



Output signal based combination of two NLMS adaptive filters – transient analysis

Tõnu Trump

Department of Radio and Telecommunication Engineering, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia; tonu.trump@lr.ttu.ee

Received 11 October 2010, accepted 20 December 2010

Abstract. A combination of two complex normalized least mean square (NLMS) adaptive filters that adapt on the same input signal at the same time is investigated. One of the filters has a large and the other one has a small step size. The outputs of the filters are combined together through a mixing parameter λ . This combination is an interesting new way of achieving simultaneously a fast initial convergence and a small steady state error of an adaptive algorithm. The mixing parameter is computed from the output signals of the individual filters. The expressions characterizing the time evolution of the mean square deviation and the excess mean square error of the combination scheme are derived. The theoretical results are verified by simulations.

Key words: adaptive algorithms, combination, analysis.

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. A large step size leads to a fast initial convergence but the algorithm also exhibits a large mean-square error in the steady state, and on the contrary, a small step size slows down the convergence but results in a small steady state error [1,2].

Variable step size adaptive schemes offer a possible solution allowing achieving both fast initial convergence and low steady state misadjustment [3–7]. How successful these schemes are depends on how well the algorithm is able to estimate the distance of the adaptive filter weights from the optimal solution. The variable step size algorithms use different criteria for calculating the proper step size at any given time instance. For example, squared instantaneous errors have been used in [4] and the squared autocorrelation of errors at adjacent time instances have been used in [6]. Paper [5] investigates an algorithm that changes the time-varying convergence parameters in such a way that the change is proportional to the negative of the gradient of the squared estimation error with respect to the convergence parameter. In [7] the norm of the projected weight error vector is used as a criterion to determine how close the adaptive filter is to its optimum performance.

Recently there has been an interest in a combination scheme that is able to optimize the trade-off between the convergence speed and the steady state error [8]. 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 a small steady state error. The outputs of the filters are combined through a mixing parameter λ . The performance of this scheme has been studied

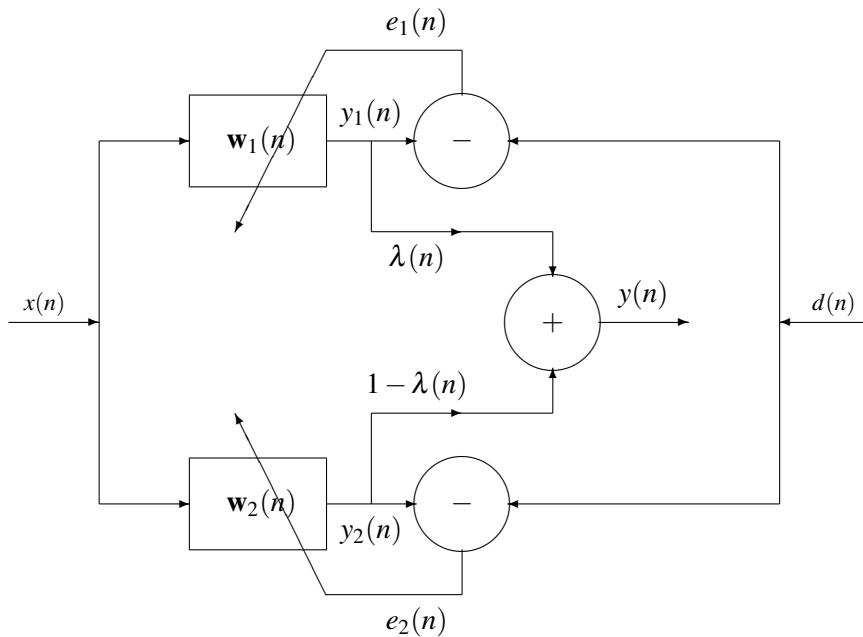


Fig. 1. The combined adaptive filter.

for some parameter update schemes [9–11]. Paper [9] uses convex combination, i.e., λ is constrained to lie between 0 and 1. Paper [10] presents a transient analysis of a slightly modified version of this scheme. The parameter λ is in those studies found by using an LMS-type adaptive scheme and computing the sigmoidal function of the result. Another approach by computing the mixing parameter using an affine combination is provided in [11]. This paper uses the ratio of time averages of the instantaneous errors of the filters. The error function of the ratio is computed to obtain λ .

In [12] a convex combination of two adaptive filters with different adaptation schemes has been investigated with the aim of improving the steady state characteristics. One of the adaptive filters uses the LMS algorithm and the other one the Generalized Normalized Gradient Decent algorithm. The combination parameter λ is computed, using stochastic gradient adaptation. In [13] the convex combination of two adaptive filters is applied in a variable filter length scheme to gain improvements in low signal to noise ratio conditions. In [14] the combination has been used to join two affine projection filters with different regularization parameters. Paper [15] uses the combination on parallel binary structured LMS algorithms. These three works use the LMS-like scheme of [16] to compute λ .

It should be noted that schemes involving two filters have been proposed earlier [17,18]. However, in those early schemes only one of the filters was adaptive, while the other used fixed filter weights. The updating of the fixed filter was accomplished by copying all coefficients from the adaptive filter, when the adaptive filter was performing better than the fixed one.

