
Applications of a Combination of Two Adaptive Filters

Tõnu Trump

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/50451>

1. Introduction

Designing a Least Mean Square (LMS) family adaptive algorithm includes solving the well-known trade-off between the initial convergence speed and the mean-square error in steady state according to the requirements of the application at hands. The trade-off is controlled by the step-size parameter of the algorithm. 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 [9,17]. In several applications it is, however, eligible to have both and hence it would be very desirable to be able to design algorithms that can overcome the named trade-off.

Variable step size adaptive schemes offer a potential solution allowing to achieve both fast initial convergence and low steady state misadjustment [1, 8, 12, 15, 18]. 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 the algorithm proposed in [15] changes the time-varying convergence parameters in such a way that the change is proportional to the negative of gradient of the squared estimation error with respect to the convergence parameter. Squared instantaneous errors have been used in [12] and the squared autocorrelation of errors at adjacent time instances in [1] to modify the step size. In reference [18] the norm of projected weight error vector is used as a criterion to determine how close the adaptive filter is to its optimum performance.

More recently there has been an interest in a combination scheme that is able to optimize the trade-off between convergence speed and steady state error [14]. 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 for some parameter update schemes [2, 6, 19]. The reference [2] uses convex combination i.e. λ is constrained to lie between 0 and 1. The reference [19] presents a transient analysis of a slightly modified version of this scheme. The parameter λ is in those papers found using an LMS type adaptive scheme and computing the sigmoidal function of the result. The reference [6] takes another approach computing the mixing parameter using an affine combination. This paper uses the ratio of time averages of the instantaneous errors of the filters. The error function of the ratio is then computed to obtain λ .

In [13] a convex combination of two adaptive filters with different adaptation schemes has been investigated with the aim to improve the steady state characteristics. One of the adaptive filters in that paper uses LMS algorithm and the other one Generalized Normalized Gradient Decent algorithm. The combination parameter λ is computed using stochastic gradient adaptation. In [24] the convex combination of two adaptive filters is applied in a variable filter length scheme to gain improvements in low SNR conditions. In [11] the combination has been used to join two affine projection filters with different regularization parameters. The work [7] uses the combination on parallel binary structured LMS algorithms. These three works use the LMS like scheme of [5] to compute λ .

It should be noted that schemes involving two filters have been proposed earlier [3, 16]. However, in those early schemes only one of the filters have been adaptive while the other one has used fixed filter weights. Updating of the fixed filter has been accomplished by copying of all the coefficients from the adaptive filter, when the adaptive filter has been performing better than the fixed one.

In this Chapter we compute the mixing parameter λ 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 [21] and [4]. In [23], the output signal based combination was used in adaptive line enhancer and in [22] it was used in the system identification application.

We will investigate three applications of the combination: system investigation, adaptive beamforming and adaptive line enhancer. We describe each of the applications in detail and present a proper analysis.

We will assume throughout the Chapter that the signals are complex-valued and that the combination scheme uses two 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 and the superscript H Hermitian transposition of a matrix. The operator $E[\cdot]$ denotes mathematical expectation, $Re\{\cdot\}$ is the real part of a complex variable and $Tr[\cdot]$ stands for the trace of a matrix.

2. Combination of Two Adaptive Filters

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

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

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \mu_i e_i^*(n) \mathbf{x}(n). \quad (2)$$

In the above $w_i(n)$ is the N vector of coefficients of the i -th adaptive filter, with $i = 1, 2$ and $x(n)$ is the known N input vector, common for both of the adaptive filters. The input process is assumed to be a zero mean wide sense stationary Gaussian process. μ_i is the step size of i -th adaptive filter. We assume without loss of generality that $\mu_1 > \mu_2$. The case $\mu_1 = \mu_2$ is not interesting as in this case the two filters remain equal and the combination renders to a single filter.

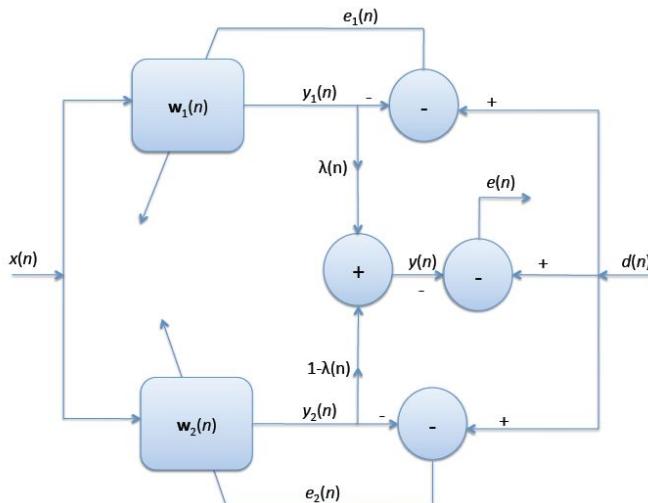


Figure 1. The combined adaptive filter.

The desired signal in 1 can be expressed as

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

where the vector w_o is the optimal Wiener filter coefficient vector for the problem at hands and the process $\zeta(n)$ is the irreducible error that is statistically independent of all the other signals.

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 $\lambda(n)$ can be any real number.

We define the *a priori* system error signal as difference between the output signal of the optimal Wiener filter at time n , given by $y_o(n) = \mathbf{w}_o^H\mathbf{x}(n) = d(n) - \zeta(n)$, 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

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

Setting the derivative to zero results in

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

where we have replaced the Wiener filter output signal $y_o(n)$ by its observable noisy version $d(n)$. Note however, that because the input signal $x(n)$ and irreducible error $\zeta(n)$ 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 in 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 to the denominator of (7). The constant 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. System Identification

In several areas it is essential to build a mathematical model of some phenomenon or system. In this class of applications, the adaptive filter can be used to find a best fit of a linear model to an unknown plant. The plant and the adaptive filter are driven by the same known input signal and the plant output provides the desired signal of the adaptive filter. The plant can be dynamic and in this case we have a time varying model. The system identification

configuration is depicted in Figure 2. As before $x(n)$ is the input signal, $v(n)$ is the measurement noise, $y(n)$ is the adaptive filter output signal and $e(n)$ is the error signal. The desired signal is $d(n) = \mathbf{w}_o^H \mathbf{x}(n) + \zeta(n)$, where \mathbf{w}_o is the vector of Wiener filter coefficients and the irreducible error $\zeta(n)$ consists of the measurement noise $v(n)$ together with the effects of the plant that can not be explained with a length N linear model. The result of pure system identification problem is the vector of adaptive filter coefficients.

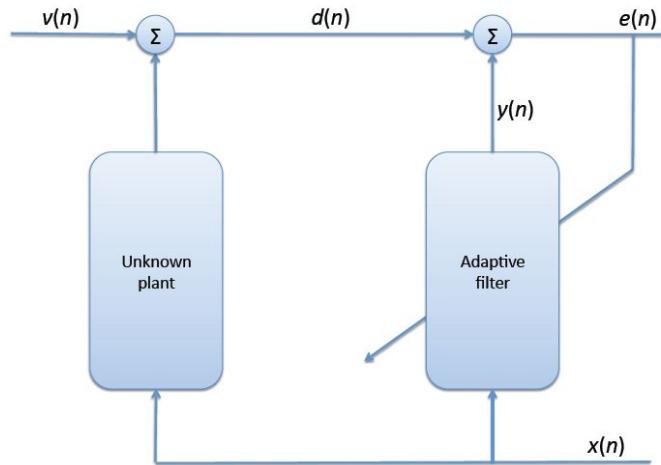


Figure 2. Block diagram of generelized sidelobe canceller.

The same basic configuration is also used to solve the echo and noise cancellation problems. In echo cancellation the unknown plant is the echo path either electrical or acoustical and the input signal $x(n)$ is the speech signal of one of the parties of telephone conversation. Speech of the other party is contained in the signal $v(n)$. The objective is to cancel the components of the desired signal that are due to the input $x(n)$.

In noise cancellation problems the signal $v(n)$ is the primary microphone signal containing noise and the signal to be cleaned. The input signal $x(n)$ is formed by the reference microphones. The reference signals are supposed to be correlated with the noise in the primary signal but not with the useful signal. The objective here is to suppress the noise and clean the signal of interest i.e. $v(n)$.

In here we are going to use the combination of two adaptive filters described in the previous Section to solve the system identification problem.

