Joint gradient-based time-delay estimation and adaptive minimum mean-squared-error filtering

Estimation conjointe du délai et filtrage adaptatif basés sur la méthode de gradient du minimum de l'erreur quadratique moyenne

Daniel Boudreau, *Communications Research Centre, Directorate of Satellite Communications, Ottawa, Ont., K2H 8S2 and*
Peter Kabal, *Department of Electrical Engineering, McGill University, Montréal, Que., H3A 2A7 and INRS-Télécommunications, Université du Québec, Verdun, Que., H3E 1H6*

A general estimation model is defined in which two observations are available; one is a noisy version of the transmitted signal, while the other is a noisy filtered and delayed version of the same transmitted signal. The time-varying delay and the filter are unknown quantities that must be estimated. A joint estimator is proposed. It is composed of an adaptive delay element in conjunction with a transversal adaptive filter. The same error signal is used to adjust the delay element and the filter such that the minimum mean squared error is attained. Two joint gradient-based adaptation algorithms are studied. The joint steepest-descent (SD) algorithm is first investigated. The possibility of a multitude of stable solutions is established and a condition of convergence is presented. A stochastic implementation of the joint SD algorithm, under the form of a joint least-mean-square (LMS) algorithm, is then presented. It is analysed in terms of convergence in the mean and in the mean square of both the delay estimate and the adaptive filter weight vector estimate. The conditions of convergence of the joint LMS algorithm are established as a function of the power spectral densities of the observed signals and the minimum mean squared error. The joint LMS algorithm is simulated under various conditions and it is shown that the adaptive delay element is very effective in reducing the mean squared error at the output of a long adaptive filter coping with two asynchronous inputs.

Un modèle mathématique constitué de deux observations est tout d'abord défini. L'une des observations est une version corrompue du signal transmis, alors que l'autre est une version corrompue, filtrée et décalée de ce même signal. Une estimation conjointe du délai et du filtre est proposée. L'estimateur conjoint est formé d'un délai adaptatif et d'un filtre adaptatif à réponse impulsionnelle finie. Le même signal d'erreur est utilisé pour corriger le délai et le filtre de telle sorte que le minimum de l'erreur quadratique moyenne soit atteint. Deux algorithmes d'adaptation basés sur la méthode de gradient sont proposés. L'algorithme conjoint à descente maximale est étudié en premier lieu. La possibilité de solutions multiples est établie et une condition de convergence est présentée. Une forme utilisant l'algorithme des moindres carrés moyens est ensuite considérée. Cette forme est analysée en termes de convergence de l'espérance et en termes de convergence de la valeur quadratique moyenne, tant pour l'estimé de délai que pour l'estimé du filtre adaptatif. Ces conditions de convergence sont établies en fonction de la densité de puissance spectrale des signaux observés et du minimum de l'erreur quadratique moyenne. L'algorithme des moindres carrés moyens est simulé sous diverses conditions. L'efficacité d'un délai adaptatif est démontrée, pour réduire l'erreur quadratique moyenne à la sortie d'un filtre adaptatif affichant une longue réponse impulsionnelle.

I. Introduction

The problem of estimating the time delay between two continuous versions of the *same* signal, each one corrupted by uncorrelated noise components, has been the subject of many research efforts in recent years. The maximum likelihood estimator for the unknown delay has been derived for a static [1]-[2] and a time-varying delay [3]. Closed-loop adaptive techniques using the minimum mean-squared-error (MMSE) or the least squares (LS) criteria have also been proposed. In these cases, the estimator structure is such that one signal is processed by an adaptive system for which the output is compared to the other signal, with the error being used to adapt a conventional adaptive transversal filter or an adaptive delay element.

In this paper, we consider a signal model that generalizes somewhat the conventional model by allowing frequency-dependent attenuation in the delayed path. We also specifically consider discrete-time signals and systems. This work finds some applications in system modelling problems, where the unknown system often has an impulse response that can be modelled as a pure time delay in series with a linear filter. This can occur in noise or echo cancellation, digital communication or geophysical exploration.

We study a joint adaptive estimator which is composed of an adaptive delay element working in conjunction with an adaptive filter. The adaptive delay element attempts to model the reference delay and can take any real value. The addition of this adaptive delay to the usual adaptive filtering operations can improve the conventional adaptive parameter estimation techniques that would otherwise be of limited usefulness, especially in the case where the main adaptive filter input and its reference signal decorrelate with time. A simple adaptive filter has the potential to model both the reference delay and the reference filter, since the overall function can be approximated by an FIR filter with the proper number of taps. This approach is inefficient in the sense that the reference delay is modelled by a shift in the adaptive filter impulse response. For a fixed filter order, this shift may result in an error that is larger than the error corresponding to perfect modelling. An additional adaptive delay estimation algorithm, specifically designed to track the reference delay variations, allows a better impulse response centring and the use of an adaptive filter with a smaller order.

In this paper, we present the results of an analysis for the joint adaptive delay and filter structure based on the MMSE performance index. A joint steepest-descent (SD) algorithm and a joint least-mean-square (LMS) algorithm are investigated. The principal contributions of this paper are the generalization of existing gradient-based time-delay estimation without reference filtering, as proposed in [4], and the analysis of a new joint algorithm for the synchronization of the input and the reference signals used by an adaptive filter. Our joint algorithms are not based on the assumption that the input signal and the reference signal fed to an adaptive filter are sampled in the same clock period. They also allow the tracking of time-varying delays in the reference path by a process separated from the adaptive filter, which is itself free to perform the task of modelling the linear reference filter or its inverse.

The paper is organized as follows. In the next section, the minimum mean-squared-error function is considered in general terms, as a joint function of the two estimates. The form of this function allows one to draw some conclusions about the general convergence behaviour of the joint algorithms. The presence of a multitude of minima in the objective function is discussed. Then the joint steepest-descent algorithm is studied in section III, where the conditions for convergence to a local minimum of the mean-squared-error function are given. The joint least-mean-square algorithm is investigated in section IV. Analytical results for the convergence in the mean and in the mean square, for both the adaptive delay estimator and the adaptive filter weight vector, are presented. Finally some experimental results are given, in order to complete the presentation.

II. General minimum mean-squared-error function

We consider a situation that generalizes the conventional model used in delay estimation by allowing frequency-dependent attenuation in the delayed path. We also specifically consider discrete-time signals and systems. The corresponding model, where $z_1(n)$ and $z_2(n)$ are the two observed signals, is of the form

$$z_1(n) = s(n) + v_1(n), \quad (1.1)$$

$$z_2(n) = h(n) \otimes s(nT - D_n) + v_2(n), \quad (1.2)$$

where n is the discrete-time index, $s(n)$ is the transmitted signal, D_n is a time delay (possibly time-varying), and $h(n)$ is the impulse response of a linear filter which is applied on a delayed-by-D_n version of the signal $s(n)$. The discrete-time noise processes, $v_1(n)$ and $v_2(n)$, are zero-mean and stationary and are assumed to be uncorrelated with each other as well as with $s(n)$. The operator $\otimes$ is the convolution operator. Note that the time-varying reference delay, D_n, is not limited to an integer number of sampling periods and can take any real value. All the discrete-time signals are assumed to be sampled versions, with sampling period T, of continuous-time signals that are strictly band-limited to the frequency range $-1/2T < f < 1/2T$. A block diagram corresponding to the mathematical model of (1.1)-(1.2) is illustrated in Fig. 1. Note that the case in which the delay, D_n, follows the linear filter is also of interest, but is not considered in this paper. See [5] for more details.

