbf{x}}{\mathbf{x}^T\mathbf{x}}e_{i,p}(n) \tag{13}$$

to obtain

$$\tilde{\mathbf{w}}_i(n) = \tilde{\mathbf{w}}_i(n-1) - \frac{\mathbf{x}}{\mathbf{x}^T\mathbf{x}}(e_{i,a}(n) - e_{i,p}(n)). \tag{14}$$

Hence we have

$$\begin{aligned}&\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n) \\ &= \tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) \\ &\quad -\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}(e_{2,a}(n)-e_{2,p}(n)) \\ &\quad -(e_{1,a}(n)-e_{1,p}(n))\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) \\ &\quad +(e_{1,a}(n)-e_{1,p}(n))(e_{2,a}(n)-e_{2,p}(n)).\end{aligned} \tag{15}$$

Substituting now the relationship between *a priori* and *a posteriori* errors for normalized LMS [1]

$$e_{i,p}(n) = e_{i,a}(n) - \mu_i e_i(n) \tag{16}$$

into the above we have

$$\begin{aligned}&\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n) \\ &= \tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) \\ &\quad -\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mu_2 e_2(n) \\ &\quad -\mu_1 e_1(n)\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) + \mu_1\mu_2 e_1(n)e_2(n).\end{aligned} \tag{17}$$

We now note that the error signal of i–th filter and its *a priori* error are related as

$$e_i(n) = e_{i,a}(n) + v(n) = \tilde{\mathbf{w}}_i^T(n-1)\mathbf{x} + v(n). \tag{18}$$

Substitution of this relation into the previous equation results in

$$\begin{aligned}&\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n) \\ &= \tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) \\ &\quad -\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mu_2[\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)+v(n)] \\ &\quad -\mu_1[\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}+v(n)]\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1) \\ &\quad +\mu_1\mu_2[\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)+v(n)][\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}+v(n)].\end{aligned} \tag{19}$$

And because $\mathbf{x}(n)$ and $v(n)$ are independent we have for the expectation

$$\begin{aligned}&E[\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n)] \\ &= (1-\mu_1)(1-\mu_2)E[\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)] \\ &\quad +\mu_1\mu_2 E[v^2(n)]\end{aligned} \tag{20}$$

In order to find an expression for $\lambda(n)$, we substitute (12) and (20) into (11) to obtain

$$\lambda(n) = \frac{\gamma(n-1)}{\xi(n-1)} \tag{21}$$

with

$$\begin{aligned}\gamma(n) &= (1-\mu_2)^2E[\tilde{\mathbf{w}}_2^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)] \\ &\quad -(1-\mu_1)(1-\mu_2)E[\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)] \\ &\quad +(\mu_2^2-\mu_1\mu_2)\sigma_v^2\end{aligned}$$

and

$$\begin{aligned}\xi(n) &= (1-\mu_1)^2 E[\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_1(n-1)] \\ &-2(1-\mu_1)(1-\mu_2)E[\tilde{\mathbf{w}}_1^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)] \\ &+(1-\mu_2)^2 E[\tilde{\mathbf{w}}_2^T(n-1)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n-1)] \\ &+(\mu_1-\mu_2)^2\sigma_v^2.\end{aligned}$$

Let us now find the steady state excess mean square error EMSE $= \lim_{n\to\infty} E[e_a^2(n)]$. To evaluate this we first note that equation (20) is in fact a difference equation in terms of $E[\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n)]$. Let us denote

$$s(n) = E[\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n)],$$

$$c_1 = \mu_1\mu_2\sigma_v^2$$

and

$$c_2 = (1-\mu_1)(1-\mu_2).$$

Then it is possible to rewrite equation (20) as

$$s(n) = c_2 s(n-1) + c_1 u(n),$$

where $u(n)$ is the unit step function. This equation can be solved using e.g. z–transform techniques [8]. Taking the z–transform of both sides of the above equation we have

$$S(z) = c_2 z^{-1} S(z) + \frac{c_2}{1-z^{-1}}. \tag{22}$$

Solving for $S(z)$ results in

$$S(z) = \frac{c_1}{(1-c_2 z^{-1})(1-z^{-1})}. \tag{23}$$

We now make use of the partial fraction expansion [8] to obtain

$$S(z) = \frac{c_1 c_2}{(1-c_2 z^{-1})(c_2-1)} + \frac{c_1}{(1-z^{-1})(1-c_2)}. \tag{24}$$