3.2 Excess Mean Square Error

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 predict 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 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 of the combined filter $MSD = E[\tilde{\mathbf{w}}^H(n)\tilde{\mathbf{w}}(n)]$ is given by

$$MSD = \lambda^2 E[\tilde{\mathbf{w}}_1^H(n)\tilde{\mathbf{w}}_1(n)] + 2\lambda(1-\lambda)Re\{E[\tilde{\mathbf{w}}_2^H(n)\tilde{\mathbf{w}}_1(n)]\} + (1-\lambda)^2 E[\tilde{\mathbf{w}}_2^H(n)\tilde{\mathbf{w}}_2(n)]. \quad (10)$$

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 of the combination, $EMSE(n) = E[|e_a(n)|^2]$,

$$E[|e_a(n)|^2] = \lambda^2 E[|e_{1,a}(n)|^2] + 2\lambda(1-\lambda)E[Re\{e_{1,a}(n)e_{2,a}^*(n)\}] + (1-\lambda)^2 E[|e_{2,a}(n)|^2]. \quad (13)$$

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

$$E[|e_a(n)|^2] = \lambda^2 E[\tilde{\mathbf{w}}_1^H \mathbf{x} \mathbf{x}^H \tilde{\mathbf{w}}_1] + 2\lambda(1-\lambda)E[Re\{\tilde{\mathbf{w}}_1^H \mathbf{x} \mathbf{x}^H \tilde{\mathbf{w}}_2\}] + (1-\lambda)^2 E[\tilde{\mathbf{w}}_2^H \mathbf{x} \mathbf{x}^H \tilde{\mathbf{w}}_2]. \quad (14)$$

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

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

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

Reformulation the relation (1) as

$$e_i(n) = d(n) - \mathbf{w}_i^H(n-1)\mathbf{x}(n) = e_o(n) + \tilde{\mathbf{w}}_i^H(n-1)\mathbf{x}(n) \quad (16)$$

and subtracting (2) from w_o we have

$$\tilde{\mathbf{w}}_i(n) = (\mathbf{I} - \mu_i \mathbf{x}\mathbf{x}^H) \tilde{\mathbf{w}}_i(n-1) - \mu_i \mathbf{x} e_o^*(n). \quad (17)$$

We next approximate the outer product of input signal vectors by its correlation matrix $\mathbf{x}\mathbf{x}^H \approx \mathbf{R}_x$. The approximation is justified by the fact that with small step size the weight error update of the LMS algorithm (17) behaves like a low pass filter with a low cutoff frequency. With this approximations we have

$$\tilde{\mathbf{w}}_i(n) \approx (\mathbf{I} - \mu_i \mathbf{R}_x) \tilde{\mathbf{w}}_i(n-1) - \mu_i \mathbf{x} e_o^*(n). \quad (18)$$

This means in fact that we apply the small step size theory [9] even if the assumption of small step size is not really true for the fast adapting filter. In our simulation study we will see, however, that the assumption works in practice rather well.

Let us now define the eigendecomposition of the correlation matrix as

$$\mathbf{Q}^H \mathbf{R}_x \mathbf{Q} = \boldsymbol{\Omega}, \quad (19)$$

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

$$\mathbf{v}_i(n) = \mathbf{Q}^H \tilde{\mathbf{w}}_i(n) \quad (20)$$

and the transformed last term of equation (18) as

$$\mathbf{p}_i(n) = \mu_i \mathbf{Q}^H \mathbf{x} e_o^*(n). \quad (21)$$

Then we can rewrite the equation (18) after multiplying both sides by \mathbf{Q}^H from the left as

$$\mathbf{v}_i(n) = (\mathbf{I} - \mu_i \boldsymbol{\Omega}) \mathbf{v}_i(n-1) - \mathbf{p}_i(n). \quad (22)$$

We note that the mean of \mathbf{p}_i is zero by the orthogonality theorem and the crosscorrelation matrix of \mathbf{p}_k and \mathbf{p}_l equals

$$E[\mathbf{p}_k \mathbf{p}_l^H] = \mu_k \mu_l \mathbf{Q}^H E[\mathbf{x} e_o^*(n) e_o(n) \mathbf{x}^H] \mathbf{Q}. \quad (23)$$

We now invoke the Gaussian moment factoring theorem to write

$$E[\mathbf{x} e_o^*(n) e_o(n) \mathbf{x}^H] = E[\mathbf{x} e_o^*(n)] E[e_o(n) \mathbf{x}^H] + E[\mathbf{x} \mathbf{x}^H] E[|e_o|^2]. \quad (24)$$

The first term in the above is zero due to the principle of orthogonality and the second term equals $\mathbf{R} J_{min}$, where $J_{min} = E[|e_o|^2]$ is the minimum mean square error produced by the corresponding Wiener filter. Hence we are left with

$$E[\mathbf{p}_k \mathbf{p}_l^H] = \mu_k \mu_l J_{min} \boldsymbol{\Omega}. \quad (25)$$

As the matrices \mathbf{I} and $\boldsymbol{\Omega}$ in (22) are both diagonal, it follows that the m -th element of vector $\mathbf{v}_i(n)$ is given by

$$\begin{aligned} v_{i,m}(n) &= (1 - \mu_i \omega_m) v_{i,m}(n-1) - p_{i,m}(n) \\ &= (1 - \mu_i \omega_m)^n v_m(0) + \sum_{i=0}^{n-1} (1 - \mu_i \omega_m)^{n-1-i} p_{i,m}(i), \end{aligned} \quad (26)$$

where ω_m is the m -th eigenvalue of \mathbf{R}_x and $v_{i,m}$ and $p_{i,m}$ are the m -th components of the vectors \mathbf{v}_i and \mathbf{p}_i respectively.

We immediately see that the mean value of $v_{i,m}(n)$ equals

$$E[v_{i,m}(n)] = (1 - \mu_i \omega_m)^n v_m(0) \quad (27)$$

as the vector \mathbf{p}_i has zero mean.

To proceed with our development for the combination of two LMS filters we note that we can express the MSD and its individual components in (10) through the transformed weight error vectors as

$$\begin{aligned} E[\tilde{\mathbf{w}}_k^H(n) \tilde{\mathbf{w}}_l(n)] &= E[\mathbf{v}_k^H(n) \mathbf{v}_l(n)] \\ &= \sum_{m=0}^{N-1} E[v_{k,m}(n) v_{l,m}^*(n)] \end{aligned} \quad (28)$$

so we also need to find the auto- and cross correlations of v .

Let us concentrate on the m -th component in the sum above corresponding to the cross term. The expressions for the component filters follow as special cases. Substituting (26) into the expression of m -th component of MSD above, taking the mathematical expectation and noting that the vector p is independent of $v(0)$ results in

$$\begin{aligned} E[v_{k,m}(n)v_{l,m}^*(n)] &= E[(1-\mu_k\omega_m)^n v_k(0)(1-\mu_l\omega_m)^n v_l^*(0)] \\ &\quad + E\left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (1-\mu_k\omega_m)^{n-1-i} (1-\mu_l\omega_m)^{n-1-j} p_{k,m}(i)p_{l,m}^*(j)\right]. \end{aligned} \quad (29)$$

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

$$[v_{k,m}(0)=v_{l,m}(0)=v_m(0)]$$

and that

$$E[p_{k,m}(i)p_{l,m}^*(j)] = \begin{cases} \mu_k\mu_l\omega_m J_{min}, & i=j \\ 0, & \text{otherwise}. \end{cases} \quad (30)$$

We then have for the m -th component of MSD

$$\begin{aligned} E[v_{k,m}(n)v_{l,m}^*(n)] &= (1-\mu_k\omega_m)^n (1-\mu_l\omega_m)^n |v_m(0)|^2 \\ &\quad + \mu_k\mu_l\omega_m J_{min} (1-\mu_k\omega_m)^{n-1} (1-\mu_l\omega_m)^{n-1} \\ &\quad \cdot \sum_{i=0}^{n-1} (1-\mu_k\omega_m)^{-i} (1-\mu_l\omega_m)^{-i}. \end{aligned} \quad (31)$$