In the joint estimation problem considered in this paper, it is required that both the time-varying delay, D_n, and the reference filter, $h(n)$, or its inverse, $h^{-1}(n)$, be estimated[†]. The adaptive filter used to estimate $h(n)$ or $h^{-1}(n)$ is a transversal filter, with a weight vector w_n of length M.

In joint MMSE delay estimation and adaptive filtering, the mean-squared-error surface is searched by both the adaptive filter estimation algorithm and the delay estimation algorithm. A system identification configuration takes the form given in Fig. 2.

In general, the output of the adaptive branch can be defined as $y(n,d_n)$, where the dependence on the adaptive delay is explicitly shown. The reference signal, $r(n)$, is defined to be one of the two observed signals $z_1(n)$ or $z_2(n)$. Then the error signal, $e(n,d_n)$, is defined as

$$e(n,d_n) = r(n) - y(n,d_n), \quad (2)$$

and the MSE function, at time n, is

$$\xi_n = E\left[\left|e(n,d_n)\right|^2\right]. \quad (3)$$

The joint estimation can be thought of as taking place in a vector space made of a weight vector subspace and a delay subspace. The two subspaces are not orthogonal, which implies that the two estimation processes are not independent (because the adaptive filter can model a reference delay). In order to obtain an expression for the MSE function, define as $u(n)$ the input to the adaptive *branch*. The signal $u(n)$ is therefore the generic representation of the observation that must be adaptively processed. It can be $z_1(n)$, as in Fig. 2, or $z_2(n)$ if one wants to estimate the inverse of the reference branch (inverse filtering). The output of the adaptive branch, $y(n,d_n)$, is assumed to be given by

$$y(n,d_n) = \boldsymbol{w}_n^H \boldsymbol{u}_n, \quad (4)$$

where the superscript H denotes complex conjugate transpose. The vector $\boldsymbol{u}_n$ is the vector of delayed input samples, stored at iteration n, in the adaptive filter delay line; i.e.,

$$\boldsymbol{u}_n = \left[u(nT - d_n), u(nT - T - d_{n-1}), \ldots, u(nT - MT + T - d_{n-M+1})\right]^T. \quad (5)$$

The input-signal autocorrelation matrix and the cross-correlation vector between this input and the reference signal are then expressed as

$$\boldsymbol{R}_n = E\left[\boldsymbol{u}_n \boldsymbol{u}_n^H\right] \quad (6)$$

and

$$\boldsymbol{p}_n = E\left[\boldsymbol{u}_n r^*(n)\right]. \quad (7)$$

The MSE function is represented by either one of the following equivalent equations,

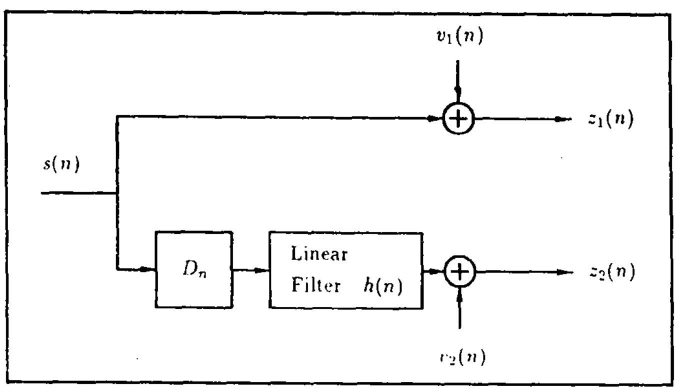

Figure 1: *Mathematical signal model.*

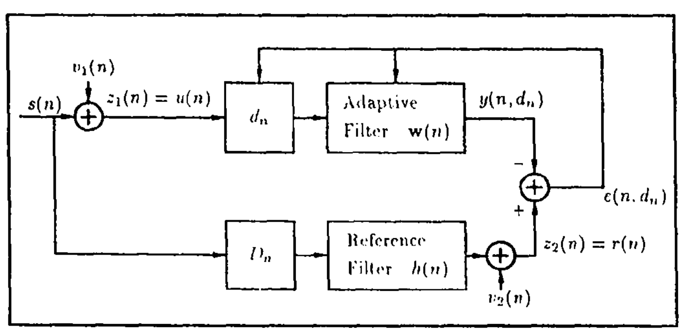

Figure 2: *Systems identification configuration with the delays before the filters.*

† Note that the inverse of any linear filtering operation $h(n)$ is denoted as $h^{-1}(n)$. Therefore $h(n) \otimes h^{-1}(n) = \delta(n)$, where $\delta(n)$ is the unit-sample sequence [6].

$$\xi_n = \begin{cases} \phi_{rr}(n,n) + \boldsymbol{w}_n^H \boldsymbol{R}_n \boldsymbol{w}_n - 2\mathrm{Re}\left[\boldsymbol{w}_n^H \boldsymbol{p}_n\right] \\ \phi_{rr}(n,n) + \phi_{yy}(n,n) - 2\mathrm{Re}\left[\phi_{yr}(n,n)\right] \end{cases} \tag{8}$$

where Re[•] is the real value operator, $\phi_{rr}(n, m)$ and $\phi_{yy}(n, m)$ are the autocorrelation functions of the reference signal $r(n)$ and the adaptive branch output respectively, and $\phi_{yr}(n, m)$ is the cross-correlation function between this output and the delayed reference signal.

The MSE espressions reflect the nature of the joint estimator operation. In the weight vector subspace, associated with the first equation of (8), the MSE function is a quadratic surface [7]. The one-dimensional delay subspace is naturally linked to the correlation fuctions of the second equation of (8). The MSE function is not, in general, unimodal with respect to d_n. In order to see this, note that ξ_n depends on correlation functions that vary according to the adaptive filter and the reference filter, as well as to the autocorrelation function of the signals $u(n)$ and $r(n)$. All of these functions are multimodal with respect to their time argument, a characteristic which in turn causes the MSE function to behave similarly with respect to d_n and produces a multitude of local extrema.

This behaviour causes a problem in the search for the minimum of ξ_n with respect to d_n. In closed-loop estimation, this phenomenon leads to false lock problems, as in phase-locked loops. These problems are generally solved by designing an acquisition procedure in which the delay estimate is varied until the algorithm falls in its tracking region, near the MSE global minimum. Once in tracking mode, the estimation algorithm can compute the derivative of the MSE function with respect to the delay value and generate a correcting signal that brings the loop into lock. For the joint delay and adaptive filtering algorithm, it is possible to use an acquisition procedure based on the *least squares* criterion and known as the *optimum lag* algorithm [8]. In this algorithm, the least sum of squared errors is computed for a representative set of delay values, and the delay and filter impulse response corresponding to the global minimum are selected. This algorithm is computationally involved, but it is well suited for an acquisition phase.

In the following, when we study the joint algorithm in tracking or steady-state conditions, we assume that such an acquisition procedure has caused the joint algorithm to lock near the minimum of the MSE function.

III. The joint steepest-descent algorithm

The joint delay estimation and adaptive filtering steepest-descent algorithm is composed of the usual SD adaptive filter algorithm, of the form [9]

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n - \mu \nabla_{\boldsymbol{w}_n} \xi_n \tag{9}$$

and of the SD adaptive delay algorithm [4]

$$d_{n+1} = d_n - \alpha \frac{\partial \xi_n}{\partial d_n}, \tag{10}$$

where μ and α are small positive constants and $\nabla_{\boldsymbol{w}_n}\xi_n$ represents the gradient of ξ_n with respect to $\boldsymbol{w}_n$.