In the present paper the mixing parameter λ is computed from output signals of the individual filters. The way of calculating the mixing parameter is optimal in the sense that it results from minimization of the mean-squared error of the combined filter. The scheme was independently proposed in [19] and [20]; the steady state performance of it was investigated in [21] and the tracking performance in [22]. In [23] the output signal based combination was used in the adaptive line enhancer. Those papers investigate the behaviour of the adaptive combination scheme in steady state, i.e., in the situation when discrete time n approaches infinity. In the main body of this paper, on the other hand, a transient analysis of the algorithm is given, i.e., the formulae, predicting the entire course of adaptation, not only steady state, are derived.

It will be assumed throughout the paper that the signals are complex-valued and the combination scheme uses two normalized LMS (NLMS) adaptive filters. The italic, boldface lower-case, and boldface upper-

case letters will be used for scalars, column vectors, and matrices, respectively. The superscript T denotes transposition of a matrix, the operator $E[\cdot]$ denotes mathematical expectation, and $tr[\cdot]$ stands for trace of a matrix.

2. ALGORITHM

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

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

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

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

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 and $\mathbf{x}(n)$ is the known N input vector, common for both adaptive filters. The input process is assumed to be a zero mean wide sense stationary Gaussian process. The desired signal $d(n)$ is a sum of the output of the filter to be identified and the Gaussian, zero mean independent an identically distributed (i.i.d.) measurement noise that is statistically independent of all other signals. The step size of the i th adaptive filter is denoted by μ_i . We assume without loss of generality that $\mu_1 > \mu_2$. The case $\mu_1 = \mu_2$ is not interesting because the two filters remain equal and the combination renders to a single filter.

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 $y_i(n) = \mathbf{w}_i^H(n-1)\mathbf{x}(n)$ and the mixing parameter λ can be any real number.

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

$$e_a(n) = y_o(n) - \lambda(n)y_1(n) - (1 - \lambda(n))y_2(n). \quad (5)$$

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

$$\frac{\partial E[|e_a(n)|^2]}{\partial \lambda(n)} = 2E[\Re\{e\{(y_o(n) - y_2(n))(y_1(n) - y_2(n))^*\} + \lambda(n)|(y_2(n) - y_1(n))|^2\}]. \quad (6)$$

Setting the derivative to zero results in

$$\lambda(n) = \frac{E[\Re\{(d(n) - y_2(n))(y_1(n) - y_2(n))^*\}]}{E[|(y_1(n) - y_2(n))|^2]}, \quad (7)$$

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.

The denominator of equation (7) comprises expectation of the squared difference of the two filter output signals. This quantity can be very small or even zero, particularly at the beginning of adaptation if the two step sizes are close to each other. Correspondingly λ computed directly from (7) may be large. To avoid this from happening, we add a small regularization constant ε to the denominator of (7). The constant ε should be selected small compared to $E[\mathbf{x}^T(n)\mathbf{x}(n)]$ but large enough to prevent division by zero in given arithmetic.

3. TRANSIENT ANALYSIS

In this section we are interested in finding expressions that characterize transient performance of the combined algorithm, i.e., we intend to derive formulae that characterize the entire course of adaptation of the algorithm. Before we can proceed we need, however, to introduce some notations. First, let us denote the weight error vector of the i th filter as

$$\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o - \mathbf{w}_i(n). \quad (8)$$

Then the equivalent weight error vector of the combined adaptive filter will be

$$\tilde{\mathbf{w}}(n) = \lambda \tilde{\mathbf{w}}_1(n) + (1 - \lambda) \tilde{\mathbf{w}}_2(n). \quad (9)$$

The mean square deviation (MSD) of the combined filter is given by

$$\begin{aligned} \text{MSD} &= E[\tilde{\mathbf{w}}^H(n)\tilde{\mathbf{w}}(n)] = \lambda^2 E[\tilde{\mathbf{w}}_1^H(n)\tilde{\mathbf{w}}_1(n)] \\ &\quad + 2\lambda(1 - \lambda) \mathcal{R}e\{E[\tilde{\mathbf{w}}_2^H(n)\tilde{\mathbf{w}}_1(n)]\} \\ &\quad + (1 - \lambda)^2 E[\tilde{\mathbf{w}}_2^H(n)\tilde{\mathbf{w}}_2(n)]. \end{aligned} \quad (10)$$

The *a priori* estimation error of an individual filter is defined as

$$e_{i,a}(n) = \tilde{\mathbf{w}}_i^H(n-1)\mathbf{x}(n). \quad (11)$$

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

$$e_a(n) = \lambda(n)e_{1,a}(n) + (1 - \lambda(n))e_{2,a}(n) \quad (12)$$

and because $\lambda(n)$ is, according to (7), a ratio of mathematical expectations and, hence, deterministic, we have for the excess mean square error (EMSE) of the combination

$$\begin{aligned} E[|e_a(n)|^2] &= \lambda^2 E[|e_{1,a}(n)|^2] \\ &\quad + 2\lambda(1 - \lambda) E[\mathcal{R}e\{e_{1,a}(n)e_{2,a}^*(n)\}] \\ &\quad + (1 - \lambda)^2 E[|e_{2,a}(n)|^2]. \end{aligned} \quad (13)$$