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

$$\begin{aligned} \sum_{i=0}^{n-1} (1-\mu_k\omega_m)^{-i} (1-\mu_l\omega_m)^{-i} &= \frac{1 - [(1-\mu_k\lambda_m)^{-1}(1-\mu_l\lambda_m)^{-1}]^n}{1 - (1-\mu_k\lambda_m)^{-1}(1-\mu_l\lambda_m)^{-1}} \\ &= \frac{(1-\mu_k\omega_m)(1-\mu_l\omega_m)}{\mu_k\mu_l\omega_m^2 - \mu_k\omega_m - \mu_l\omega_m} - \frac{(1-\mu_k\omega_m)^{-n+1}(1-\mu_l\omega_m)^{-n+1}}{\mu_k\mu_l\omega_m^2 - \mu_k\omega_m - \mu_l\omega_m}. \end{aligned} \quad (32)$$

After substitution of the above into (31) and simplification we are left with

$$E[v_{k,m}(n)v_{l,m}^*(n)] = (1-\mu_k\omega_m)^n(1-\mu_l\omega_m)^n \left[|v_m(0)|^2 + \frac{J_{min}}{\omega_m^2 - \frac{\omega_m}{\mu_l} - \frac{\omega_m}{\mu_k}} \right] - \frac{J_{min}}{\omega_m^2 - \frac{\omega_m}{\mu_l} - \frac{\omega_m}{\mu_k}}, \quad (33)$$

which is our result for a single entry to the MSD crossterm vector. It is easy to see that for the terms involving a single filter we get an expressions that coincide with the one available in the literature [9].

Let us now focus on the cross term

$$EMS E_{kl} = E[\tilde{\mathbf{w}}_k^H(n-1)\mathbf{x}(n)\mathbf{x}^H(n)\tilde{\mathbf{w}}_l(n-1)],$$

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

$$\begin{aligned} EMS E_{kl} &= E[\tilde{\mathbf{w}}_k^H(n-1)\mathbf{R}_x\tilde{\mathbf{w}}_l(n-1)] \\ &= Tr\{E[\mathbf{R}_x\tilde{\mathbf{w}}_l(n-1)\tilde{\mathbf{w}}_k^H(n-1)]\} \\ &= Tr\{\mathbf{R}_x E[\tilde{\mathbf{w}}_l(n-1)\tilde{\mathbf{w}}_k^H(n-1)]\}. \end{aligned} \quad (34)$$

Let us now recall that according to (20) for any of the filters $\tilde{\mathbf{w}}_i(n) = \mathbf{Q}\mathbf{v}_i(n)$ so that we are justified to write

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

The EMSE of the combined filter can now be computed as

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

where the components of type $E[v_{k,i}(n-1)v_{l,i}(n-1)]$ are given by (33). To compute $\lambda(n)$ we use (15) substituting (35) for its individual components.

4. Adaptive Sensor Array

In this Chapter we describe how to use the combination of two adaptive filters in an adaptive beamformer. The beamformer we employ here is often termed as Generalized Side-lobe Canceller [9].

Let ϕ denote the angle of incidence of a planar wave impinging a linear sensor array, measured with respect to the normal to the array. The electrical angle θ is related to the incidence angle as

$$\theta = \frac{2\pi\delta}{\lambda} \sin \phi, \quad (37)$$

where λ is the wavelength of the incident wave and δ is the spacing between adjacent sensors of the linear array.

Suppose that the signal impinging the array of $M=N+1$ sensors is given by

$$\mathbf{u}(n) = \mathbf{A}(\Theta)\mathbf{s}(n) + \mathbf{v}(n), \quad (38)$$

where $\mathbf{s}(n)$ is the vector of emitter signals, Θ is a collection of directions of arrivals, $\mathbf{A}(\Theta)$ is the array steering matrix with its columns $\mathbf{a}(\theta)$ defined as responses toward the individual sources $s(n)$ and $\mathbf{v}(n)$ is a vector of additive circularly symmetric Gaussian noise. The M vectors

$$\mathbf{a}(\theta) = [1, e^{j\theta}, \dots, e^{j(M-1)\theta}]^T \quad (39)$$

are called the steering vectors of the respective sources. We assume that the source of interest is located at the electrical angle θ_0 .

The block diagram of the Generalized Sidelobe Canceller is shown in Figure 3. The structure consists of two branches. The upper branch is the steering branch, that directs its beam toward the desired source. The lower branch is the blocking branch that blocks the signals impinging at the array from the direction of the desired source and includes an adaptive algorithm that minimizes the mean square error between the output signals of the branches.

The weights in steering branch \mathbf{w}_s are selected from the condition

$$\mathbf{w}_s^H \mathbf{a}(\theta_0) = g \quad (40)$$

i.e. we require the response in the direction of the source of interest θ_0 to equal a constant g . Common choices for g are $g=M$ and $g=1$. Here we have used $g=M$.

The signal at the output of the upper branch is given by

$$d(n) = \mathbf{w}_s^H \mathbf{u}(n). \quad (41)$$

In the lower branch we have a blocking matrix, that will block any signal coming from the direction θ_b . The columns of the $M \times M-1$ blocking matrix C_b are defined as being the orthogonal complement of the steering vector $a(\theta_b)$ in the upper branch

$$\mathbf{a}^H(\theta_b) \mathbf{C}_b = \mathbf{0}. \quad (42)$$

The vector valued signal $x(n)$ at the output of the blocking matrix is formed as

$$\mathbf{x}(n) = \mathbf{C}_b^H \mathbf{u}(n). \quad (43)$$

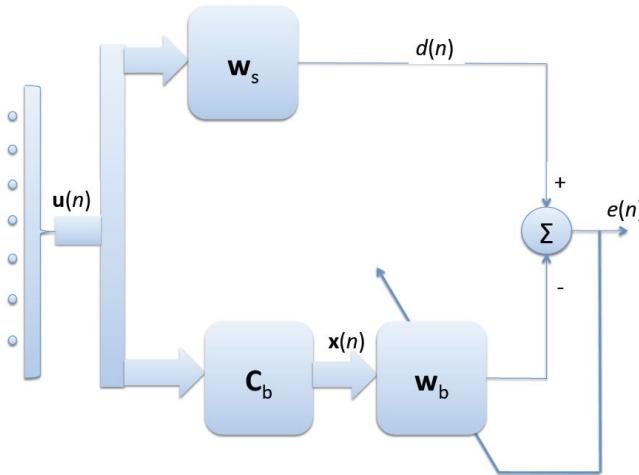


Figure 3. Block diagram of generelized sidelobe canceller.

The output of the algorithm is

$$e(n) = d(n) - \mathbf{w}_b^H(n) \mathbf{x}(n). \quad (44)$$

The signals $x(n)$ and $d(n)$ can be used as the input and desired signals respectively in an adaptive algorithm to select the blocking weights w_b . In this Chapter we use the combination of two adaptive filters that gives us fast initial convergence and low steady state misadjustment at the same time.

4.1 Signal to Interference and Noise Ratio

The EMSE of the adaptive algorithm can be analysed as it is done in Section 3.1. In this application we are also interested in signal to interference and noise ratio (SINR) at the array output. To evaluate this we first note that the power that signal of interest generates at the array output is according to (40)

$$P_s = \mathbf{w}_s^H \mathbf{a}(\theta_0) \sigma_{s_0}^2 \mathbf{a}^H(\theta_0) \mathbf{w}_s = |g|^2 \sigma_{s_0}^2, \quad (45)$$

where $\sigma_{s_0}^2$ is the variance of the useful signal arriving from the angle θ_0 .

To find the interference and noise power we first define the reduced signal vector $\check{\mathbf{s}}$ and a reduced DOA collection $\check{\Theta}$ where we have left out the signal of interest and the steering vector corresponding to the useful signal but kept all the interferers and the interference steering vectors. The corresponding array steering matrix is $\check{\mathbf{A}}(\check{\Theta})$.

The correlation matrix of interference and noise in the signal $x(n)$, which is the input signal to our adaptive scheme, is then given by

$$\check{\mathbf{R}}_x = \mathbf{C}_b^H \check{\mathbf{A}}(\check{\Theta}) E[\check{\mathbf{s}}\check{\mathbf{s}}^H] \check{\mathbf{A}}^H(\check{\Theta}) \mathbf{C}_b + \mathbf{C}_s^H \mathbf{C}_b \sigma_v^2, \quad (46)$$

where σ_v^2 is the noise variance, the first component in the summation is due to the interfering sources and the second component is due to the noise.