The delay-estimation part of the joint algorithm can be studied by using a truncated Taylor expansion of the MSE function around a certain delay value $d_n = \vartheta_n$, where the MSE function is minimum. Keeping only the first three terms of the series, we get

$$\begin{aligned} \xi\{d_n, \boldsymbol{w}_n\} &\approx \xi_n\{\vartheta_n, \boldsymbol{w}_n\} \\ &+ (d_n - \vartheta_n)\dot{\xi}_n\{\vartheta_n, \boldsymbol{w}_n\} + 1/2(d_n - \vartheta_n)^2 \ddot{\xi}_n\{\vartheta_n, \boldsymbol{w}_n\}, \end{aligned} \tag{11}$$

where the dot denotes a derivative with respect to the delay value d_n. This approximation is used in order to linearize the delay estimation algorithm. The linearized SD algorithm is obtained by combining (10) and (11), and assuming that ϑ_n is a minimum of $\xi_n\left(\dot{\xi}_n\{\vartheta_n, \boldsymbol{w}_n\} = 0\right)$. It is given by

$$d_{n+1} = \left(1 - \alpha\ddot{\xi}_n\{\vartheta_n, \boldsymbol{w}_n\}\right)d_n + \alpha\vartheta_n\ddot{\xi}_n\{\vartheta_n, \boldsymbol{w}_n\}, \tag{12}$$

which models the behaviour of a first-order delay-lock loop [10].

A. Convergence of the joint SD algorithm

A necessary condition for a specific d_n and $\boldsymbol{w}_n$ to be a stationary solution of the joint SD algorithms is that both of the following equations be satisfied [11]:

$$\begin{cases} \nabla_{\boldsymbol{w}_n} \xi_n = 0 \\ \dfrac{\partial \xi_n}{\partial d_n} = 0. \end{cases} \tag{13}$$

Note that the first equation of (13) is in fact a *necessary and sufficient* condition for convergence. This is so because ξ_n is quadratic with respect to $\boldsymbol{w}_n$, implying that there is a unique minimum in $\boldsymbol{w}_n$ for a given value d_n. When the first equation of (13) is satisfied, this unique solution is attained, and any further modifications of d_n will increase ξ_n. This is the case because the adaptive filter models both the relative delay and the reference filter in the minimum MSE sense. Then, this solution corresponds also to a minimum with respect to d_n. The sufficiency of the condition is due to the uniqueness of the minimum with respect to $\boldsymbol{w}_n$.

If the adaptation factors μ and α are chosen sufficiently small, the process always reaches a limit point [12]. The next proposition gives a condition on μ and α that ensures convergence of the joint algorithms under specific conditions. This condition is derived in [11] for joint carrier phase acquisition and adaptive equalization as encountered in digital communications. It is reformulated here for the problem at hand. This condition is general in that it establishes the stability range for the two adaptation factors such that the MSE is reduced at each iteration, when the two adaptive processes are alternated. It is also important because it confirms that, with the right parameters, the joint SD algorithm converges eventually to a stationary point (i.e., (13) is satisfied).

Proposition 1. *Let the set of positive integers be divided arbitrarily into two disjoint subsets* κ_1 *and* κ_2*, each containing an infinite number of positive integers. Let* $\alpha_n = 0$ *when* $n \in \kappa_1$*, and* $\mu_n = 0$ *when* $n \in \kappa_2$*. Let* $\lambda_{max}(n)$ *be the maximum eigenvalue of the signal autororrelation matrix* $\boldsymbol{R}_n$ *and let* ϑ_n *be the delay value closest to* d_n *for which* $\xi\{d_n, \boldsymbol{w}_n\}$ *is minimum. The MSE will converge to a stationary point if*

$$0 < \mu_n < \frac{1}{\lambda_{max}(n)} \tag{14}$$

for $n \in \kappa_1$*, and*

$$0 < \alpha_n < 2\left[\frac{\partial^2}{\partial d_n^2}\xi\{\vartheta_n, \boldsymbol{w}_n\}\right]^{-1} \tag{15}$$

for $n \in \kappa_2$.

The formal proof of this proposition is given in [5]. It is easily seen that when $\alpha_n = 0$, the usual SD adaptive filtering conditions apply and (14) is the conventional condition for convergence. When $\mu_n = 0$, the MSE function evaluated at $d_n = \vartheta_n$ is constant and condition (15) guarantees the convergence of the linearized SD algorithm given in (12).

This proposition states that d_n and $\mathbf{w}_n$ may be adjusted in any alternating fashion, and the MSE will converge to a stationary point if μ_n satisfies (14) during the adjustment of $\mathbf{w}_n$, and α_n satisfies (15) during the adjustment of d_n. The above condition is important because it confirms that, with the right parameters used *in alternation*, the MSE is reduced at each iteration and the joint SD algorithm converges eventually to a stationary point.

B. Steady-state delay estimation properties of the algorithm

In this subsection, we briefly study the system and signal components that directly influence the stability and the delay tracking behaviour of the joint SD algorithm. In order to proceed, we assume that the reference filter $h(n)$ is time-invariant, that the signal-to-noise ratios (SNRs) are high and that the adaptive filter has fully adapted to $h(n)$ and is at least as long as this impulse response. These assumptions imply that in steady state, the i^{th} adaptive filter coefficient, w_{ni}, at iteration n, is approximately

$$w^{*}_{ni} \approx \begin{cases} h(i) & \text{System identification (cancellation)} \\ h^{-1}(i) & \text{Inverse filtering (equalization),} \end{cases} \tag{16}$$

where $h(i)$ is the i^{th} weight of the reference path filter. In the analysis, we use the linearized delay adaptation algorithm of (12) with $\xi\{d_n,\mathbf{w}_n\} = \xi_n$ and $\vartheta_n = D_n$ for the cancellation configuration, and $\vartheta_n = -D_n$ for the equalization structure. Furthermore, in steady state, we assume that $d_n = \pm D_n$ in which case the error is minimum and the corresponding MSE equals the MMSE. Then $\ddot{\xi}_n|d_n = \pm D_n = \ddot{\xi}_{min}$ and is constant with time. From (12), the stability range for α is

$$0 < \alpha < 2/\ddot{\xi}_{min}. \tag{17}$$

The time constant of delay adaptation can be defined by fitting the geometric ratio $1-\alpha\ddot{\xi}_{min}$ to an exponential with time constant τ_{del}:

$$1-\alpha\ddot{\xi}_{min} = e^{-1/\tau_{del}} \approx 1 - 1/\tau_{del}$$

The time constant of delay adaptation is therefore

$$\tau_{del} \approx \frac{1}{\alpha\ddot{\xi}_{min}}. \tag{18}$$

We assume a configuration in which the reference delay, D_n, varies slowly enough so that all the samples in the reference filter delay line are approximately affected by the same delay. Then it can be shown [5] that v_n, the *delay-dependent* term of the MSE function, takes the form

$$v_n^{(C)} \approx -2\,\mathrm{Re}\left[\sum_l \rho_h(l)\phi_{ss}(-lT + D_n - d_n)\right], \tag{19}$$

$$v_n^{(E)} \approx -2\,\mathrm{Re}\left[\phi_{ss}(D_n + d_n)\right], \tag{20}$$

where the superscripts (C) and (E) stand respectively for cancellation and equalization, and $\rho_h(k)$ is the deterministic autocorrelation of the reference filter impulse response and is defined as

$$\rho_h(k) = \sum_i h(k+i)h^*(i). \tag{21}$$

Note that d_n is negative in the equalization case. Comparing (19) and (20), we note that the cancellation configuration is influenced by the form of both the deterministic autocorrelation $\rho_h(n)$ and the input signal autocorrelation $\phi_{ss}(\tau)$, while the equalization configuration is a function of only $\phi_{ss}(\tau)$. Since $\phi_{ss}(\tau)$ exhibits a maximum at $\tau = 0$, $v_n^{(E)}$ has a global minimum at $d_n = -D_n$. In the cancellation scenario, the characteristics of the delay tracking loop are functions of the reference filter, $h(n)$, but because $\rho_h(n)$ has a maximum at $n = 0$, there is a single global minimum corresponding to $d_n = D_n$.