As $e_{i,a}(n) = \tilde{\mathbf{w}}_i^H(n-1)\mathbf{x}(n)$, the expression of the EMSE becomes

$$\begin{aligned} E[|e_a(n)|^2] &= \lambda^2 E[\tilde{\mathbf{w}}_1^H(n-1)\mathbf{xx}^H\tilde{\mathbf{w}}_1(n-1)] \\ &\quad + 2\lambda(1 - \lambda) E[\mathcal{R}e\{\tilde{\mathbf{w}}_1^H(n-1)\mathbf{xx}^H\tilde{\mathbf{w}}_2(n-1)\}] \\ &\quad + (1 - \lambda)^2 E[\tilde{\mathbf{w}}_2^H(n-1)\mathbf{xx}^H\tilde{\mathbf{w}}_2(n-1)]. \end{aligned} \quad (14)$$

In what follows we often drop the explicit time index n as we have done in (14), if it is not necessary to avoid a confusion.

Noting that $y_i(n) = \mathbf{w}_i^H(n-1)\mathbf{x}(n)$, we can rewrite the expression for $\lambda(n)$ in (7) as

$$\lambda(n) = \frac{E[\tilde{\mathbf{w}}_2^H\mathbf{xx}^H\tilde{\mathbf{w}}_2] - E[\mathcal{R}e\{\tilde{\mathbf{w}}_2^H\mathbf{xx}^H\tilde{\mathbf{w}}_1\}]}{E[\tilde{\mathbf{w}}_1^H\mathbf{xx}^H\tilde{\mathbf{w}}_1] - 2E[\mathcal{R}e\{\tilde{\mathbf{w}}_1^H\mathbf{xx}^H\tilde{\mathbf{w}}_2\}] + E[\tilde{\mathbf{w}}_2^H\mathbf{xx}^H\tilde{\mathbf{w}}_2]}. \quad (15)$$

We thus need to investigate the evolution of terms of the type $\text{EMSE}_{k,l} = E[\tilde{\mathbf{w}}_k^H(n-1)\mathbf{x}(n)\mathbf{x}^H(n)\tilde{\mathbf{w}}_l(n-1)]$ in order to reveal the time evolution of $\text{EMSE}(n)$ and $\lambda(n)$. To do so, we concentrate first on the mean square deviation defined in (10).

For a single NLMS filter we have after subtraction of (1) from \mathbf{w}_o and expressing $e_i(n)$ through the error of the corresponding Wiener filter $e_o(n)$

$$\tilde{\mathbf{w}}_i(n) = \left(\mathbf{I} - \frac{\mu}{\mathbf{x}^H \mathbf{x}} \mathbf{x} \mathbf{x}^H \right) \tilde{\mathbf{w}}_i(n-1) - \frac{\mu}{\mathbf{x}^H \mathbf{x}} \mathbf{x} e_o^*(n). \quad (16)$$

At this point we make two approximations. First, we approximate the outer product of input signal vectors by its correlation matrix $\mathbf{x} \mathbf{x}^H \approx \mathbf{R}_x$. Second, we approximate the inner product of the input signal vectors by N times of its power $\mathbf{x}^H \mathbf{x} \approx N \sigma_x^2$. With those approximations we have

$$\tilde{\mathbf{w}}(n) \approx \left(\mathbf{I} - \frac{\mu}{N \sigma_x^2} \mathbf{R}_x \right) \tilde{\mathbf{w}}(n-1) - \frac{\mu}{N \sigma_x^2} \mathbf{x} e_o^*(n). \quad (17)$$

This means, in fact, that we apply the small step size theory [2] even if the assumption of a small step size is not really true for the fast adapting filter. In our simulation study we will see, however, that the assumption works rather well in practice.

Let us now define the eigendecomposition of the correlation matrix as

$$\mathbf{Q}^H \mathbf{R}_x \mathbf{Q} = \mathbf{\Omega}, \quad (18)$$

where \mathbf{Q} is a unitary matrix whose columns are the orthogonal eigenvectors of \mathbf{R}_x and $\mathbf{\Omega}$ is a diagonal matrix having eigenvalues associated with the corresponding eigenvectors on its main diagonal. We also define the transformed weight error vector as

$$\mathbf{v}(n) = \mathbf{Q}^H \tilde{\mathbf{w}}(n) \quad (19)$$

and the transformed last term of equation (17) as

$$\mathbf{p}(n) = \frac{\mu}{N \sigma_x^2} \mathbf{Q}^H \mathbf{x} e_o^*(n). \quad (20)$$

Then we can rewrite equation (17) after multiplying both sides by \mathbf{Q}^H from the left as

$$\mathbf{v}(n) = \left(\mathbf{I} - \frac{\mu}{N \sigma_x^2} \mathbf{\Omega} \right) \mathbf{v}(n-1) - \mathbf{p}(n). \quad (21)$$