It follows from the standard Wiener filtering theory that the minimum interference and noise power at the array output is given by

$$\check{J}_{min} = \sigma_{int,v}^2 - \check{\mathbf{p}}^H \check{\mathbf{R}}^{-1} \check{\mathbf{p}}, \quad (47)$$

where the desired signal variance excluding the signal from the source of interest is

$$\sigma_{int,v}^2 = \mathbf{w}_s^H \check{\mathbf{A}} \check{\mathbf{R}} \check{\mathbf{A}}^H \mathbf{w}_s + \sigma_v^2 \mathbf{w}_s^H \mathbf{w}_s \quad (48)$$

and the crosscorrelation vector between the adaptive filter input signal and desired signal excluding the signal from source of interest is

$$\check{\mathbf{p}} = \mathbf{C}_b^H \check{\mathbf{A}}(\check{\Theta}) E[\check{\mathbf{s}}\check{\mathbf{s}}^H] \check{\mathbf{A}}^H(\check{\Theta}) \mathbf{w}_s + \sigma_v^2 \mathbf{C}_b^H \check{\mathbf{A}}(\check{\Theta}) \mathbf{w}_s. \quad (49)$$

We can now find the eigendecomposition of $\check{\mathbf{R}}_x$ and use the resulting eigenvalues in (35) and (36) to find the excess mean square error due to interference and noise only $EMSE_{int,v}$. The error power can be computed as minimum interference and noise power at the array output plus excess mean square error due to interference and noise only

$$P_{v,int} = \check{J}_{min} + EMSE_{int,v}(n) \quad (50)$$

and the signal to noise ratio is thus given by

$$SNR(n) = \frac{P_s}{P_{v,int}(n)}. \quad (51)$$

5. Adaptive Line Enhancer

Adaptive line enhancer is a device that is able to separate the input into two components. One of them consists mostly of the narrow-band signals that are present at the input and the other one consists mostly of the broadband noise. In the context of this paper the signal is considered to be of narrow band if its bandwidth is small as compared to the sampling frequency of the system.

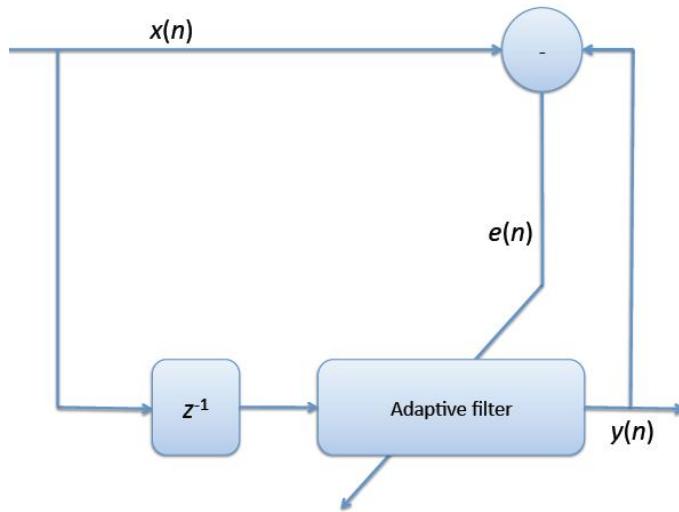


Figure 4. The adaptive line enhancer.

We assume that the broadband noise is zero mean, white and Gaussian and that the narrow band component is centred. One is usually interested in the narrow band components and the device is often used to clean narrow band signals from noise before any further processing. The line enhancer is shown in Figure 4. Note that the input signal to the adaptive filter of the line enhancer is delayed by Δ sample times and the input vector is thus $x(n-\Delta)$. The desired signal is $d(n) = x(n)$. The line enhancer is in fact a Δ step predictor. The device is able to predict the narrow band components that have long correlation times but it cannot pre-

dict the white noise and hence, only a prediction of narrow band components appears in the filter output signal $y(n)$. The signal $y(n)$ is also the output of the system.

Let us now find the autocorrelation function of the enhancer output signal $y(n)$. We make the standard assumption from independence theory which states that the filter weights and the input signal are independent [9].

The l -th autocorrelation lag of the filter output process $r(l) = E[y(n)y^*(n+l)]$, equals

$$r(l) = E[\mathbf{w}^H(n-\Delta)\mathbf{x}(n-\Delta)\mathbf{x}^H(n-\Delta+l)\mathbf{w}(n-\Delta+l)]. \quad (52)$$

The input signal $x(n)$ consists of two uncorrelated components $s(n)$, the sum of narrow band signals, and $v(n)$, the additive noise

$$x(n) = s(n) + v(n). \quad (53)$$

We can decompose the impulse response of the adaptive filter into two components. One of them is the optimal Wiener filter for the problem

$$\mathbf{w}_o = E[\mathbf{x}(n-\Delta)\mathbf{x}^H(n-\Delta)]^{-1}E[\mathbf{x}(n-\Delta)x(n)] \quad (54)$$

and the other one, $\tilde{\mathbf{w}}(n)$, represents the estimation errors.

$$\tilde{\mathbf{w}}(n) = \mathbf{w}_o - \mathbf{w}(n). \quad (55)$$

The output signal can hence be expressed as

$$y(n) = y_o(n) - \tilde{y}(n). \quad (56)$$

Substituting (53) and (55) into (52) and noticing that the cross-correlation between the Wiener filter output and that of the filter defined by weight errors is

$$E[y_o(n)\tilde{y}^*(l)] = E[\mathbf{w}_o^H\mathbf{x}(n-\Delta)\mathbf{x}^H(l-\Delta)\tilde{\mathbf{w}}(n-1)] = 0 \quad (57)$$

because of the adopted independence assumption and because $E[\tilde{\mathbf{w}}(n-1)] = E[\mathbf{w}_o - \mathbf{w}(n-1)] = 0$, we have

$$\begin{aligned} r(l) &= E[\mathbf{w}_o^H\{\mathbf{s}(n-\Delta) + \mathbf{v}(n-\Delta)\}\{\mathbf{s}^H(n-\Delta+l) + \mathbf{v}^H(n-\Delta+l)\}\mathbf{w}_o] \\ &\quad + E[\tilde{\mathbf{w}}^H(n-1)\{\mathbf{s}(n-\Delta) + \mathbf{v}(n-\Delta)\} \\ &\quad \cdot \{\mathbf{s}^H(n-\Delta+l) + \mathbf{v}^H(n-\Delta+l)\}\tilde{\mathbf{w}}(n-1+l)]. \end{aligned} \quad (58)$$

Developing and grouping terms in the above equation results in

$$\begin{aligned} r(l) = & E[\mathbf{w}_o^H \mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l) \mathbf{w}_o] \\ & + E[\mathbf{w}_o^H \mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l) \mathbf{w}_o] \\ & + E[\tilde{\mathbf{w}}^H(n-1) \mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l) \tilde{\mathbf{w}}(n-1+l)] \\ & + E[\tilde{\mathbf{w}}^H(n-1) \mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l) \tilde{\mathbf{w}}(n-1+l)]. \end{aligned} \quad (59)$$

Using the fact that \mathbf{w}_o is deterministic and the properties of the trace operator we further obtain

$$\begin{aligned} r(l) = & \mathbf{w}_o^H E[\mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l) \mathbf{w}_o] \\ & + \mathbf{w}_o^H E[\mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l) \mathbf{w}_o] \\ & + E[\text{Tr}\{\tilde{\mathbf{w}}(n-1+l) \tilde{\mathbf{w}}^H(n-1) \mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l)\}] \\ & + E[\text{Tr}\{\tilde{\mathbf{w}}(n-1+l) \tilde{\mathbf{w}}^H(n-1) \mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l)\}]. \end{aligned} \quad (60)$$

We now invoke the independence assumption saying that the weight vector $\tilde{\mathbf{w}}^H(n-1)$ is independent from the signals $\mathbf{s}(n-\Delta)$ and $\mathbf{v}(n-\Delta)$. This leads us to

$$\begin{aligned} r(l) = & \mathbf{w}_o^H E[\mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l) \mathbf{w}_o] \\ & + \mathbf{w}_o^H E[\mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l) \mathbf{w}_o] \\ & + \text{Tr}\{E[\tilde{\mathbf{w}}(n-1+l) \tilde{\mathbf{w}}^H(n-1)] E[\mathbf{s}(n-\Delta) \mathbf{s}^H(n-\Delta+l)]\} \\ & + \text{Tr}\{E[\tilde{\mathbf{w}}(n-1+l) \tilde{\mathbf{w}}^H(n-1)] E[\mathbf{v}(n-\Delta) \mathbf{v}^H(n-\Delta+l)]\}. \end{aligned} \quad (61)$$

To proceed we need to find the matrix $\mathbf{K}(l) = E[\tilde{\mathbf{w}}(n-1+l) \tilde{\mathbf{w}}^H(n-1)]$.