Based on (19), (20) and (17), the following sufficient range of convergence can be computed for the delay gain factor:

$$0 < \alpha < \frac{-1}{\Phi_{max}\,\mathrm{Re}\left[\rho_h''(0)\right]} \qquad \text{Cancellation}$$

and

$$0 < \alpha < \frac{3T^2}{\Phi_{max}\pi^2} \qquad \text{Equalization,}$$

where Φ_{max} is the maximum value of the input signal power spectral density, $\Phi_{ss}(e^{j\omega})$, and the prime denotes the derivative with respect to the continuous-time correlation argument. It is easy to show that $\rho_h''(0)$ is proportional to the square of the reference filter bandwidth, as well as to $\rho_h(0)$. This implies that the convergence properties of the delay SD algorithm are related to the power distribution, across the total bandwidth, of the input signal and the reference filter in the cancellation case, and of the signal only in the equalization case. This is a behaviour essentially similar to the adaptive-filter convergence, which is related to the distribution of the eigenvalues of the input-signal autocorrelation matrix [7].

IV. The joint least-mean-square algorithm

In order to implement the joint steepest-descent algorithm presented in the previous section, the MSE gradient with respect to the adaptive weight vector and the MSE derivative with respect to the adaptive delay both must be estimated. This can be accomplished in various ways; in particular, by approximating the derivatives with difference equations [13], or by approximating the MSE function, $\xi_n = E[\,|e(n,d_n)|^2]$, with the instantaneous squared error, $\gamma_n = |e(n,d_n)|^2$, and by applying the SD algorithm. This last option corresponds to the least-mean-square algorithm [14] and is the subject of this section.

Consider a cancellation configuration where it is assumed that the delay, d_n, propagates instantaneously into the adaptive filter delay line. The adaptive branch output can be expressed as

$$y(n,d_n) = \mathbf{w}_n^H \mathbf{u}(nT - d_n), \tag{22}$$

and $\mathbf{u}(nT - d_n)$ is the delayed vector of input samples defined as

$$\mathbf{u}(nT-d_n) = \left[u(nT-d_n), u(nT-T-d_n), \ldots, u(nT-MT+T-d_n)\right]^T. \tag{23}$$

In the adaptive weight vector subspace, the LMS algorithm that we consider is then given by [7]

$$\mathbf{w}_{n+1} = \mathbf{w}_n + 2\mu e^*(n,d_n)\mathbf{u}(nT-d_n). \tag{24}$$

The error $e(n,d_n)$ is represented by (2). In the adaptive delay subspace, the derivative estimate is given by

$$\hat{\nabla}_{d_n}\xi_n = \frac{\partial|e(n,d_n)|^2}{\partial d_n} = -2\,\mathrm{Re}\left[e^*(n,d_n)\frac{\partial y(n,d_n)}{\partial d_n}\right]. \tag{25}$$

The LMS adaptive delay algorithm is obtained by using the derivative estimate of (25) in the SD adaptive delay algorithm, defined in (10).

The joint LMS algorithm is defined by

$$w_{n+1} = w_n + 2\mu e^*(n, d_n) u(nT - d_n) \quad (26)$$

and

$$d_{n+1} = d_n + 2\alpha \mathrm{Re}\left[e^*(n, d_n) w_n^T \dot{u}(nT - d_n)\right]. \quad (27)$$

In order to ease the derivations, all signals and systems are considered *real* in the analyses. At this point, we are interested in the convergence of the joint LMS algorithm ((26) and (27)) from an arbitrary initial condition.

With the help of the ordinary differential equations (ODE) method [15], it is shown in [5] that the joint LMS algorithm, when the gain factors are of the form $\mu = \alpha = 1/n$, converges to a local minimum of the MSE function, like the exact version of the joint SD algorithm. This result, even if it does not apply directly to algorithm (26)-(27), is important by itself since it shows that if the adaptation factors are chosen sufficiently small, the estimates produced by the algorithm will be, on average, close to a stable stationary point of the MSE function. Furthermore, the above result shows that if the gain factors are small but constant, convergence cannot be attained in the sense that there exists an integer N such that $\theta(n+1) = \theta(n)$ for $N \le n$, but the difference between the parameter estimate and a stable stationary point will be small as n becomes large and can be made smaller by decreasing the gain factors.

A. The joint LMS algorithm in steady state

The quality of the joint LMS algorithm can be studied by considering the quality of the two estimates that it generates. Since the delay and weight vector estimates are random variables, the joint algorithm can be analysed in terms of convergence in the mean and in the mean square of either estimate. Because of the coupling between the two adaptive processes, the gradient error will affect the delay tracking and the derivative uncertainty will itself influence the adaptive filter. These mutual effects can be included in the delay variance and weight-noise vector correlation matrix, in steady-state conditions. The bounds for μ and α are determined for both types of convergence. The results for the delay estimator are given first. Then the weight vector estimator is considered and finally the two sets of results are combined, to obtain some misadjustment expressions for the joint LMS algorithm.

In the course of the analyses, in addition to the general real signals and systems assumption already mentioned, the following assumptions are used:

1) The input and noise signals are zero-mean Gaussian random processes. The noise signals are also assumed to be white noise processes.

2) The adaptive system is in *steady state* and the reference system is stationary; i.e., the reference delay is constant at $D_n = D$ and the reference filter is also fixed in time.

3) Independence theory holds; i.e., the zero-mean input data vectors are uncorrelated with each other and with $r(k)$. Then

$$E\left[u_n u_k^T\right] = 0 \quad for \quad k = 0, 1, \ldots, n-1, \quad (28.1)$$

$$E\left[u_n r(k)\right] = 0 \quad for \quad k = 0, 1, \ldots, n-1. \quad (28.2)$$

The terminology *independence theory* is common in the analysis of adaptive algorithms (see [7], for example)†.

4) In steady state, the adaptive weight vector, $\mathbf{w}_n$, can be expressed as

$$w_n = w_{opt} + \eta_n. \quad (29)$$

where $\mathbf{w}_{opt}$ is the optimum Wiener solution given by

$$w_{opt} = R^{-1} p_{n|d_n = D}, \quad (30)$$

and η_n is the weight-noise vector.

5) In the analysis of the delay estimator, the vector η_n is a zero-mean stationary Gaussian vector, uncorrelated with the data vectors (because of (28.1)-(28.2)) and such that

$$E\left[\eta_i \eta_j\right] = 0 \quad for \quad i \ne j. \quad (31)$$

The noise vector correlation matrix, defined as

$$K_\eta = E\left[\eta_n \eta_n^T\right], \quad (32)$$

is therefore diagonal with the values $E[\eta_i^2(n)]$ on the main diagonal. In the analysis of the weight vector estimate, the delay estimate is assumed stationary.

6) The system is in cancellation configuration (see Fig. 2). The results can be extended in a straightforward manner to the equalization case.

7) When the signal-to-noise ratios are assumed high, the adaptive-filter Wiener solution for $d_n = D$ is approximately equal to the reference filter (in practice, this amounts to SNRs greater than 10 dB).