We note that the mean of \mathbf{p} is zero by the orthogonality theorem and its correlation matrix equals

$$E[\mathbf{p} \mathbf{p}^H] = \frac{\mu^2}{N^2 \sigma_x^4} \mathbf{Q}^H E[\mathbf{x} e_o^*(n) e_o(n) \mathbf{x}^H] \mathbf{Q}. \quad (22)$$

We now invoke the Gaussian moment factoring theorem to write

$$E[\mathbf{x} e_o^*(n) e_o(n) \mathbf{x}^H] = E[\mathbf{x} e_o^*(n)] E[e_o(n) \mathbf{x}^H] + E[\mathbf{x} \mathbf{x}^H] E[|e_o|^2]. \quad (23)$$

The first term in the above is zero due to the principle of orthogonality and the second term equals $\mathbf{R} J_{\min}$. Hence we are left with

$$E[\mathbf{p} \mathbf{p}^H] = \frac{\mu^2}{N^2 \sigma_x^4} J_{\min} \mathbf{\Omega}, \quad (24)$$

where $J_{\min} = E[|e_o|^2]$ is the minimum mean square error produced by the corresponding Wiener filter. As the matrices \mathbf{I} and $\mathbf{\Omega}$ in (21) are diagonal, it follows that the m th element of vector $\mathbf{v}(n)$ is given by

$$\begin{aligned} v_m(n) &= \left(1 - \frac{\mu}{N \sigma_x^2} \omega_m \right) v_m(n-1) - p_m(n) \\ &= \left(1 - \frac{\mu}{N \sigma_x^2} \omega_m \right)^n v_m(0) + \sum_{i=0}^{n-1} \left(1 - \frac{\mu}{N \sigma_x^2} \omega_m \right)^{n-1-i} p_m(i), \end{aligned} \quad (25)$$

where ω_m is the m th eigenvalue of \mathbf{R}_x and v_m and p_m are the m th components of the vectors \mathbf{v} and \mathbf{p} , respectively.

We can now express the MSD and its individual components in (10) through the transformed weight error vectors as

$$E[\tilde{\mathbf{w}}^H(n)\tilde{\mathbf{w}}(n)] = E[\mathbf{v}^H(n)\mathbf{v}(n)] = \sum_{m=0}^{N-1} E[v_m(n)v_m^*(n)]. \quad (26)$$

Let us concentrate on the m th component in the sum above corresponding to the cross term and denote it as $\Upsilon_m = E[v_{k,m}(n)v_{l,m}^*(n)]$. The expressions for the component filters follow as special cases. Substituting (25) into the expression of Υ_m above, taking the mathematical expectation, and noting that the vector \mathbf{p} is independent of $\mathbf{v}(0)$ results in

$$\begin{aligned} \Upsilon_m &= E \left[\left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^n v_k(0) \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^n v_l^*(0) \right] \\ &+ E \left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{n-1-i} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{n-1-j} p_{k,m}(i)p_{l,m}^*(j) \right]. \end{aligned} \quad (27)$$

We now note that most likely the two component filters are initialized to the same value

$$v_{k,m}(0) = v_{l,m}(0) = v_m(0) \quad (28)$$

and that

$$E[p_{k,m}(i)p_{l,m}^*(j)] = \begin{cases} \frac{\mu_k \mu_l}{N^2 \sigma_x^4} \omega_m J_{\min}, & i = j \\ 0, & \text{otherwise} \end{cases} \quad (29)$$

We then have for the m th component of MSD

$$\begin{aligned} \Upsilon_m &= \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^n \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^n |v_m(0)|^2 \\ &+ \frac{\mu_k \mu_l}{N^2 \sigma_x^4} \omega_m J_{\min} \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{n-1} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{n-1} \sum_{i=0}^{n-1} \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{-i} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{-i}. \end{aligned} \quad (30)$$

The sum over i above can be recognized as a geometric series with n terms. The first term is equal to 1 and the geometric ratio equals $\left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{-1} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{-1}$. Hence we have

$$\begin{aligned} \sum_{i=0}^{n-1} \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{-i} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{-i} &= \frac{\left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right) \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)}{\frac{\mu_k \mu_l}{N^2 \sigma_x^4} \omega_m^2 - \frac{\mu_k}{N\sigma_x^2} \omega_m - \frac{\mu_l}{N\sigma_x^2} \omega_m} \\ &- \frac{\left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^{-n+1} \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^{-n+1}}{\frac{\mu_k \mu_l}{N^2 \sigma_x^4} \omega_m^2 - \frac{\mu_k}{N\sigma_x^2} \omega_m - \frac{\mu_l}{N\sigma_x^2} \omega_m}. \end{aligned} \quad (31)$$

After substitution of the above into (30) and simplification we are left with

$$\begin{aligned} \Upsilon_m &= E[v_{k,m}(n)v_{l,m}^*(n)] \\ &= \left(1 - \frac{\mu_k}{N\sigma_x^2} \omega_m\right)^n \left(1 - \frac{\mu_l}{N\sigma_x^2} \omega_m\right)^n \left[|v_m(0)|^2 + \frac{J_{\min}}{\omega_m^2 - \frac{N\sigma_x^2 \omega_m}{\mu_l} - \frac{N\sigma_x^2 \omega_m}{\mu_k}} \right] \\ &- \frac{J_{\min}}{\omega_m^2 - \frac{N\sigma_x^2 \omega_m}{\mu_l} - \frac{N\sigma_x^2 \omega_m}{\mu_k}}, \end{aligned} \quad (32)$$

which is our result for a single entry to the MSD cross term vector. It is easy to see that for the terms involving a single filter we get expressions that coincide with the one available in the literature [2].