5.1 Weight error correlation matrix

In this Section we investigate the combination of two adaptive filters and derive the expressions for the crosscorrelation matrix between the output signals of the individual filters $y_i(n)$ and $y_k(n)$. The autocorrelation matrices of the individual filter output signals follow directly using only one signal in the formulae.

For the problem at hands we can rewrite the equation (18) noting that we have introduced a Δ samples delay in the signal path as

$$\tilde{\mathbf{w}}_i(n) \approx (\mathbf{I} - \mu_i \mathbf{R}_x) \tilde{\mathbf{w}}_i(n-1) - \mu_i \mathbf{x}(n-\Delta) e_o^*(n). \quad (62)$$

For the weight error correlation matrix we then have

$$\begin{aligned}
 \mathbf{K}_{i,k,l}(n) = & E[\tilde{\mathbf{w}}_i(n+l)\tilde{\mathbf{w}}_k^H(n)] = E[(\mathbf{I} - \mu_i \mathbf{R}_x)\tilde{\mathbf{w}}_i(n+l-1)\tilde{\mathbf{w}}_k(n-1)^H(\mathbf{I} - \mu_k \mathbf{R}_x)] \\
 & - E[(\mathbf{I} - \mu_i \mathbf{R}_x)\tilde{\mathbf{w}}_i(n+l-1)\mu_k \mathbf{x}^H(n-\Delta)e_o(n)] \\
 & - E[\mu_i \mathbf{x}(n-\Delta)e_o^*(n)(\mathbf{I} - \mu_k \mathbf{R}_x)\tilde{\mathbf{w}}_k^H(n-1)] \\
 & + E[\mu_i \mu_k \mathbf{x}(n-\Delta)e_o^*(n)e_o(n)\mathbf{x}^H(n-\Delta)].
 \end{aligned}$$

The second and third terms of the above equal zero because we have made the usual independence theory assumptions which state, that the weight errors $\tilde{\mathbf{w}}_i(n)$ are independent of the input signal $\mathbf{x}(n-\Delta)$. To evaluate the last term we assume that the adaptive filters are long enough to remove all the correlation between $e_o(n)$ and $\mathbf{x}(n-\Delta)$. In this case we can rewrite the above as

$$\mathbf{K}_{i,k,l}(n) = (\mathbf{I} - \mu_i \mathbf{R}_x) \mathbf{K}_{i,k,l}(n-1) (\mathbf{I} - \mu_k \mathbf{R}_x) + \mu_i \mu_k J_{min} \mathbf{R}_x, \quad (63)$$

where $J_{min} = E[|e_o|^2]$ is the minimum mean square error produced by the corresponding Wiener filter.

We now assume that the signal to noise ratio is low so that the input signal is dominated by the white noise process $v(n)$. In this case we can approximate the correlation matrix of the input process by unit matrix as

$$\mathbf{R}_x \approx \sigma_v^2 \mathbf{I}, \quad (64)$$

where σ_v^2 is the noise variance. Later in the simulation study we will see that the theory developed this way actually works well with quite moderate signal to noise ratios. Then substituting (64) into (63) yields

$$\mathbf{K}_{i,k,l}(n) = (\mathbf{I} - \mu_i \sigma_v^2 \mathbf{I}) \mathbf{K}_{i,k,l}(n-1) (\mathbf{I} - \mu_k \sigma_v^2 \mathbf{I}) + \mu_i \mu_k J_{min} \sigma_v^2 \mathbf{I}. \quad (65)$$

In steady state, when $n \rightarrow \infty$ we have

$$\mathbf{K}_{i,k,l}(\infty) = (1 - \mu_i \sigma_v^2) \mathbf{K}_{i,k,l}(\infty) (1 - \mu_k \sigma_v^2) + \mu_i \mu_k J_{min} \sigma_v^2 \mathbf{I}. \quad (66)$$

Solving the above for $\mathbf{K}_{i,k,l}(\infty)$ we have

$$\mathbf{K}_{i,k,l}(\infty) = \frac{\mu_i \mu_k J_{min}}{\mu_i \sigma_x^2 + \mu_k \sigma_x^2 - \mu_i \mu_k \sigma_v^2} \mathbf{I}. \quad (67)$$

5.2. Second order statistics of line enhancer output signal

As we see from the previous discussion, the correlation matrix of the weight error vector is diagonal. We therefore have that the matrix $\mathbf{K}_{i,k}(l) = E[\tilde{\mathbf{w}}_i(n-1+l)\tilde{\mathbf{w}}_k^H(n-1)]$ has in steady state, when $n \rightarrow \infty$, elements different from zero only alongside the main diagonal and the elements at this diagonal equal to $\frac{\mu_i\mu_k J_{min}}{\mu_i\sigma_x^2 + \mu_k\sigma_x^2 - \mu_i\mu_k\sigma_v^2}$. Substituting $\mathbf{K}_{i,k}(l)$ into (61) we now have that the l -th correlation lag of the output signal is equal to

$$\begin{aligned} r_{i,k}(l) &= \mathbf{w}_o^H E[\mathbf{s}(n-\Delta)\mathbf{s}^H(n-\Delta+l)]\mathbf{w}_o \\ &\quad + \mathbf{w}_o^H E[\mathbf{v}(n-\Delta)\mathbf{v}^H(n-\Delta+l)]\mathbf{w}_o \\ &\quad + r_s(l)N \frac{\mu_i\mu_k J_{min}}{\mu_i\sigma_x^2 + \mu_k\sigma_x^2 - \mu_i\mu_k\sigma_v^2} \\ &\quad + \text{Tr}\{\mathbf{K}_{i,k}(l)E[\mathbf{v}(n-\Delta)\mathbf{v}^H(n-\Delta+l)]\}, \end{aligned} \quad (68)$$

where $r_s(l)$ is the l -th autocorrelation lag of the input signal $s(n)$.

As the noise v has assumed to be white, the matrix $E[\mathbf{v}(n-\Delta)\mathbf{v}^H(n-\Delta+l)]$ has nonzero elements σ_v^2 only along the l -th diagonal and the rest of the matrix is filled with zeroes. Then

$$\begin{aligned} r_{i,k}(l) &= \mathbf{w}_o^H E[\mathbf{s}(n-\Delta)\mathbf{s}^H(n-\Delta+l)]\mathbf{w}_o \\ &\quad + \sigma_v^2 \sum_{i=0}^{N-l-1} w_o^*(i)w_o(i+l) \\ &\quad + r_s(l)N \frac{\mu_i\mu_k J_{min}}{\mu_i\sigma_x^2 + \mu_k\sigma_x^2 - \mu_i\mu_k\sigma_v^2} + r_0 \end{aligned} \quad (69)$$

where $r_0 = N\sigma_v^2 \frac{\mu_i\mu_k J_{min}}{\mu_i\sigma_x^2 + \mu_k\sigma_x^2 - \mu_i\mu_k\sigma_v^2}$, if $l=0$ and zero otherwise.

From (4) we see that the autocorrelation lags of the combination output signal $y(n)$ can be composed from its components $r_{i,k}(l)$ as follows

$$\begin{aligned} r(l) &= \lambda(n)^2 E[y_1(n)y_1^*(n+l)] + 2\lambda(n)(1-\lambda(n))E[\text{Re}\{y_1(n)y_2^*(n+l)\}] \\ &\quad + (1-\lambda(n))^2 E[y_2(n)y_2^*(n+l)] \\ &= \lambda(n)^2 r_{1,1}(l) + 2\lambda(n)(1-\lambda(n))\text{Re}\{r_{1,2}(l)\} + (1-\lambda(n))^2 r_{2,2}(l). \end{aligned}$$

The autocorrelation matrix of y is a Toeplitz matrix \mathbf{R} having the autocorrelation lags $r(l)$ along its first row.

