SCIENTIFIC REPORT
PROJECT PN-III-P4-PCE-2021-0438

Novel Sparsity-Aware Adaptive Algorithms for Acoustic Applications (NARALAN)

Stage 2022 – Development of Optimized Adaptive Algorithms with Individual Control Factors

1 Optimized LMS Algorithms with Individual Control Factors

The least-mean-square (LMS) algorithm and its variants [such as the normalized LMS (NLMS)] are frequently used for system identification problems [1]–[8]. The main parameter that controls the behavior of these algorithms is the step-size. In this context, the variable step-size (VSS) algorithms were designed to achieve a proper compromise between fast convergence rate and low misadjustment, e.g., see [9]–[14] and the references therein.

In the framework of system identification, a natural optimization criterion to follow is the minimization of the system misalignment. However, most of the VSS algorithms were developed assuming that the unknown system is time invariant, which is rarely the case in real-world applications. Consequently, these algorithms could require some additional features to control their behavior, like system change detectors or extra parameters, which are not easy tasks in practice.

In this section, we present a family of optimized LMS-based algorithms, which act like VSS adaptive filters. First, we consider that the unknown system is variable in time (based on a first-order Markov model) and the minimization of the system misalignment is carried out under this circumstance. Second, we extend the approach to the more general case of an LMS algorithm with individual control factors, which is somehow similar to the idea behind sparse adaptive filters [15]. Simulations performed in the context of acoustic echo cancellation (which is a very challenging system identification problem) indicate the good behavior of these algorithms.

1.1 System Model

Let us consider a system identification problem, where an adaptive filter is used to model an unknown system, both driven by the same zero-mean input signal $x(n)$. In this context, the reference signal at the discrete-time index n is

$$d(n) = \mathbf{h}^T(n)\mathbf{x}(n) + v(n), \tag{1}$$

where $\mathbf{h}(n) = \left[\begin{array}{cccc} h_0(n) & h_1(n) & \cdots & h_{L-1}(n) \end{array} \right]^T$ is the impulse response (of length L) of the system that we need to identify, superscript ^{T} is the transpose operator, $\mathbf{x}(n) = \left[\begin{array}{cccc} x(n) & x(n-1) & \cdots & x(n-L+1) \end{array} \right]^T$ is a vector containing the L most recent samples of the input signal, and $v(n)$ is the system noise, usually considered as a zero-mean white Gaussian noise signal.

In the following, let us consider that $\mathbf{h}(n)$ follows a simplified first-order Markov model, i.e.,

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{w}(n), \tag{2}$$

where $\mathbf{w}(n)$ is a zero-mean white Gaussian noise signal vector, which is uncorrelated with $\mathbf{h}(n-1)$. The correlation matrix of $\mathbf{w}(n)$ is assumed to be $\mathbf{R_w} = \sigma_w^2 \mathbf{I}_L$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix and the variance, σ_w^2, captures the uncertainties in $\mathbf{h}(n)$. This model can be valid in many system identification problems, like in acoustic echo cancellation (since the acoustic echo paths are time-variant systems) [16], [17].

The objective is to estimate $\mathbf{h}(n)$ with an adaptive filter, defined by $\widehat{\mathbf{h}}(n) = \left[\begin{array}{cccc} \widehat{h}_0(n) & \widehat{h}_1(n) & \cdots & \widehat{h}_{L-1}(n) \end{array} \right]^T$. It is interesting to notice that equations (1) and (2) define a state variable model, similar to Kalman filtering [18]–[20].

1.2 An Optimized LMS-Based Algorithm

The LMS algorithm [7], [8] is defined by the update:

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mu \mathbf{x}(n) e(n), \tag{3}$$

where μ is the step-size parameter and

$$e(n) = d(n) - \widehat{\mathbf{h}}^T(n-1)\mathbf{x}(n) \tag{4}$$

is the error signal of the adaptive filter. Also, let us define the a posteriori misalignment as $\mathbf{c}(n) = \mathbf{h}(n) - \widehat{\mathbf{h}}(n)$, so that, developing (3), the update results in

$$\mathbf{c}(n) = \mathbf{c}(n-1) + \mathbf{w}(n) - \mu \mathbf{x}(n) e(n). \tag{5}$$