Let us now focus on the cross term

$$\text{EMSE}_{k,l} = E [\tilde{\mathbf{w}}_k^H(n-1)\mathbf{x}(n)\mathbf{x}^H(n)\tilde{\mathbf{w}}_l(n-1)], \quad (33)$$

appearing in the EMSE equation (14). Due to the independence assumption we can rewrite this by using the properties of the trace operator as

$$\begin{aligned} \text{EMSE}_{k,l} &= E [\tilde{\mathbf{w}}_k^H(n-1)\mathbf{R}_x\tilde{\mathbf{w}}_l(n-1)] \\ &= \text{tr} \{E [\mathbf{R}_x\tilde{\mathbf{w}}_l(n-1)\tilde{\mathbf{w}}_k^H(n-1)]\} \\ &= \text{tr} \{\mathbf{R}_xE [\tilde{\mathbf{w}}_l(n-1)\tilde{\mathbf{w}}_k^H(n-1)]\}. \end{aligned} \quad (34)$$

Let us now recall that for any of the filters $\tilde{\mathbf{w}}_i(n) = \mathbf{Q}\mathbf{v}_i(n)$ to write

$$\begin{aligned} \text{EMSE}_{k,l} &= \text{tr} \{\mathbf{R}_xE [\mathbf{Q}\mathbf{v}_l(n-1)\mathbf{v}_k^H(n-1)\mathbf{Q}^H]\} \\ &= \text{tr} \{E [\mathbf{v}_k^H(n-1)\mathbf{Q}^H\mathbf{R}_x\mathbf{Q}\mathbf{v}_l(n-1)]\} \\ &= \text{tr} \{E [\mathbf{v}_k^H(n-1)\mathbf{\Omega}\mathbf{v}_l(n-1)]\} \\ &= \sum_{i=0}^{N-1} \omega_i E [v_{k,i}^*(n-1)v_{l,i}(n-1)]. \end{aligned} \quad (35)$$

The EMSE of the combined filter can now be computed as

$$\text{EMSE} = \sum_{i=0}^{N-1} \omega_i E [|\lambda(n)v_{k,i}(n-1) + (1-\lambda(n))v_{l,i}(n-1)|^2], \quad (36)$$

where the components of type $E[v_{k,i}(n-1)v_{l,i}(n-1)]$ are given by (32). To compute $\lambda(n)$, we use (15), substituting (35) for its individual components.

4. SIMULATION RESULTS

A simulation study was carried out with the aim of verifying the approximations made in Section 3. In particular we are interested in how well the small step-size theory applies to our combination scheme of two adaptive filters.

We have combined two 64 tap long adaptive filters. In order to obtain a practical algorithm, the expectation operators in both the numerator and denominator of (7) have been replaced by exponential averaging of the type

$$P_u(n) = (1-\gamma)P_u(n-1) + \gamma u^2(n), \quad (37)$$

where $u(n)$ is the signal to be averaged, $P_u(n)$ is the averaged quantity, and $\gamma = 0.01$. The averaged quantities were then used in (7) to obtain λ . We have selected the sample echo path model number one from [24], to be the unknown system to identify. The impulse response of the echo path is shown in Figure 2.

The curves in Figures 3–6 are averaged over 100 trials. The input signal x is formed from the Gaussian white noise with unity variance by passing it through the filter with the transfer function

$$H(z) = \frac{1}{1 - 0.5z^{-1} - 0.1z^{-2}} \quad (38)$$

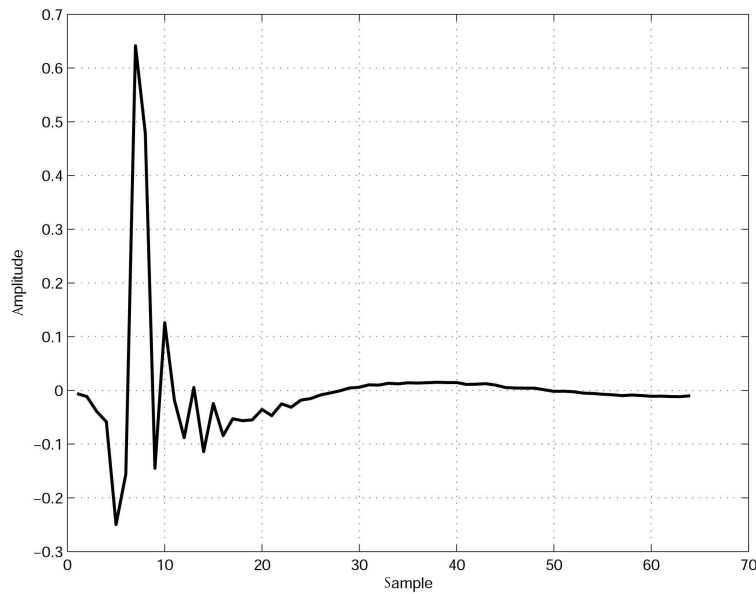


Fig. 2. The true impulse response.