Note that Assumption 3 can hardly be justified in practice, but has been used with success in the analysis of stochastic algorithms [7]. The Gaussian assumption about η_n is also commonly used in the analysis of the LMS algorithm [16]-[17]. The noise vector properties put forth in Assumption 5 follow largely from these assumptions and will prove to be useful in the analyses. Note in particular, that K_η was found to be approximately equal to $\mu\xi_{min}I$ in [9], for the LMS algorithm. The validity of this approximation is directly related to the validity of Assumption 3. In most cases, it is only asymptotically valid as the adaptation constant μ vanishes.

1. *Results for the LMS delay estimator in steady state*

The LMS delay tracking algorithm in (27) is analyzed in terms of convergence of the delay estimate in the mean and in the mean square. The analysis parallels and extends that of Messer [4] and can be found in [5].

For $d_n = D$, the output of the adaptive branch can be expressed as

$$y(n, D) = w_{opt}^T u(nT - D) + \eta_n^T u(nT - D). \quad (33)$$

The first term on the right is defined as the optimum output, $\hat{r}(n)$, since it represents the adaptive branch output for perfect modelling in the MSE sense. The second term on the right is defined as the output steady-state noise, $\chi(n, D)$. Define $e_{min}(n, D)$ as the error between the optimum adaptive branch and the reference branch; i.e.,

$$e_{min}(n, D) = r(n) - \hat{r}(n), \quad (34)$$

† Note that (28.1)-(28.2) is different from the usual independence theory assumption since u_n and u_k are influenced by different delays. But since we are in steady state, $d_n = D$ and (28.1)-(28.2) is close to the usual form.

and the corresponding MSE as

$$\xi_{min} = E\left[e_{min}^2(n, D)\right]. \tag{35}$$

Then, relying largely on Assumptions 3 and 5, we have the following two propositions:

Proposition 2. *In steady-state conditions, the delay estimator, given by the LMS delay tracking algorithm operating jointly with an adaptive filter, is an unbiased estimator if*

$$0 < \alpha < \frac{2}{\ddot{\xi}_{min}}. \tag{36}$$

Proposition 3. *In steady-state conditions, the delay estimator, given by the LMS delay tracking algorithm operating jointly with an adaptive filter, is convergent in the mean square if*

$$0 < \alpha < \frac{\ddot{\xi}_{min}}{2\sigma_G^2}, \tag{37}$$

where the quantity σ^2_G is given in (38).

The quantity σ^2_G can be shown to be [5]

$$\begin{aligned}\sigma_G^2 = {} & 3\left(\phi''_{\hat{r}\hat{r}}(0)\right)^2 + 6\phi''_{\hat{r}\hat{r}}(0)\phi''_{uu}(0)\mathrm{tr}\left[K_\eta\right] + 3\left(\phi''_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right)^2 \\ & + \left(\phi_{rr}(0) - \phi_{\hat{r}\hat{r}}(0) + \phi_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right)\left(\phi^{(4)}_{\hat{r}\hat{r}}(0) + \phi^{(4)}_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right) \\ & + 2\phi''_{\hat{r}\hat{r}}(0)\left(\phi''_{rr}(0) - \phi''_{\hat{r}\hat{r}}(0) - \phi''_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right),\end{aligned} \tag{38}$$

which, for high signal-to-noise ratios $\left(\phi''_{rr} \approx \phi''_{\hat{r}\hat{r}}\right)$, can be approximated by

$$\begin{aligned}\sigma_G^2 \approx {} & 3\left(\phi''_{\hat{r}\hat{r}}(0)\right)^2 + 4\phi''_{\hat{r}\hat{r}}(0)\phi''_{uu}(0)\mathrm{tr}\left[K_\eta\right] + 3\left(\phi''_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right)^2 \\ & + \left(\phi_{rr}(0) - \phi_{\hat{r}\hat{r}}(0) + \phi_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right)\left(\phi^{(4)}_{\hat{r}\hat{r}}(0) + \phi^{(4)}_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right),\end{aligned} \tag{39}$$

where tr[•] is the trace operator, K_η is the weight-noise correlation matrix defined in (32), $\phi(\tau)$ denotes a correlation between two random processes, the prime denotes a derivative with respect to τ, and $\phi^{(4)}(0)$ denotes $\partial^4\phi(\tau)/\partial\tau^4$ at $\tau = 0$.

Note that, in interpreting the propositions, it is important to keep in mind that the result is true if no false lock occurs; i.e., if no noise samples force the delay estimate to lock on a local solution, or if the adaptive filter does not compensate at all for the delay reference.

The steady-state delay estimate variance is given by

$$v_{ss} = \lim_{n\to\infty} E\left[\left(d_n - D\right)^2\right] = \frac{\alpha\sigma_N^2}{2\ddot{\xi}_{min} - 4\alpha\sigma_G^2}, \tag{40}$$

where σ^2_N is given by

$$\sigma_N^2 = 4\left(\phi_{rr}(0) - \phi_{\hat{r}\hat{r}}(0) + \phi_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right)\left(\phi''_{\hat{r}\hat{r}}(0) + \phi_{uu}(0)\mathrm{tr}\left[K_\eta\right]\right). \tag{41}$$

2. *Results for the LMS adaptive filter in steady state*

As in the case of the LMS delay tracking algorithm, the LMS weight vector adaptive algorithm of (26) can be analysed in terms of convergence in the mean and the mean square of the weight vector estimate. That type of analysis has been performed by many authors and the details concerning our problem can be found in [5]. Due to the assumptions made, in particular the instantaneous propagation of the adaptive delay value through the adaptive-filter delay line, the behaviour of the filter is not affected in many different ways by the delay element. The following two propositions characterize the convergence of the weight vector.

Proposition 4. *In steady-state conditions, the weight vector estimator, given by the adaptive filter LMS algorithm operating jointly with a mean-square convergent delay tracking algorithm, converges in the mean if*

$$0 < \mu < \frac{1}{\lambda_{max}}, \tag{42}$$

*where λ_{max} denotes the maximum eigenvalue of the input signal autocorrelation matrix **R**. The weight vector estimate experiences a bias given by*

$$\boldsymbol{b} = 1/2\, v_{ss} \boldsymbol{R}^{-1} \ddot{\boldsymbol{p}}(D), \tag{43}$$

where $\ddot{\boldsymbol{p}}(D)$ represents the second derivative of the cross-correlation vector with respect to the delay d_n.

Proposition 5. *In steady-state conditions, the weight vector estimator, given by the adaptive filter LMS algorithm operating jointly with a mean-square convergent delay tracking algorithm, is convergent in the mean square if*

$$0 < \mu < \frac{1}{\sum_{i=1}^{M}\lambda_i}, \tag{44}$$

*where λ_i is the i^{th} eigenvalue of the $M \times M$ input signal autocorrelation matrix **R**.*

Note that the convergence condition of (42) and (44) are identical to the usual conditions for convergence of an LMS adaptive filter [7], but that the effect of the delay estimator on the adaptive filter is to add a bias to the weight vector estimate.