Moreover, based on (1) and (2), the error signal from (4) can be rewritten as

$$e(n) = \mathbf{x}^T(n)\mathbf{c}(n-1) + \mathbf{x}^T(n)\mathbf{w}(n) + v(n). \tag{6}$$

Since we deal with a system identification problem, a reasonable optimization criterion to follow is the minimization of the system misalignment. Therefore, taking the ℓ_2 norm in (5), then mathematical expectation on both sides, we obtain

$$\begin{aligned} &E\left[\|\mathbf{c}(n)\|_2^2\right] = E\left[\|\mathbf{c}(n-1)\|_2^2\right] + L\sigma_w^2 \\ &- 2\mu E\left[\mathbf{x}^T(n)\mathbf{c}(n-1)e(n)\right] - 2\mu E\left[\mathbf{x}^T(n)\mathbf{w}(n)e(n)\right] \\ &+ \mu^2 E\left[e^2(n)\mathbf{x}^T(n)\mathbf{x}(n)\right]. \end{aligned} \tag{7}$$

Next, taking (6) into account, assuming that the input signal is a zero-mean Gaussian process with the variance σ_x^2, and removing the uncorrelated products, the first cross-correlation term from the second line of (7) results in

$$\begin{aligned} &E\left[\mathbf{c}^T(n-1)\mathbf{x}(n)e(n)\right] \\ &\approx \mathrm{tr}\left\{E\left[\mathbf{c}(n-1)\mathbf{c}^T(n-1)\right] E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\right\} \\ &= \sigma_x^2 E\left[\|\mathbf{c}(n-1)\|_2^2\right], \end{aligned} \tag{8}$$

where $\mathrm{tr}\{\cdot\}$ denotes the trace of a matrix. Similarly, the last cross-correlation term from the second line of (7) can be developed as

$$\begin{aligned} &E\left[\mathbf{w}^T(n)\mathbf{x}(n)e(n)\right] \\ &\approx \mathrm{tr}\left\{E\left[\mathbf{w}(n)\mathbf{w}^T(n)\right] E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\right\} = L\sigma_x^2\sigma_w^2 \end{aligned} \tag{9}$$

and, finally, the last term of (7) results in

$$\begin{aligned} &E\left[e^2(n)\mathbf{x}^T(n)\mathbf{x}(n)\right] = \mathrm{tr}\left\{E\left[v^2(n)\mathbf{x}(n)\mathbf{x}^T(n)\right]\right\} \\ &+ \mathrm{tr}\left\{E\left\{\left[\mathbf{c}^T(n-1)\mathbf{x}(n) + \mathbf{w}^T(n)\mathbf{x}(n)\right]^2 \mathbf{x}(n)\mathbf{x}^T(n)\right\}\right\}. \end{aligned} \tag{10}$$

The last term in (10) can be developed based on the Gaussian moment factoring theorem [7] (i.e., the Isserlis' theorem). Thus, after several computations, (10) becomes

$$\begin{aligned} & E\left[e^2(n)\mathbf{x}^T(n)\mathbf{x}(n)\right] \\ & = L\sigma_x^2\sigma_v^2 + (L+2)\sigma_x^4\left\{E\left[\|\mathbf{c}(n-1)\|_2^2\right] + L\sigma_w^2\right\}, \end{aligned} \tag{11}$$

where $\sigma_v^2 = E\left[v^2(n)\right]$ is the variance of the additive noise.

Denoting $m(n) = E\left[\|\mathbf{c}(n)\|_2^2\right]$ and introducing (8), (9), and (11) in (7), we obtain

$$\begin{aligned} m(n) &= \left[1 - 2\mu\sigma_x^2 + (L+2)\mu^2\sigma_x^4\right] m(n-1) \\ &+ L\mu^2\sigma_x^2\left[\sigma_v^2 + (L+2)\sigma_x^2\sigma_w^2\right] - 2L\mu\sigma_x^2\sigma_w^2 + L\sigma_w^2. \end{aligned} \tag{12}$$

At this point, let us consider that the step-size μ is time dependent, in order to evaluate $\partial m(n)/\partial\mu(n) = 0$. This leads to an optimal step-size:

$$\mu(n) = \frac{1}{(L+2)\sigma_x^2 + \xi}, \tag{13}$$

where $\xi = L\sigma_v^2/\left[m(n-1) + L\sigma_w^2\right]$. So, the filter update is