Calculating the inverse transform of $S(z)$ results in

$$s(n) = c_2^n \frac{c_1 c_2}{c_2-1} u(n) + \frac{c_1}{1-c_2} u(n). \tag{25}$$

The step size of the NLMS algorithm is usually selected to be less than one and consequently $c_2 = (1-\mu_1)(1-\mu_2) < 1$. This means that $s(n)$ is determined by the second term of the above equation when n approaches infinity. Hence we can write for the cross term in steady state

$$\lim_{n\to\infty} E[\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n)] = \frac{\mu_1\mu_2\sigma_v^2}{\mu_1+\mu_2-\mu_1\mu_2}. \tag{26}$$

Analogously we can obtain for the case of both $\mathbf{w_i}$ being the same

$$\lim_{n\to\infty} E[\tilde{\mathbf{w}}_i^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_i(n)] = \frac{\mu_i\sigma_v^2}{2-\mu_i}. \tag{27}$$

It follows from (5) that we can express the *a priori* error of the combination as

$$e_a(n) = \lambda(n) e_{1,a} + (1-\lambda(n))\, e_{2,a} \tag{28}$$

and because λ is according to (6) a ratio of mathematical expectations and, hence, deterministic, we have for the excess mean square error of the combination

$$E[e_a^2] = \lambda^2 E[e_{1,a}^2] + 2\lambda(1-\lambda)E[e_{1,a}e_{2,a}] + (1-\lambda)^2 E[e_{2,a}^2]. \tag{29}$$

We can now use the equations (26) and (27) in (29) to obtain for the excess mean square error of the combined algorithm

$$\begin{aligned}\text{EMSE} &= \lim_{n\to\infty} E[e_a^2] = \left[\frac{\lambda^2(\infty)\mu_1}{2-\mu_1}\right. \\ &+\frac{2\lambda(\infty)(1-\lambda(\infty))\mu_1\mu_2}{\mu_1+\mu_2-\mu_1\mu_2} \\ &+\left.\frac{(1-\lambda(\infty))^2\mu_2}{2-\mu_2}\right]\sigma_v^2.\end{aligned} \tag{30}$$

Let us proceed finding an expression for mean square deviation MSD $= \lim_{n\to\infty} E\|\tilde{\mathbf{w}}_i(n)\|^2$. To do so we need to invoke the independence theory [1, 2]. In particular let us assume that the sequence $\mathbf{x}(n)$ is independent and identically distributed. It then follows that weights computed at step $n-1$, $\mathbf{w}(n)$ are independent of the input at stage n, $\mathbf{x}(n)$. We can then interpret the expectations as conditional expectations, conditioned on the weight error vectors and write

$$E[\tilde{\mathbf{w}}_1^T(n)\mathbf{x}\mathbf{x}^T\tilde{\mathbf{w}}_2(n)] = \tilde{\mathbf{w}}_1^T(n)\mathbf{R_x}\tilde{\mathbf{w}}_2(n), \tag{31}$$

where $\mathbf{R_x} = E[\mathbf{x}\mathbf{x}^T]$ is the input signal correlation matrix. To proceed we need to assume that the input signal $\mathbf{x}$ is white $\mathbf{R_x} = \sigma_x^2\mathbf{I}$, and in this case

$$\begin{aligned}\text{MSD} &= \lim_{n\to\infty} E[\tilde{\mathbf{w}}^T(n)\tilde{\mathbf{w}}(n)] = \left[\frac{\lambda^2(\infty)\mu_1}{2-\mu_1}\right. \\ &+\frac{2\lambda(\infty)(1-\lambda(\infty))\mu_1\mu_2}{\mu_1+\mu_2-\mu_1\mu_2} \\ &+\left.\frac{(1-\lambda(\infty))^2\mu_2}{2-\mu_2}\right]\frac{\sigma_v^2}{\sigma_x^2}.\end{aligned} \tag{32}$$