to get a coloured input signal. The measurement noise is Gaussian white noise, statistically independent of x . The solid lines represent the simulation results and the dashed lines are the theoretical results. The theory predicts the simulation results rather well and because of that in Figs 3–6 the theoretical curve overlaps with the simulation result.

As seen from Figure 3, there is a rapid convergence in the beginning, followed by a stabilization period. When the slow adapting filter gets better than the fast one between sample times 15 000 and 20 000, a second convergence occurs. One can observe a good resemblance of simulation and theoretical curves.

Figure 4 depicts the time evolution of the combination parameter λ in this simulation. At the beginning of the test case the combination parameter is close to one. Correspondingly the output signal of the fast

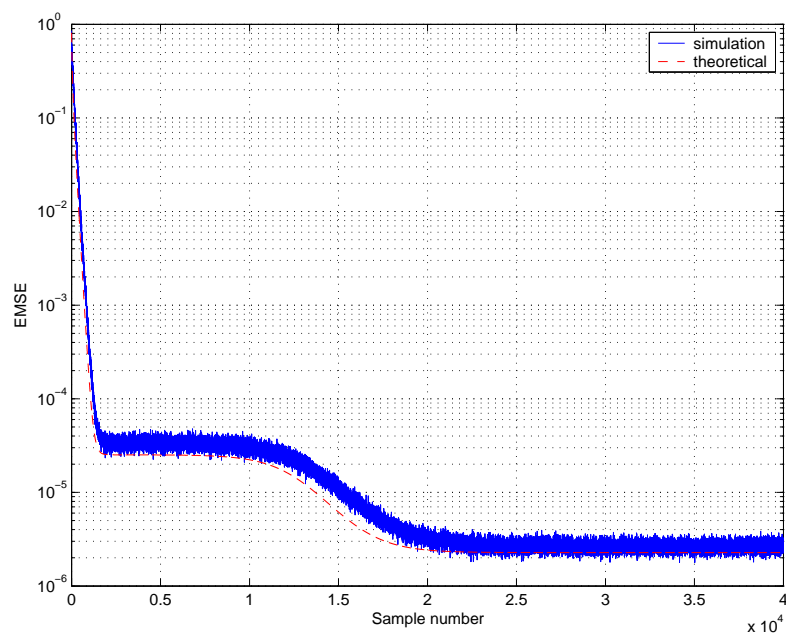


Fig. 3. Time-evolutions of EMSE with $\mu_1 = 0.5$, $\mu_2 = 0.05$, and $\sigma_v^2 = 10^{-4}$.

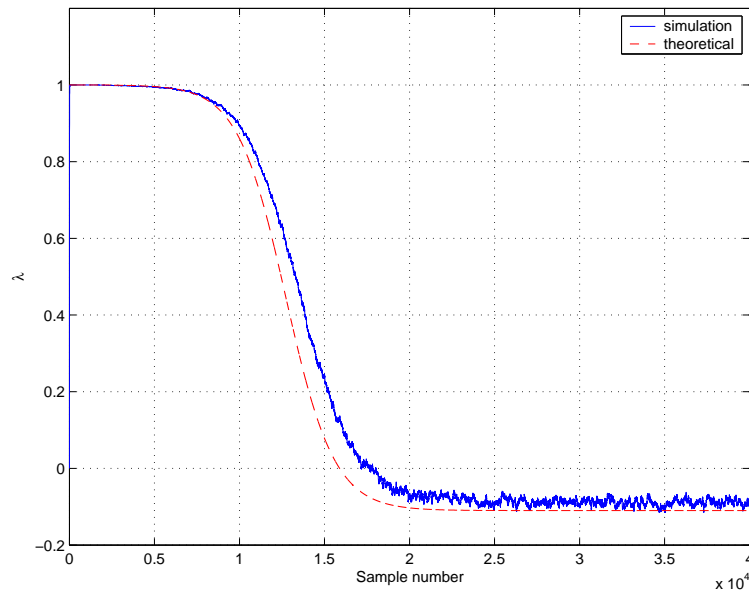


Fig. 4. Time-evolutions of λ with $\mu_1 = 0.5$, $\mu_2 = 0.05$, and $\sigma_v^2 = 10^{-4}$.

filter is used as the output of the combination. After a while, when the slow filter catches up the fast one and becomes better, λ changes towards zero and eventually becomes a small negative number. In this state the slow but more accurate filter determines the combined output. Again, one can see that there is a clear similarity between the lines.