3. *Excess mean squared error and misadjustment with the joint LMS algorithm*

From (8), the steady-state MSE function is

$$\xi_{ss} = \phi_{rr}(0) + E\left[\boldsymbol{w}_n^T \boldsymbol{R} \boldsymbol{w}_n\right] - 2E\left[\boldsymbol{w}_n^T \boldsymbol{p}_n\right], \tag{45}$$

where the values of the estimates take on their steady-state form. Neglecting some terms involving the square of v_{ss}, we can transform (45) into

$$\xi_{ss} = \xi_{min} + v_{ss}\ddot{\xi}_{min}/2 + \frac{\mu\left(\xi_{min} + \ddot{\xi}_{min}v_{ss}/2\right)\mathrm{tr}[\boldsymbol{R}]}{I - \mu\mathrm{tr}[\boldsymbol{R}]} \tag{46}$$

The excess MSE is given by the expression $\xi_{ex} = \xi_{ss} - \xi_{min}$, which can be transformed into

$$\xi_{ex} = \xi_{ex}^d + \xi_{ex}^f + \xi_{ex}^{df}, \tag{47}$$

where the excess MSE specific to the adaptive delay element is defined as

$$\xi_{ex}^d = \frac{v_{ss}\ddot{\xi}_{min}}{2}; \tag{48}$$

the excess MSE specific to the adaptive filter is defined as

$$\xi_{ex}^{f} = \frac{\mu \xi_{min} \mathrm{tr}[\boldsymbol{R}]}{1-\mu \mathrm{tr}[\boldsymbol{R}]}; \tag{49}$$

and the cross-product excess MSE is defined as

$$\xi_{ex}^{df} = \frac{\mu \ddot{\xi}_{min} v_{ss} \mathrm{tr}[\boldsymbol{R}]}{2(1-\mu \mathrm{tr}[\boldsymbol{R}])}. \tag{50}$$

The expression for ξ_{ex}^{d} is valid for pure LMS delay estimation [4], and the expression for ξ_{ex}^{f} is valid for an adaptive LMS filter operating without an adaptive delay [7].

The misadjustment is defined as the ratio of the excess MSE to ξ_{min}. Therefore, the misadjustment expression can be shown to be

$$M = M^{d} + M^{f} + M^{d} M^{f}, \tag{51}$$

where M^d and M^f are the misadjustments specific to the adaptive delay element and to the adaptive filter respectively. They are obtained by dividing ξ_{ex}^{d} and ξ_{ex}^{f} by ξ_{min}.

B. Discussion of the LMS algorithm analysis

The joint steepest-descent algorithm and its stochastic counterpart, the joint LMS algorithm, represent the generalizations of either the conventional SD (LMS) delay tracking algorithm [4] or the conventional SD (LMS) adaptive transversal filter algorithm [14]. It is therefore not surprising to find that all the results concerning the delay algorithm degenerate to those of [4] when the signals are properly interpreted, and that the adaptive-filter derivations come down to the LMS adaptive-filter results when the delay, D, and the variance are set equal to zero.

Another point to note is that, as long as the delay estimation algorithm is convergent in the mean square (the steady-state delay variance v_{ss} is finite), the conditions for convergence of the LMS adaptive filter are identical to the usual conditions for a similar adaptive filter operating alone or with a fixed delay element. The convergence depends on the eigenvalues of the input-signal autocorrelation matrix. Note also that, because of the adaptive delay element, the weight vector estimate is biased.

As (37) and (40) suggest, the convergence of the LMS adaptive delay element depends on ξ_{min}, σ_G^2 and σ_N^2. Using the high SNR assumption $\left(\phi_{r\hat{r}}^{''} \approx \phi_{\dot{r}\dot{r}}^{''}\right)$ and the fact that

$$\xi_{min} = \phi_{rr}(0) - \phi_{\hat{r}\hat{r}}(0), \tag{52}$$

(39) and (41) can take the form†

$$\begin{aligned} \sigma_G^2 \approx\ & 3/4\ddot{\xi}_{min}^2 - 1/2\xi_{min}\xi_{min}^{[4]} \\ & + \left[\xi_{min}\phi_{uu}^{[4]}(0) - 1/2\xi_{min}^{[4]}\phi_{uu}(0) - 2\ddot{\xi}_{min}\phi_{uu}^{''}(0)\right]\mathrm{tr}\left[\boldsymbol{K}_\eta\right] \\ & + \left[3\left(\phi_{uu}^{''}(0)\right)^2 + \phi_{uu}(0)\phi_{uu}^{(4)}(0)\right]\mathrm{tr}^2\left[\boldsymbol{K}_\eta\right] \end{aligned} \tag{53}$$

and

$$\begin{aligned} \sigma_N^2 \approx\ & 2\xi_{min}\ddot{\xi}_{min} \\ & + \left[2\ddot{\xi}_{min}\phi_{uu}(0) - 4\xi_{min}\phi_{uu}^{''}(0)\right]\mathrm{tr}\left[\boldsymbol{K}_\eta\right] \\ & - 4\phi_{uu}(0)\phi_{uu}^{''}(0)\mathrm{tr}^2\left[\boldsymbol{K}_\eta\right]. \end{aligned} \tag{54}$$

Equations (53) and (54) indicate that the convergence of the LMS adaptive delay element depends on the input signal power $\phi_{uu}(0)$ and the minimum MSE ξ_{min}. If μ is small, $\mathrm{tr}[\mathbf{K}_\eta]$ is small and we have $\sigma_G^2 \approx 3/4\ddot{\xi}_{min}^2$ and $\sigma_N^2 \approx 2\xi_{min}\ddot{\xi}_{min}$. We therefore see, from (36) and (37), that the upper bound for convergence in the mean square is about one-third of the upper bound for convergence in the mean. The steady-state delay variance is also approximately given by $v_{ss} \approx \alpha\xi_{min}$.

The delay estimate variance is encountered in the excess MSE and misadjustment expressions, such as (47) and (51). Once the delay variance is computed or fixed, these two quantities are seen to be functions of two terms specific to the adaptive delay element and to the adaptive filter respectively, and of a cross-product term (note that, since the delay-specific term is a function of v_{ss}, it is indirectly a function of the adaptive filter). The expressions for ξ_{ex}^{d} and ξ_{ex}^{f} are identical to those obtained for the respective adaptive algorithms operating alone [4], [7]. The cross-product terms, ξ_{ex}^{df} and M^{df}, are essentially the result of gradient and derivative estimation errors in the two adaptation processes. For stationary input and reference processes, the estimation noise in one adaptive algorithm is increased by the gradient estimation noise present in the other adaptive system. Therefore, the total misadjustment, M, is not merely the sum of the adaptive delay element and adaptive filter misadjustment expression M^d and M^f, but also includes a term due to the joint estimation noise. Note, however, that the cross-product misadjustment, M^{df}, is equal to the product of M^d and M^f, making it a second-order term that, in practical situations, can be one order of magnitude smaller than the individual terms.

As a final remark, note that the key quantities in the analyses are ξ_{min} and its second derivative. These quantities can be estimated from *a priori* knowledge of the transmitted signal and from the estimation of the received signal's autocorrelation functions. Some possible estimation procedures are given in [5].

C. Experimental results with the joint LMS algorithm

Using the analysis results, it is possible to compute the adaptive delay gain factor, α, as a function of the adaptive filter gain factor, μ. In order to perform this task, we combine the expression for v_{ss}, given in (40), with equations (53) and (54) and the expression for $\mathrm{tr}[\mathbf{K}_\eta]$ given by [5]:

$$\mathrm{tr}\left[\boldsymbol{K}_\eta\right] = \mu M \frac{\xi_{min} + \ddot{\xi}_{min} v_{ss}/2}{1-\mu \mathrm{tr}[\boldsymbol{R}]}.$$

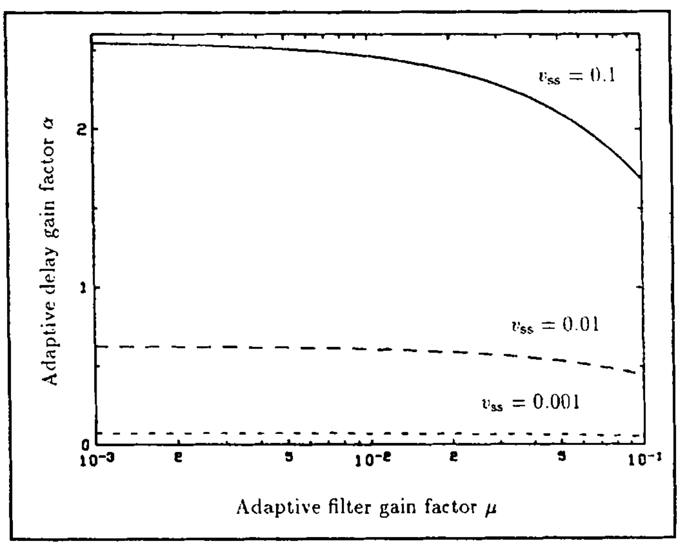

Figure 3: *Theoretical curve of α versus μ; SNR = 10 dB; small-dash curve: v_{ss} = 0.001; large-dash curve: v_{ss} = 0.01; continuous curve: v_{ss} = 0.1.*

† Note that these expressions are *exact* for white input and noise signals.