Thus far we have evaluated the terms $E[y_i(n)y_k^*(n+l)]$, what remains is to find an expression for the steady state combination parameter $\lambda(\infty)$. For this purpose we can use (15), not-

ing that $y_i(n) = \mathbf{w}_i^H(n-1)\mathbf{x}(n-\Delta)$. All the terms in the expression (15) are similar and we need to evaluate

$$\gamma_{ik} = E[\tilde{\mathbf{w}}_i^H(n-1)\mathbf{x}(n-\Delta)\mathbf{x}^H(n-\Delta)\tilde{\mathbf{w}}_k(n-1)]. \quad (70)$$

Due to the independence assumption we can rewrite (70) using the properties of trace operator as

$$\begin{aligned} \gamma_{ik} &= Tr\{E[\mathbf{x}(n-\Delta)\mathbf{x}^H(n-\Delta)\tilde{\mathbf{w}}_k(n-1)\tilde{\mathbf{w}}_i^H(n-1)]\} \\ &= Tr\{\mathbf{R}_x E[\tilde{\mathbf{w}}_k(n-1)\tilde{\mathbf{w}}_i^H(n-1)]\} = Tr\{\mathbf{R}_x \mathbf{K}_{i,k,0}(n-1)\}. \end{aligned} \quad (71)$$

We are now ready to find $\lambda(\infty)$ by substituting (71) and (67) into (15).

The power spectrum of the output process $y(n)$ is given by

$$P(f) = \lim_{K \rightarrow \infty} \frac{1}{K} E[|Y_K(f)|^2] = \sum_{l=-\infty}^{\infty} r(l) e^{-j2\pi lf}, \quad (72)$$

where $Y_K(f)$ is the length K discrete Fourier transform of the signal $y(n)$ and f is the frequency. There is a number of methods to compute an estimate of the power spectrum from the correlation matrix of a signal. In this paper we have used the Capon method [20].

$$\hat{P}(f) = \frac{K}{\mathbf{a}^H(f) \mathbf{R}^{-1} \mathbf{a}(f)}, \quad (73)$$

where $\mathbf{a}(f) = [1 e^{-j2\pi f} \dots e^{-j2\pi(M-1)f}]^T$ and \mathbf{R} is the $K \times K$ Toeplitz correlation matrix of the signal of interest. The Capon method was chosen because the signals we are interested in are sine waves in noise and the Capon method gives a more distinct spectrum estimate than the Fourier transform based methods in this situation.

6. Simulation Results

In this Section we present the results of our simulation study.

In order to obtain a practical algorithm, the expectation operators in both numerator and denominator of (7) have been replaced by exponential averaging of the type

$$P_{av}(n) = (1 - \gamma)P_{av}(n-1) + \gamma p(n), \quad (74)$$

where $p(n)$ is the quantity to be averaged, $P_{av}(n)$ is the averaged quantity and γ is the smoothing parameter. The averaged quantities were then used in (7) to obtain λ . The curves

shown in the Figures to follow are averages over 100 independent trials. We often show the simulation results and the theoretical curves in the same Figures. In several cases the curves overlap and are therefore indistinguishable.

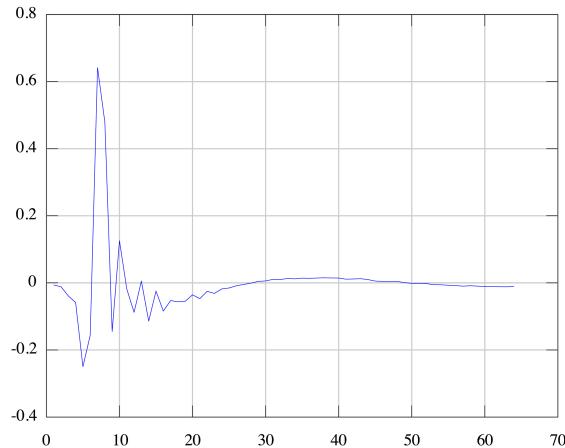


Figure 5. The true impulse response.

6.1 System Identification

We have selected the sample echo path model number one shown in Figure 5 from [10], to be the unknown system to identify and combined two 64 tap long adaptive filters.

In the Figures below the noisy blue line represents the simulation result and the smooth red line is the theoretical result. The curves are averaged over 100 independent trials.

In the system identification example we use Gaussian white noise with unity variance as the input signal. The measurement noise is another white Gaussian noise with variance $\sigma_v^2=10^{-3}$. The step sizes are $\mu_1=0.005$ for the fast adapting filter and $\mu_2=0.005$ for the slowly adapting filter. Figure 6 depicts the evolution of EMSE in time. One can see that the system converges fast in the beginning. The fast convergence is followed by a stabilization period between sample times 1000-7000 followed by another convergence to a lower EMSE level between the sample times 8000-12000. The second convergence occurs when the mean squared error of the filter with small step size surpasses the performance of the filter with large step size. One can observe that the there is a good accordance between the theoretical and the simulated curves so that the theoretical and the simulation curves are difficult to distinguish from each other.

The combination parameter λ is shown in Figure 7. At the beginning, when the fast converging filter gives smaller EMSE than the slowly converging one, λ is close to unity. When the

slow filter catches up the fast one λ starts to decrease and obtains a small negative value at the end of the simulation example. The theoretical and simulated curves fit well.

In the Figure 8 we show the time evolution of mean square deviation of the combination in the same test case. Again one can see that the theoretical and simulation curves fit well.

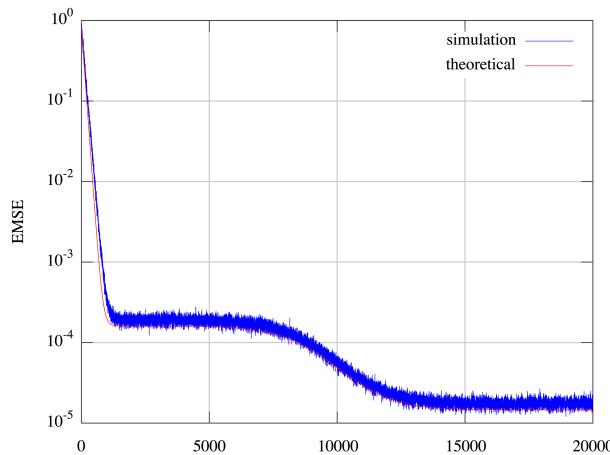


Figure 6. Time-evolutions of EMSE with $\mu_1 = 0.005$ and $\mu_2 = 0.0005$ and $\sigma_v^2 = 10^{-3}$.

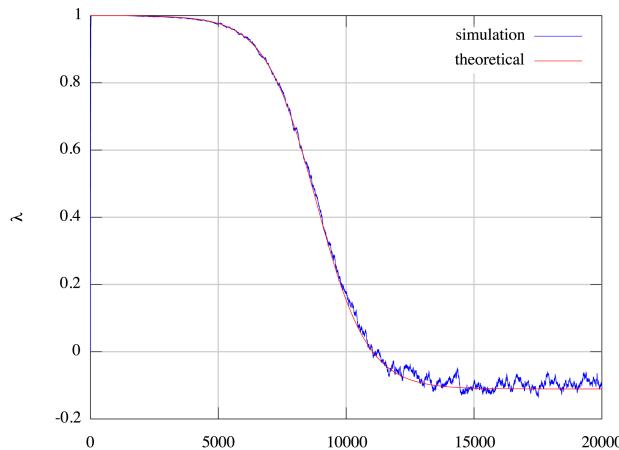


Figure 7. Time-evolutions of λ with $\mu_1 = 0.005$ and $\mu_2 = 0.0005$ and $\sigma_v^2 = 10^{-3}$.

6.2 Adaptive beamforming

In the beamforming example we have used a 8 element linear array with half wave-length spacing. The noise power is 10^{-4} in this simulation example. The useful signal which is 10 dB stronger than the noise arrives from the broadside of the array. There are three strong interferers at -35° , 10° and 15° with $SNR_1 = 33$ dB and $SNR_2 = SNR_3 = 30$ dB respectively. The step sizes of the adaptive combination are $\mu_1 = 0.05$ and $\mu_2 = 0.006$.