The reason for λ becoming negative at the end of adaptation lies in that we have two adaptive filters working in parallel on the same input signal and the criterion for selecting λ is minimization of the mean square error of the output signal. At the end of adaptation some of the additive noise $v(n)$ will be cancelled out if λ is allowed to become negative to provide the smallest possible mean square error.

In Figure 5 we have made the difference between the step sizes small. One can see that the characteristic horizontal part of the learning curve has almost disappeared. We have also increased the measurement noise level. The simulation and theoretical curves show a good match.

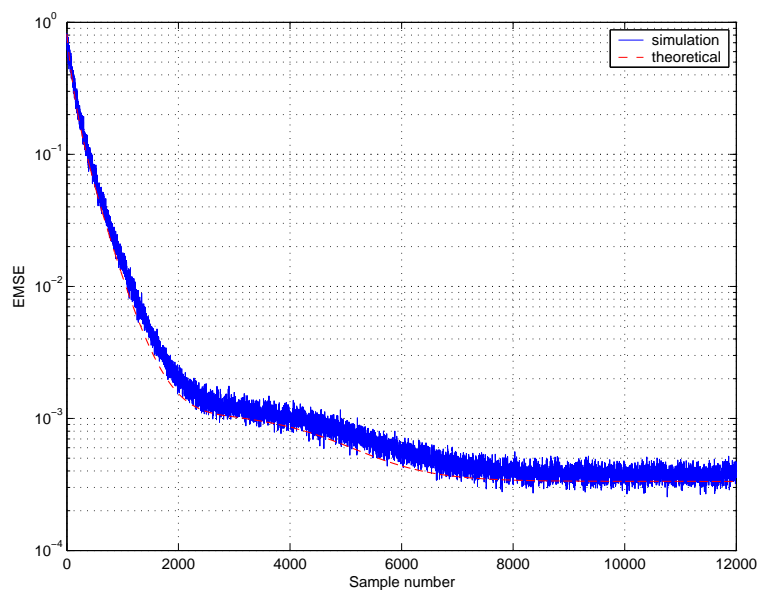


Fig. 5. EMSE with $\mu_1 = 0.2$, $\mu_2 = 0.1$, and $\sigma_v^2 = 10^{-2}$.

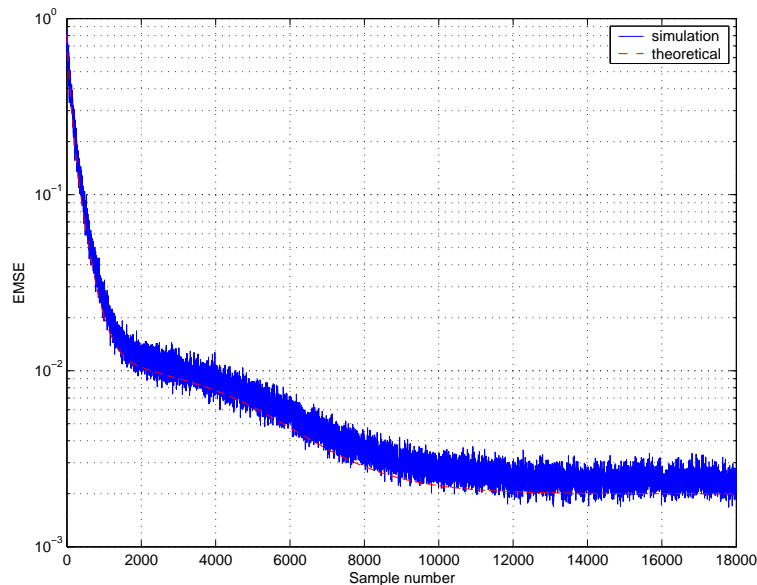


Fig. 6. EMSE with $\mu_1 = 0.2$, $\mu_2 = 0.05$, and $\sigma_v^2 = 10^{-1}$.

In Figure 6 we have increased the measurement noise level even more. One can see that the theoretical simulation results agree well.

5. CONCLUSIONS

We have investigated a combination of two NLMS adaptive filters that are simultaneously applied to the same input signals. The mixing parameter λ was computed, using the output signals of the individual filters and the desired signal. The transient behaviour of the algorithm was investigated, using the assumption of a small step size, and the expressions for evolution of $EMSE(n)$ and $\lambda(n)$ were derived. Finally, it was shown in the simulation study that the derived formulae fit the simulation results well.