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mu(n)\mathbf{x}(n)e(n). \tag{14}$$

To update the parameter $m(n)$, the step-size from (13) is introduced in (12) and, after some computations, it results in

$$m(n) = \left[1 - \mu(n)\sigma_x^2\right]\left[m(n-1) + L\sigma_w^2\right]. \tag{15}$$

Summarizing, the resulting optimized LMS-based algorithm is defined by the relations (4), (14), and (15), using the initialization $\widehat{\mathbf{h}}(0) = \mathbf{0}$ and $m(0) = \varepsilon$ (where ε is a positive constant). This algorithm is very similar to the simplified Kalman filter from [20], which was obtained as an approximation of the general Kalman filter; similar works can be found in [21] and [22]. Also, the resulting algorithm can be obtained following the line of the NLMS algorithm and a joint-optimization problem on both the normalized step-size and regularization parameters [14]. For the sake of consistency, we will refer this algorithm as the joint-optimized NLMS (JO-NLMS).

1.3 An Optimized LMS-Based Algorithm with Individual Control Factors

In this subsection, we extend the previous approach to an LMS-based algorithm with individual control factors. This algorithm is defined by the update [23]:

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mu\mathbf{K}(n-1)\mathbf{x}(n)e(n), \tag{16}$$

where $\mathbf{K}(n)$ is a diagonal matrix ($L \times L$) containing the control factors at time index n. The filter update (16) can be also rewritten in terms of the a posteriori misalignment as

$$\mathbf{c}(n) = \mathbf{c}(n-1) + \mathbf{w}(n) - \mu\mathbf{K}(n-1)\mathbf{x}(n)e(n). \tag{17}$$

Next, taking the ℓ_2 norm in (17), then mathematical expectation on both sides (also removing the uncorrelated products), we obtain

$$\begin{aligned} E\left[\|\mathbf{c}(n)\|_2^2\right] &= E\left[\|\mathbf{c}(n-1)\|_2^2\right] + L\sigma_w^2 \\ &- 2\mu E\left[\mathbf{x}^T(n)\mathbf{K}(n-1)\mathbf{w}(n)e(n)\right] \\ &- 2\mu E\left[\mathbf{x}^T(n)\mathbf{K}(n-1)\mathbf{c}(n-1)e(n)\right] \\ &+ \mu^2 E\left[e^2(n)\mathbf{x}^T(n)\mathbf{K}^2(n-1)\mathbf{x}(n)\right]. \end{aligned} \tag{18}$$

At this point, we need to process the last three terms in (18). Based on (6) and assuming that the input signal is white, the first cross-correlation term becomes

$$\begin{aligned} & E\left[\mathbf{x}^T(n)\mathbf{K}(n-1)\mathbf{w}(n)e(n)\right] \\ & = E\left\{\operatorname{tr}\left[\mathbf{w}(n)\mathbf{w}^T(n)\mathbf{K}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\right]\right\} \\ & = \sigma_w^2\sigma_x^2\operatorname{tr}\left[\mathbf{K}(n-1)\right] \end{aligned} \quad (19)$$

and the second cross-correlation term results in

$$\begin{aligned} & E\left[\mathbf{x}^T(n)\mathbf{K}(n-1)\mathbf{c}(n-1)e(n)\right] \\ & = \operatorname{tr}\left\{E\left[\mathbf{c}(n-1)\mathbf{c}^T(n-1)\right]E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\mathbf{K}(n-1)\right\} \\ & = \sigma_x^2\operatorname{tr}\left[\mathbf{R_c}(n-1)\mathbf{K}(n-1)\right], \end{aligned} \quad (20)$$

where $\mathbf{R_c}(n) = E\left[\mathbf{c}(n)\mathbf{c}^T(n)\right]$.