Table 1
Excess mean squared errors and misadjustments for different combinations of α's and μ's. The signal-to-noise ratio is 10 dB.

μ	α	ξ^f_{ex}	ξ^d_{ex}	ξ_{ex}	M	M_{th}
0.1	0.5	0.00312	0.00193	0.00563	40.5%	39.4%
0.05	0.5	0.00141	0.00193	0.00308	22.1%	25.4%
0.1	0.1	0.00312	0.00010	0.00313	22.5%	23.3%
0.01	0.5	0.00026	0.00193	0.00195	14.0%	16.0%
0.05	0.25	0.00141	0.00051	0.00163	11.7%	14.2%

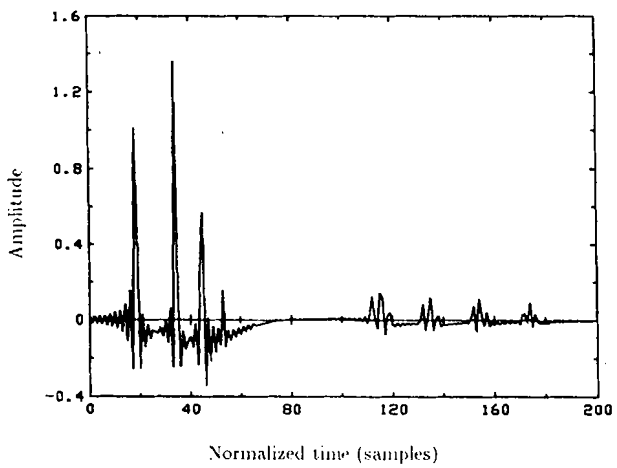

Figure 4: *Impulse response of the reverberant room.*

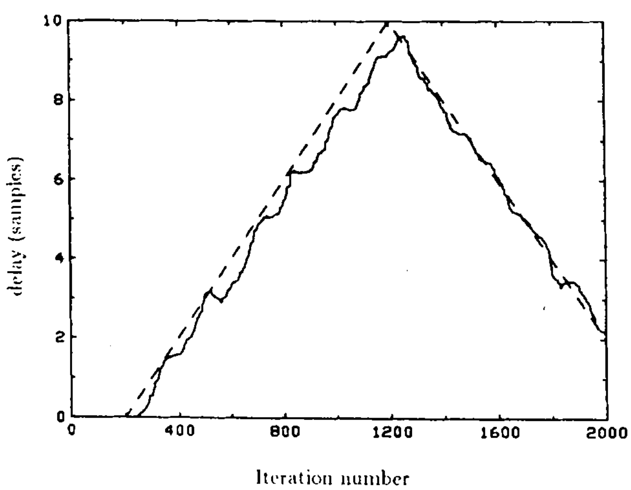

Figure 5: *LMS adaptive delay response to a reference delay ramp of 0.01 sample/sampling period and for a 200-tap reference impulse response; dashed curve: reference delay; μ = 0.01, α = 0.02. White Gaussian input.*

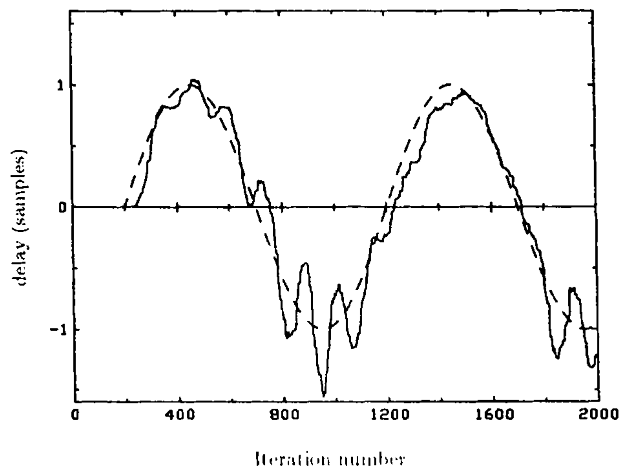

Figure 6: *LMS adaptive delay response to a sinusoidal reference delay variation and for a 200-tap reference impulse response; dashed curve: reference delay; μ = 0.01, α = 0.02. White Gaussian input.*

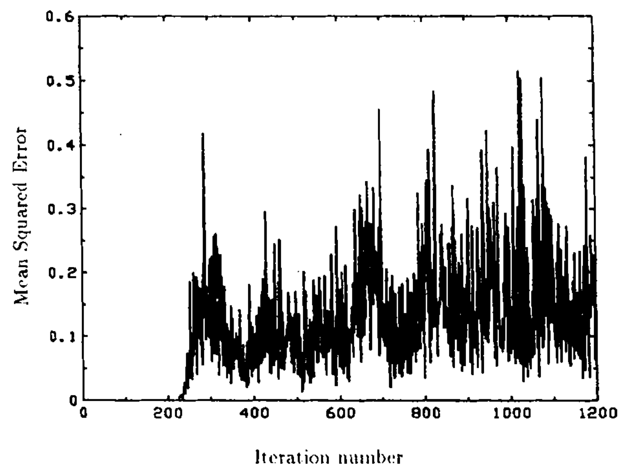

Figure 7: *Learning curve for the joint algorithm coping with a reference delay ramp of 0.01 sample/sampling period (corresponding to Fig. 5): μ = 0.01, α = 0.02.*

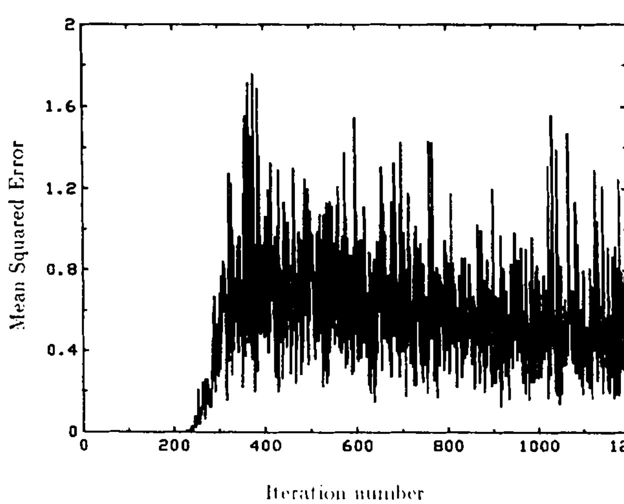

Figure 8: *Learning curve for the single adaptive filter coping with a reference delay ramp of 0.01 sample/sampling period (note the scale difference compared to Fig. 7); μ = 0.01. White Gaussian input.*

Using 21-coefficient adaptive and reference filters with white input and noise signals and for a signal-to-noise ratio of 10 dB, we obtain the plots of Fig. 3.