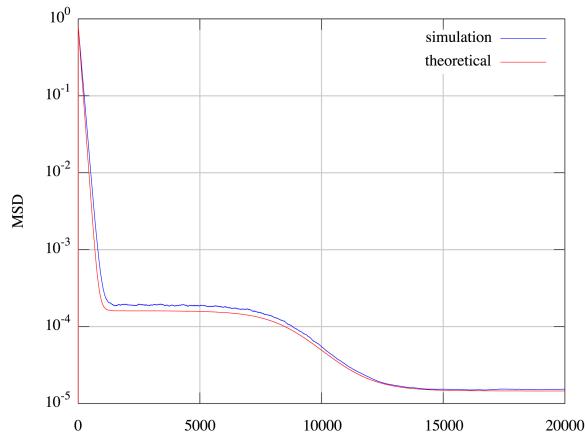


Figure 8. Time-evolutions of MSD with $\mu_1 = 0.005$ and $\mu_2 = 0.0005$ and $\sigma_v^2 = 10^{-3}$.

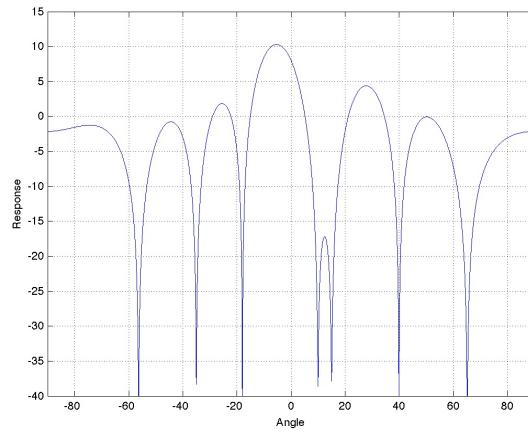


Figure 9. The antenna pattern.

The steady state antenna pattern is shown in Figure 6. One can see that the algorithm has formed deep nulls in the directions of the interferers while the response in the direction of the useful signal is equal to the number of antennas i.e. 8.

The evolution of EMSE in this simulation example is depicted in Figure 10. One can see a rapid convergence at the beginning of the simulation example. Then the EMSE value stabilizes at a certain level and after a while a second convergence occurs. The dashed red line is the theoretical result and the solid blue line is the simulation result. One can see that the two overlap and are indistinguishable in black and white print.

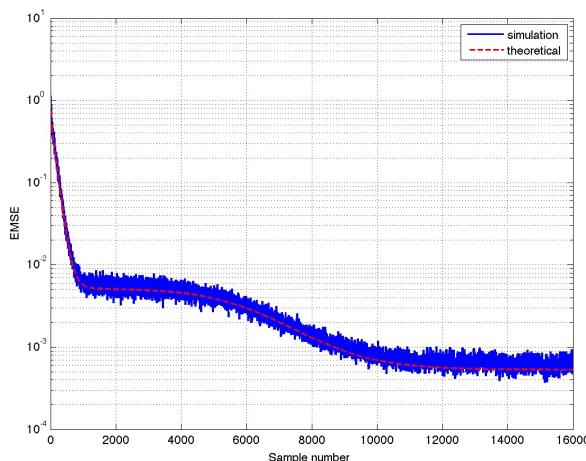


Figure 10. Time evolution of EMSE.

The time evolution of λ for this simulation example is shown in Figure 11. At the beginning λ is close to one forcing the output signal of the fast adapting filter to the output of the combination. Eventually the slow filter catches up with the fast one and λ starts to decrease obtaining at the end of the simulation example a small negative value so that the output signal is dominated by the output signal of the slowly converging filter. One can see that the simulation and theoretical curves for λ evolution are close to each other.

The signal to interference and noise ratio evolution is show in Figure 12. One can see a fast improvement of SINR at the beginning of the simulation example followed by a stabilization region. After a while a new region of SINR improvement occurs and finally the SINR stabilizes at an improved level. Again the theoretical result matches the simulation curve well making the curves indistinguishable in black and white print.

6.3. Adaptive Line Enhancer

In order to illustrate the adaptive line enhancer application we have used length $K = 32$ correlation sequences to form $K \times K$ correlation matrices for the Capon method. The narrow band signals were just sine waves in our simulations.

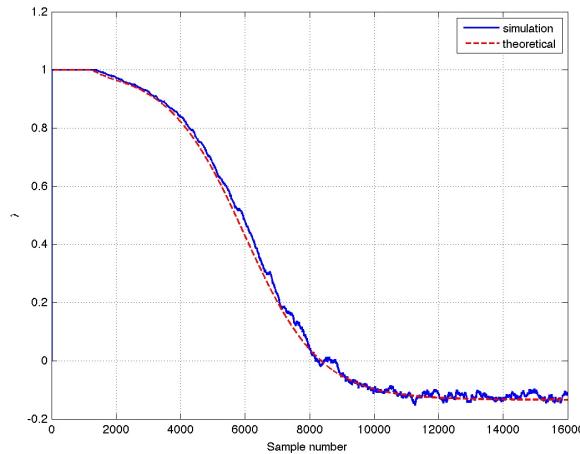


Figure 11. Time evolution of λ .

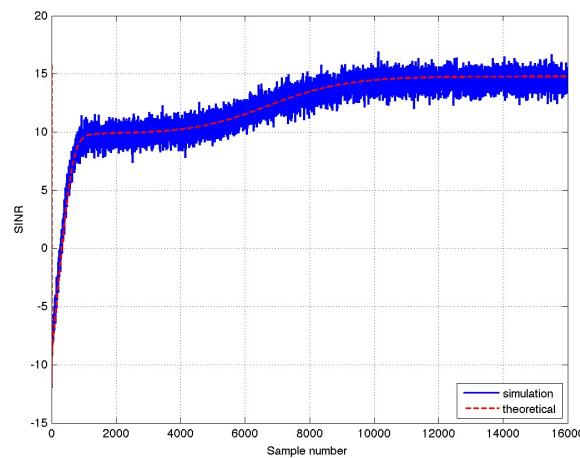


Figure 12. Time evolution of SINR.

The input signal consist of three sine waves and additive noise with unity variance. The sine waves with frequencies 0.1 and 0.4 have amplitudes equal to one and the third sine wave with normalized frequency 0.25 has amplitude equal to 0.5. The spectra of the input signal $x(n)$ and the output signal $y(n)$ are shown in Figure 13. The step sizes used were $\mu_1 = 0.5$ and $\mu_2 = 0.05$, the filter is $N=16$ taps long and the delay $\Delta = 10$.

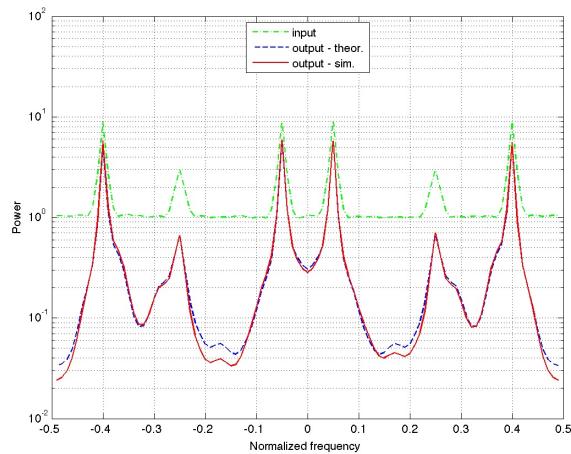


Figure 13. Line enhancer output signal spectrum.

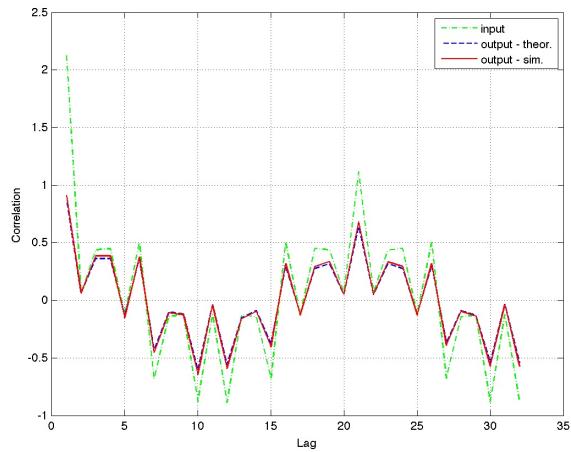


Figure 14. Line enhancer output signal autocorrelation.

In Figure 14 we show the correlation functions of input and output signals in the second simulation example. We can see that the theoretical correlation matches the correlation computed from simulations well.