It remains to evaluate the limiting value of $\lambda(n)$ as n approaches infinity, $\lambda(\infty)$. This can be done by substituting equations (26) and (27) into (21) which results in

$$\lambda(\infty) = \lim_{n\to\infty} E[\lambda(n)] = \frac{\eta}{\kappa} \tag{33}$$

with

$$\begin{aligned}\eta &= [(1-\mu_2)^2\mu_2\rho \\ &-(1-\mu_1)(1-\mu_2)\mu_1\mu_2(2-\mu_2) \\ &+(\mu_2^2-\mu_1\mu_2)(2-\mu_2)\rho](2-\mu_1)\end{aligned}$$

and

$$\begin{aligned}\kappa &= (1-\mu_1)^2\mu_1(2-\mu_2)\rho \\ &-2(1-\mu_1)(1-\mu_2)\mu_1\mu_2(2-\mu_1)(2-\mu_2) \\ &+(1-\mu_2)^2\mu_2(2-\mu_1)\rho \\ &+(\mu_1-\mu_2)^2(2-\mu_1)\rho(2-\mu_2),\end{aligned}$$

where $\rho = \mu_1+\mu_2-\mu_1\mu_2$.

It is well known that NLMS achieves the fastest initial convergence if the step size is selected to equal unity. By this reason it makes sense to choose the step size of the fast converging filter $\mu_1 = 1$ and in this case the above simplifies to

$$\lambda(\infty) = \lim_{n\to\infty} E[\lambda(n)] = \frac{\mu_2}{2(\mu_2-1)}. \tag{34}$$

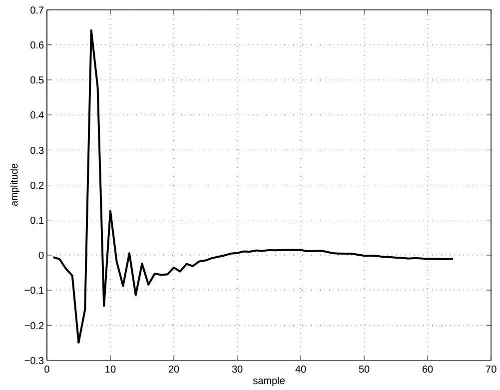

Figure 2: True impulse response used in the simulations.

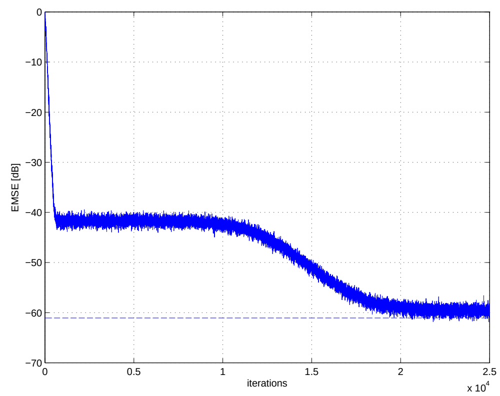

Figure 3: Time–evolutions of MSE.

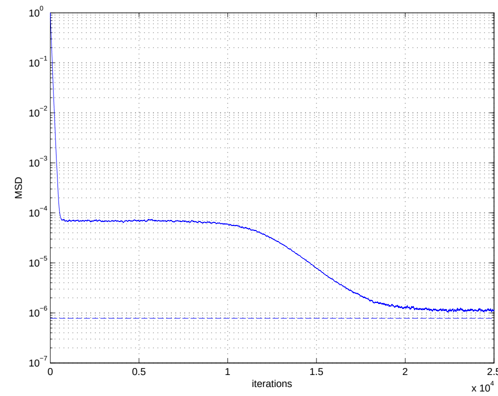

Figure 4: Time–evolutions of MSD.

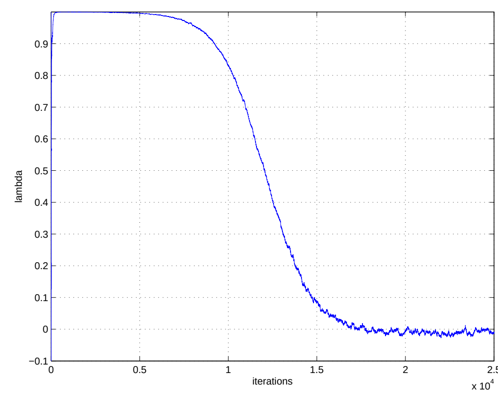

Figure 5: Time–evolutions of the combination parameter λ.

4. SIMULATION RESULTS

A simulation study was carried out with the aim of verifying the approximations made in the previous Section.

We have combined two 64 tap long adaptive filters with $\mu_1 = 1$ and $\mu_2 = 0.025$. We have assumed that signals are ergodic and, hence, in order to obtain a practical algorithm, the expectation operators in (6) have been replaced by exponential averaging of the type

$$P_u(n) = (1-\gamma)P_u(n-1) + \gamma u^2(n), \tag{35}$$

where $u(n)$ is the signal to be averaged, $P_u(n)$ is the averaged quantity and $\gamma = 0.01$.