The gain factor α increases with v_{ss}, and for a typical variance of 0.01 the value of α is approximately constant with μ, and is around 0.5. This indicates that, for low variance, the adaptive filter does not significantly influence the behaviour of the adaptive delay. The upper bound on α for convergence in the mean square (36) is not significantly influenced by the delay variance and is approximately constant for $\mu < 0.01$ (see [5]). These critical values for α and μ are retained as indications of the values that should be used in the simulations.

An important result from the previous sections is the expression for the excess MSE at the output of the joint LMS algorithm given by (47). We verify these results by computing the theoretical value of ξ^f_{ex}, using (49), and by obtaining ξ^d_{ex} as well as ξ_{ex} through simulations. The results, for five different conbinations of α and μ, are presented in Table 1 for low-order adaptive and reference filters (21 coefficients). The corresponding measured total misadjustment, M, is obtained from ξ_{ex} through division by ξ_{min}, while the theoretical total misadjustment, M_{th}, is obtained using (51). This table shows the good agreement between the measured and the theoretical quantities. Note that since the cross-product term $M^d M^f$ is a second-order component, its effect is small or negligible, as can be seen from the fact that ξ_{ex} is always approximately equal to the sum of ξ^f_{ex} and ξ^d_{ex}.

D. Results with a long reference impulse response

In practice, the reference-filter impulse response can exhibit a fairly large number of coefficients. For example, the typical impulse response associated with a reverberant room, and encountered in audio surveillance [18], has more than 200 coefficients. We generated such an impulse response using the method proposed by Allen and Berkley [19]. The response that we obtained simulates the audio channel between a source of sound and a microphone located in a closed room, with specific wall-reflection coefficients. We arbitrarily selected the parameters to simulate the behaviour of a room measuring 6 m by 6 m, with a height of 3 m. The reflection coefficient for each wall is 0.8, the sound source is assumed to be located about 0.5 m away from one of the corners, and the location of the receiver is about 1 m from the same corner. The corresponding impulse response is given in Fig. 4. Note that the response is not symmetrical with respect to any point, and that it exhibits three large reflection peaks as well as five smaller ones. This reference impulse response is used in a system identification configuration (see Fig. 2), with a 200-coefficient adaptive filter and with both spectrally white Gaussian and audio input signals.

With a white Gaussian input, the delay tracking of the joint algorithm is shown in Figs. 5 and 6, for a reference delay ramp and a sinusoidal reference delay in noiseless conditions. The rate of change of the linear delay is higher than what is typically encountered in audio surveillance [18].

The inaccuracies in the delay estimation are related to the excess MSE, which is proportional to the number of coefficients in the adaptive filter (see (49)). Note the different behaviour of positive and negative delay tracking, especially in Fig. 6. This difference is related to the fact that the reference impulse response is not symmetrical with respect to any of its points. In order to appreciate the effectiveness of the joint algorithm, the learning curve corresponding to the joint algorithm coping with a linearly changing delay (corresponding to Fig. 5) is illustrated in Fig. 7. The learning curve, corresponding to a system identification configuration in which there is no adaptive delay, i.e., a configuration in which the adaptive filter *alone* copes with the modelling of both the linear reference delay and the reference filter, is illustrated in Fig. 8. These curves were obtained by averaging 10 error curves. Note the scale difference between Fig. 7 and Fig. 8. It is obvious from these figures that the joint algorithm generates an MSE lower than that for the single adaptive filter. Similar results can be obtained with a speech input, although the joint LMS algorithm must be normalized in this case to take into account the power variations in the input signal [5].

V. Conclusions

In this article, we have studied the joint SD and joint LMS algorithms for time-delay estimation and adaptive filtering. The presence of a multitude of stationary points in the objective function was established, and the steady-state behaviour of the two algorithms was investigated. The coupling between the two LMS adaptive algorithms was shown to give a misadjustment expression equal to the sum of the individual misadjustments plus a cross-product term. The analyses were used to obtain a theoretical view of the application of such algorithms in specific environments. The simulation results that we provided give a flavour of the sort of behaviour that one can expect when using an adaptive delay element with a conventional adaptive filter. When a long impulse response filter has to be estimated, the additional computations incurred by the LMS adaptive delay algorithm are not significant, especially given the reduction in excess MSE that is attainable. This conclusion supports the use of the joint algorithm when the main input and the reference signal are believed to exhibit some form of non-synchronous behaviour.

Acknowledgements

This work was performed with the financial assistance of the Canadian Department of Communications and of the Natural Sciences and Engineering Research Council of Canada. The provision of computing equipment by Bell-Northern Research, l'Institut National de la Recherche Scientifique and the Communications Research Centre is also greatly acknowledged. The comments of the various reviewers are appreciated.

References

1. C.H. Knapp and G.C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-24, no. 4, Aug. 1976, pp. 320-327.
2. B. Champagne, M. Eizenman and S. Pasupathy, "Exact maximum likelihood time delay estimation for short observation intervals," *IEEE Trans. Acoust. Speech Signal Process.*, to be published.
3. J.A. Stuller, "Maximum-likelihood estimation of time-varying delay - part I," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-35, no. 3, Mar. 1987, pp. 300-313.
4. H. Messer and Y. Bar-Ness, "Closed-loop least mean square time delay estimator," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-35, no. 4, April 1987, pp. 413-424.
5. D. Boudreau, *Joint Time Delay Estimation and Adaptive Filtering Techniques*, Ph.D. dissertation, McGill U., Montreal, Que., Nov. 1990.
6. A.V. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, N. J.: Prentice-Hall, 1989.
7. S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, N. J.: Prentice-Hall, 1986.
8. N. Kalouptsidis, G. Carayannis and D.G. Manolakis, "Fast design of multichannel FIR least-squares filters with optimum lag," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-32, no. 1, Feb. 1984, pp. 48-59.
9. B. Widrow et al., "Adaptive noise cancelling: principles and applications," *Proc. IEEE*, vol. 63, no. 12, Dec. 1975, pp. 1692-1716.
10. H. Meyr, "Delay-lock tracking of stochastic signals," *IEEE Trans. Comm.*, vol. COM-24, no. 3, Mar. 1976, pp. 331-339.
11. R.W. Chang, "Joint optimization of automatic equalization and carrier acquisition for digital communication," *Bell Syst. Tech. J.*, July-Aug. 1970, pp. 1069-1105.
12. A. Goldstein, *Constructive Real Analysis*, New York: Harper and Row, 1967.
13. B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, N. J.: Prentice-Hall, 1985.
14. B. Widrow, J.M. McCool, M.G. Larimore and C.R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, Aug. 1976, pp. 1151-1162.
15. L. Ljung, "Analysis of recursive stochastic algorithm," *IEEE Trans. Automat. Control*, vol. AC-22, no. 9, Aug. 1977, pp. 551-575.
16. J. Krolik, M. Eizenman and S. Pasupathy, "Time delay estimation of signals with uncertain spectra," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-36, no. 12, Dec. 1988, pp. 1801-1811.
17. D.H. Youn, N. Ahmed and G.C. Carter, "An adaptive approach for time delay estimation of band-limited signals," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-31, no. 3, June 1983, pp. 780-784.
18. D. O'Shaughnessy, P. Kabal, D. Bernardi, L. Barbeau, C.-C. Chu and J.-L. Moncet, "Applying speech enhancement to audio surveillance," *J. Foren. Sci.*, vol. 35, no. 5, Sept. 1990, pp. 1163-1172.
19. J.B. Allen and D.A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, Apr. 1979, pp. 943-950.