REFERENCES

1. Sayed, A. H. *Adaptive Filters*. John Wiley and Sons, Hoboken, NJ, USA, 2008.
2. Haykin, S. *Adaptive Filter Theory, Fourth Edition*. Prentice Hall, 2002.
3. Harris, R. W., Chabries, D. M., and Bishop, F. A. Variable step (vs) adaptive filter algorithm. *IEEE Trans. Acoust. Speech Signal Process.*, 1986, **34**, 309–316.
4. Kwong, R. H. and Johnston, E. W. A variable step size LMS algorithm. *IEEE Trans. Signal Process.*, 1992, **40**, 1633–1642.
5. Matthews, V. J. and Xie, Z. A stochastic gradient adaptive filter with gradient adaptive step size. *IEEE Trans. Signal Process.*, 1993, **41**, 2075–2087.
6. Arenas-Garcia, J., Figueiras-Vidal, A. R., and Sayed, A. H. A robust variable step-size LMS-type algorithm: Analysis and simulations. *IEEE Trans. Signal Process.*, 1997, **45**, 631–639.
7. Shin, H. C. and Sayed, A. H. Variable step-size NLMS and affine projection algorithms. *IEEE Signal Process. Lett.*, 2004, **11**, 132–135.
8. Martinez-Ramon, M., Arenas-Garcia, J., Navia-Vazquez, A., and Figueiras-Vidal, A. R. An adaptive combination of adaptive filters for plant identification. In *Proc. 14th International Conference on Digital Signal Processing*. Santorini, Greece, 2002, 1195–1198.
9. Arenas-Garcia, J., Figueiras-Vidal, A. R., and Sayed, A. H. Mean-square performance of convex combination of two adaptive filters. *IEEE Trans. Signal Process.*, 2006, **54**, 1078–1090.
10. Silva, M. T. M., Nascimento, V. H., and Arenas-Garcia, J. A transient analysis for the convex combination of two adaptive filters with transfer of coefficients. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Dallas, TX, USA, 2010, 3842–3845.

11. Bershada, N. J., Bermudez, J. C., and Tourneret, J. H. An affine combination of two LMS adaptive filters – transient mean-square analysis. *IEEE Trans. Signal Process.*, 2008, **56**, 1853–1864.
12. Mandic, D., Vayanos, P., Boukis, C., Jelfs, B., Goh, S. I., Gautama, T., and Rutkowski, T. Collaborative adaptive learning using hybrid filters. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Honolulu, Hawaii, 2007, 901–924.
13. Zhang, Y. and Chambers, J. A. Convex combination of adaptive filters for a variable tap-length LMS algorithm. *IEEE Signal Process. Lett.*, 2006, **10**, 628–631.
14. Kim, K., Choi, Y., Kim, S., and Song, W. Convex combination of affine projection filters with individual regularization. In *Proc. 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*. Shimonoseki, Japan, 2008, 901–904.
15. Fathiyan, A. and Eshghi, M. Combining several PBS-LMS filters as a general form of convex combination of two filters. *J. Appl. Sci.*, 2009, **9**, 759–764.
16. Azpicueta-Ruiz, L. A., Figueiras-Vidal, A. R., and Arenas-Garcia, J. 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, 2008, 3301–3304.
17. Ochiai, K. Echo canceller with two echo path models. *IEEE Trans. Commun.*, 1977, **25**, 589–594.
18. Armbruster, W. Wideband acoustic echo canceller with two filter structure. In *Signal Processing VI, Theories and Applications* (Vanderwalle, J., Boite, R., Moonen, M., and Oosterlinck, A., eds). Elsevier Science Publishers B.V, 1992, 1611–1617.
19. Trump, T. An output signal based combination of two NLMS adaptive algorithms. In *Proc. 16th International Conference on Digital Signal Processing*. Santorini, Greece. IEEE 2009, 6 pp.
20. Azpicueta-Ruiz, L. A., Figueiras-Vidal, A. R., and Arenas-Garcia, J. A new least squares adaptation scheme for the affine combination of two adaptive filters combination of two adaptive filters. In *Proc. IEEE International Workshop on Machine Learning for Signal Processing*. Cancun, Mexico, 2008, 327–332.
21. Trump, T. Steady state analysis of an output signal based combination of two NLMS adaptive filters. In *Proc. 17th European Signal Processing Conference*. Glasgow, Scotland, 2009, 1720–1724.
22. Trump, T. Tracking performance of a combination of two NLMS adaptive filters. In *Proc. IEEE Workshop on Statistical Signal Processing*. Cardiff, UK. IEEE 2009, 181–184.
23. Trump, T. Output statistics of a line enhancer based on a combination of two adaptive filters. *Central European J. Engineering*, 2011, **1**, 244–252.
24. *ITU-T Recommendation G.168 Digital Network Echo Cancellers*. ITU-T, 2009; <http://itu.int> (accessed 27 July 2011).

Väljundsignaalidel baseeruv adaptiivsete NLMS-filtrite kombinatsioon: koonduvusanalüüs

Tõnu Trump

On uuritud kahe adaptiivse NLMS-filtri, mis adapteeruvad samadele sisendsignaalidele, kombinatsiooni. Ühel filtril on suur ja teisel väike sammu pikkus. Filtrite väljundid kombineeritakse λ -parameetri abil, mille leiame filtrite väljundsignaalidest. Selline kombinatsioon on uus viis saavutada üheaegselt adaptiivse algoritmi kiire koonduvus ja väike väljakujunenud oleku viga. Artiklis on tuletatud valemid, mis võimaldavad arvutada ruutkeskmise vea kogu filtri töö kestel. Teoreetilised tulemused on arvutikatsetega kontrollitud.