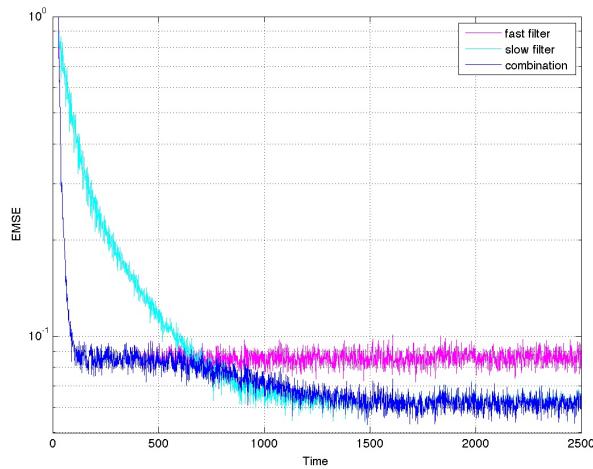


Figure 15. Evolution of EMSE of the two component filters and the combination in time.

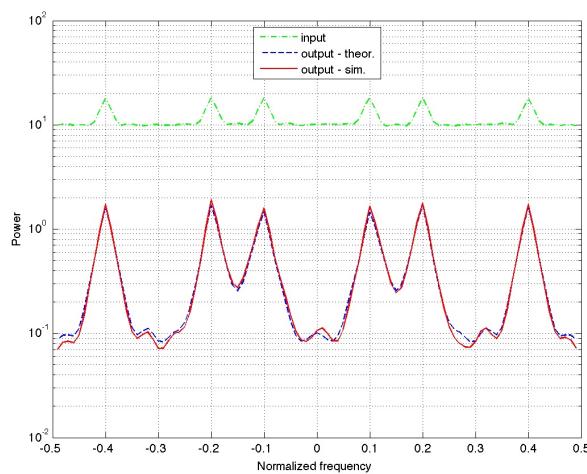


Figure 16. Line enhancer output signal spectrum.

The evolution of the excess mean square error of the combination together with that of the individual filters is shown in Figure 15. We see the fast initial convergence, which is due to the fast adapting filter. After the initial convergence there is a period of stabilization followed by a second convergence between the sample times 500 and 1500, when the error power of the slowly adapting filter bypasses that of the fast one.

In our final simulation example (Figure 16) we use three unity amplitude sinusoids with frequencies 0.1, 0.2 and 0.4. We have increased the noise variance to 10 so that the noise power is 20 times the power of each of the individual sinusoids. The adaptive filter is $N=16$ taps long and the delay $\Delta = 10$. The step sizes of the individual filters in the combined structure are $\mu_1 = 0.5$ and $\mu_2 = 0.005$. One can see that even in such noisy conditions there is still a reasonably good match between the theoretical and simulation results.

7. Conclusions

In order to make the LMS type adaptive algorithm work properly one has to select a suitable step size. The step size has to be smaller than $\frac{2}{\omega_{max}}$, where ω_{max} is the largest eigenvalue of the input signal autocorrelation matrix, in order to guarantee stability of the algorithm. Given that the stability condition is fulfilled a large step size allows the algorithm to initially converge fast but the mean square error in steady state remains large too. On the other hand if one selects a small step size it is possible to achieve a small steady state error but the initial convergence speed of the algorithm is reduced in this case. In this Chapter we have investigated the combination of two adaptive filters, which is a new and interesting way of achieving fast initial convergence and low steady state error of an adaptive filter at the same time, solving thus the trade-off one has in step size selection. We were looking at three applications of the technique - system identification, adaptive beamforming and adaptive line enhancing. In all three applications we saw that the combination worked as expected allowing the algorithm to converge fast to a certain level and then after a while providing a second convergence to a lower mean square error value.

Author details

Tõnu Trump*

Address all correspondence to: tonu.trump@gmail.com

Tallinn University of Technology, Estonia

References

- [1] Aboulnasr, T., & Mayyas, K. (1997). A robust variable step-size LMS-type algorithm: Analysis and simulations. *IEEE Transactions on Signal Processing*, 45, 631-639.
- [2] Arenas-Garcia, J., Figueiras-Vidal, A. R., & Sayed, A. H. (2006). Mean-square performance of convex combination of two adaptive filters. *IEEE Transactions on Signal Processing*, 54, 1078-1090.
- [3] Armbruster, W. (1992). Wideband Acoustic Echo Canceller with Two Filter Structure, Signal Processing VI, Theories and Applications. VanderwalleJBoiteR.MooneM.OosterlinckA., Elsevier Science Publishers B.V.
- [4] Azpicueta-Ruiz, L. A., Figueiras-Vidal, A. R., & Arenas-Garcia, J. (2008a). A new least squares adaptation scheme for the affine combination of two adaptive filters. *Proc. IEEE International Workshop on Machine Learning for Signal Processing*, Cancun, Mexico, 327-332.
- [5] Azpicueta-Ruiz, L. A., Figueiras-Vidal, A. R., & Arenas-Garcia, J. (2008b). A normalized adaptation scheme for the convex combination of two adaptive filters. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, 3301-3304.
- [6] Bershad, N. J., Bermudez, J. C., & Tourneret, J. H. (2008). An affine combination of two LMS adaptive filters - transient mean-square analysis. *IEEE Transactions on Signal Processing*, 56, 1853-1864.
- [7] Fathiyan, A., & Eshghi, M. (2009). Combining several PBS-LMS filters as a general form of convex combination of two filters. *Journal of Applied Sciences*, 9, 759-764.
- [8] Harris, R. W., Chabries, D. M., & Bishop, F. A. (1986). Variable step (vs) adaptive filter algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34, 309-316.
- [9] Haykin, S. (2002). Adaptive Filter Theory, Fourth Edition,. Prentice Hall.
- [10] ITU-T Recommendation G.168 Digital Network Echo Cancellers. (2009). *ITU-T*.
- [11] Kim, K., Choi, Y., Kim, S., & Song, W. (2008). Convex combination of affine projection filters with individual regularization. *Proc. 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, Shimonoseki, Japan, 901-904.
- [12] Kwong, R. H., & Johnston, E. W. (1992). A variable step size LMS algorithm. *IEEE Transactions on Signal Processing*, 40, 1633-1642.
- [13] Mandic, D., Vayanos, P., Boukis, C., Jelfs, B., Goh, S., I., Gautama, T., & Rutkowski, T. (2007). Collaborative adaptive learning using hybrid filters. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Honolulu, Hawaii, 901-924.

- [14] Martinez-Ramon, M., Arenas-Garcia, J., Navia-Vazquez, A., & Figueiras-Vidal, A. R. (2002). An adaptive combination of adaptive filters for plant identification. *Proc. 14th International Conference on Digital Signal Processing*, Santorini, Greece, 1195-1198.
- [15] Mathews, V. J., & Xie, Z. (1993). A stochastic gradient adaptive filter with gradient adaptive step size. *IEEE Transactions on Signal Processing*, 41, 2075-2087.
- [16] Ochiai, K. (1977). Echo canceller with two echo path models. *IEEE Transactions on Communications*, 25, 589-594.
- [17] Sayed, A. H. (2008). Adaptive Filters. John Wiley and sons.
- [18] Shin, H. C., & Sayed, A. H. (2004). Variable step-size NLMS and affine projection algorithms. *IEEE Signal Processing Letters*, 11, 132-135.
- [19] Silva, M. T. M., Nascimento, V. H., & Arenas-Garcia, J. (2010). A transient analysis for the convex combination of two adaptive filters with transfer of coefficients. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, TX, USA, 3842-3845.
- [20] Stoica, P., & Moses, R. (2005). Spectral Analysis of Signals. Prentice Hall.
- [21] Trump, T. (2009). An output signal based combination of two NLMS adaptive algorithms. *Proc. 16th International Conference on Digital Signal Processing*, Santorini, Greece.
- [22] Trump, T. (2011a). analysisOutput signal based combination of two NLMS adaptive filters - transient, *Proceedings of the Estonian Academy of Sciences*, 60(4), 258-268.
- [23] Trump, T. (2011b). Output statistics of a line enhancer based on a combination of two adaptive filters. *Central European Journal of Engineering*, 1, 244-252.
- [24] Zhang, Y., & Chambers, J. A. (2006). Convex combination of adaptive filters for a variable tap-length LMS algorithm. *IEEE Signal Processing Letters*, 10, 628-631.