We have selected the sample echo path model number one, shown in Figure 2 from [9], to be the unknown system to identify.

White noise with variance $\sigma_x^2 = 1$ was used as the input signal and another white noise, with variance $\sigma_v^2 = 6 \cdot 10^{-6}$, statistically independent of x as the measurement noise. The curves are averaged over 100 trials.

Figure 3 shows with solid line the evolution of excess mean–squared error $E[e_a^2]$ in time. One can observe that the adaptive algorithm converges fast in the beginning and undergoes a second convergence around sample 15000. The second convergence occurs when the filter with small step size becomes more accurate than the one with large step size. The theoretical steady state EMSE, computed from (30) is shown with dashed line. One can see that the experimental line approaches the theoretical steady state result.

Figure 4 depicts the corresponding time evolution of the mean square deviation MSD $= E[\|\tilde{\mathbf{w}}(n)\|^2]$. Again, as with the MSE curve we see a fast initial convergence followed by a second convergence around sample–time 15000. The simulation result is shown with solid line and the theoretical steady state MSD, computed form (32) with dashed line.

Figure 5 shows the evolution of the combination parameter λ. It can be seen that at the beginning, when the fast converging filter is better than the slower one, λ is close to one. When the filter with small step size becomes better than the fast one, λ decreases and eventually becomes a small negative number. This is because we have two dependent estimators of the same unknown system and we select parameter λ by minimizing the mean square error of the combination.

Dependence of the excess mean square error from the step size of the more accurate filter is depicted in Figure 6. Statistically independent white Gaussian noises were used as the input signal and the measurement noise. The input signal power was $\sigma_x^2 = 1$ the noise power $\sigma_v^2 = 10^{-4}$ and the step size, μ_2, was varied from 0.02 to 0.38. The circles are the simulation results after 32000 iterations and the solid line is the EMSE computed from (30). The simulation results are averaged over 100 trials. One can see that there is a reasonable fit between the theory and simulations.

Figure 7 depicts the dependence of the steady state mean square deviation, $\lim_{n\to\infty} E\|\tilde{\mathbf{w}}_i(n)\|^2$, of the algorithm as a function of μ_2. All the conditions are the same as in the previous Figure. The circles are the simulation results after 32000 iterations and the solid line is the theoretical MSD computed from (32). Again, one can observe that the theory matches the simulation results well.

Finally we compare the different possibilities of computing the parameter λ in the Figure 8. The curves here are averaged over 500 independent trials. Here $\mu_1 = 1$ and $\mu_2 = 0.1$

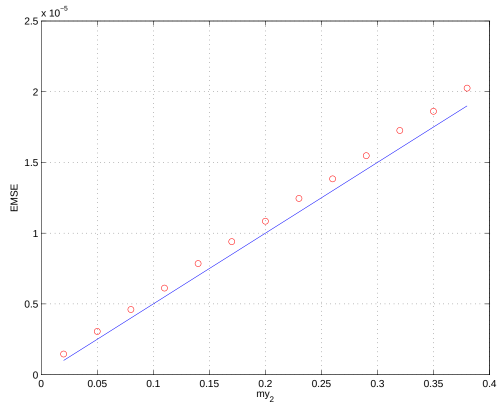

Figure 6: Excess mean square error as a function of μ_2.

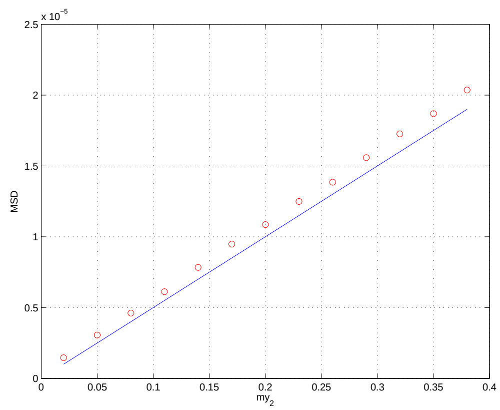

Figure 7: Mean square deviation of filter coefficients as a function of μ_2.

in order to make the differences between the different algorithms more visible. The solid red curve is the λ used in this paper. The dashed blue line is the λ from [4].

$$\lambda(n) = 1 - \kappa \mathrm{erf}\left(\frac{\hat{e}_1^2(n)}{\hat{e}_1^2(n)}\right), \tag{36}$$

where $\mathrm{erf}(x) = \frac{2}{\pi}\int_0^{\pi} e^{-t^2/2}dt$ is the error function, $\hat{e}_i^2(n) = \frac{1}{K}\sum_{m=n-K+1}^{n} e_i^2(m)$, where $K = 100$ as suggested in [4] and $\kappa = 1 - \frac{\mu_2/\mu_1}{2(\mu_2/\mu_1 - 1)}$.

The dashed – dotted green line is the λ from [7], computed using the normalized LMS like adaptation scheme.

$$\lambda(n) = \frac{1}{1 + e^{-a(n)}}, \tag{37}$$

where

$$a(n+1) = a(n) + \frac{\mu_a}{p(n)}\lambda(n)[1-\lambda(n)]e(n)[e_2(n) - e_1(n)], \tag{38}$$

and $p(n) = 0.9p(n-1) + 0.1[e_2(n) - e_1(n)]^2$. One can see that the curves corresponding to the algorithm investigated in this paper and the one from [4] converge to a negative value of λ in steady state. The transition region of the algorithm of this paper occurs slightly later than that of the algorithm from [4]. The LMS like algorithm from [7] has its transition region approximately simultaneously with the algorithm investigated in this paper but λ from [7] remains positive in steady state.

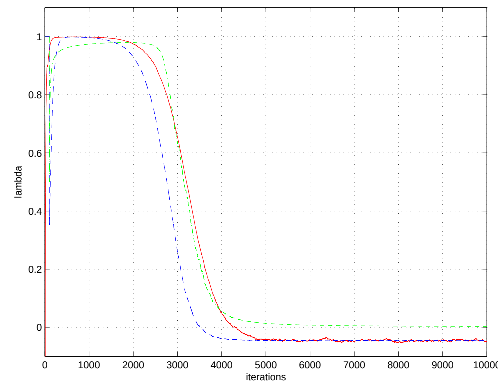

Figure 8: Evolution of λ computed using three different algorithms as a function of time.

5. CONCLUSIONS

In this paper we have investigated an affine combination of two adaptive filters that are simultaneously applied to the same input signals. It was proposed to compute the combination parameter λ using the output signals of the individual filters and the desired signal. The steady state performance of the algorithm was investigated and expressions for steady state excess mean square error and mean square deviation were derived. Finally it was shown in the simulation study that the derived formulae fit the simulation results well.

REFERENCES

[1] A. H. Sayed, *"Fundamentals of Adaptive Filtering,"* John Wiley and sons, 2003.

[2] S. Haykin, *"Adaptive Filter Theory, Fourth Edition,"* Prentice Hall, 2002.

[3] M. Martinez–Ramon, J. Arenas-Garcia, A. Navia–Vazquez and A. R. Figueiras-Vidal "An Adaptive Combination of Adaptive Filters for Plant Identification," *In Proc. 14th International Conference on Digital Signal Processing,* Santorini, Greece, 2002, pp. 1195–1198.

[4] N. J. Bershad, J. C. Bermudez and J. H. Tourneret, "An Affine Combination of Two LMS Adaptive Filters – Transient Mean–Square Analysis," *IEEE Trans. on Signal Processing,* Vol. 56, May 2008, pp. 1853–1864.

[5] N. J. Bershad, J. C. Bermudez and J. H. Tourneret, "On Performance Bounds for an Affine Combination of Two LMS Adaptive Filters," *In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing,* Las Vegas, Nevada, pp. 3297–3100, March–April, 2008.

[6] J. Arenas-Garcia, A. R. Figueiras-Vidal and A. H. Sayed, "Mean-Square Performance of Convex Combination of Two Adaptive Filters," *IEEE Trans. on Signal Processing,* Vol. 54, March 2006, pp. 1078–1090.

[7] L. A. Azpicueta–Ruiz, A. R. Figueiras-Vidal and J. Arenas-Garcia, "A Normalized Adaptation Scheme for the Convex Combination of Two Adaptive Filters," *In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing,* Las Vegas, Nevada, pp. 3301–3304, March–April, 2008.

[8] A. V. Oppenheim, R. W. Schafer *"Discrete–Time Signal Processing,"* Prentice Hall, 1989.

[9] ITU-T Recommendation G.168 "Digital Network Echo Cancellers," ITU-T, 2007.