The last term in (18) is more difficult to process. However, for a certain stationarity degree of the input signal and for large values of L ($L \gg 1$), we can treat the term $\|\mathbf{K}(n-1)\mathbf{x}(n)\|_2^2$ as a deterministic quantity. Also, we can use the orthogonality principle, i.e., assuming that the adaptive filter has converged to a certain degree. Using these assumptions, the last term in (18) can be developed as

$$\begin{aligned} & E\left[e^2(n)\mathbf{x}^T(n)\mathbf{K}^2(n-1)\mathbf{x}(n)\right] \\ & = \operatorname{tr}\left\{E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\mathbf{K}^2(n-1)\right\}E\left[e^2(n)\right] \\ & = \sigma_x^2\operatorname{tr}\left[\mathbf{K}^2(n-1)\right]\left[\sigma_v^2 + L\sigma_w^2\sigma_x^2 + \sigma_x^2 m(n-1)\right]. \end{aligned} \quad (21)$$

Finally, using (19)–(21) in (18), it results in

$$\begin{aligned} m(n) = {} & m(n-1) + L\sigma_w^2 \\ & - 2\mu\sigma_x^2\operatorname{tr}\left\{\left[\mathbf{R_c}(n-1) + \sigma_w^2\mathbf{I}_L\right]\mathbf{K}(n-1)\right\} \\ & + \mu^2\sigma_x^2\operatorname{tr}\left[\mathbf{K}^2(n-1)\right]\left[\sigma_v^2 + L\sigma_w^2\sigma_x^2 + \sigma_x^2 m(n-1)\right]. \end{aligned} \quad (22)$$

As in Section 1.2, we can impose $\partial m(n)/\partial\mu(n) = 0$ (considering that the step-size parameter is time dependent), to obtain an optimal step-size:

$$\mu(n) = \frac{\operatorname{tr}\left\{\left[\mathbf{R_c}(n-1) + \sigma_w^2\mathbf{I}_L\right]\mathbf{K}(n-1)\right\}}{\operatorname{tr}\left[\mathbf{K}^2(n-1)\right]\left\{\sigma_v^2 + \sigma_x^2\left[m(n-1) + L\sigma_w^2\right]\right\}}. \quad (23)$$

Also, using (23) in (22), the update of $m(n)$ becomes

$$\begin{aligned} m(n) = {} & m(n-1) + L\sigma_w^2 \\ & - \frac{\sigma_x^2\left\{\operatorname{tr}\left\{\left[\mathbf{R_c}(n-1) + \sigma_w^2\mathbf{I}_L\right]\mathbf{K}(n-1)\right\}\right\}^2}{\operatorname{tr}\left[\mathbf{K}^2(n-1)\right]\left\{\sigma_v^2 + \sigma_x^2\left[m(n-1) + L\sigma_w^2\right]\right\}}. \end{aligned} \quad (24)$$

Let us focus on the main term of the numerator in (24). First, we introduce the notation:

$$\begin{aligned} \mathbf{\Gamma}(n) &= \mathbf{R_c}(n) + \sigma_w^2\mathbf{I}_L, \\ \boldsymbol{\gamma}(n) &= \begin{bmatrix} \gamma_0(n) & \gamma_1(n) & \cdots & \gamma_{L-1}(n) \end{bmatrix}^T, \\ \mathbf{k}(n) &= \begin{bmatrix} k_0(n) & k_1(n) & \cdots & k_{L-1}(n) \end{bmatrix}^T, \end{aligned}$$

where $\boldsymbol{\gamma}(n)$ and $\mathbf{k}(n)$ are two vectors containing the diagonal elements of $\mathbf{\Gamma}(n)$ and $\mathbf{K}(n)$, respectively. Since $\mathbf{K}(n)$ is a diagonal matrix, we can use the Cauchy-Schwarz inequality:

$$\left\{\operatorname{tr}\left[\mathbf{\Gamma}(n-1)\mathbf{K}(n-1)\right]\right\}^2 \leq \|\boldsymbol{\gamma}(n-1)\|_2^2\,\|\mathbf{k}(n-1)\|_2^2. \quad (25)$$

It is known that the equality in (25) is obtained if and only if $\boldsymbol{\gamma}(n-1)$ and $\mathbf{k}(n-1)$ are linearly dependent, i.e., $\mathbf{k}(n-1) = q\boldsymbol{\gamma}(n-1)$, with $q > 0$. Hence,

$$\begin{aligned} &\text{tr}\left[\mathbf{K}(n-1)\right] = q\text{tr}\left[\boldsymbol{\Gamma}(n-1)\right] = q\left[m(n-1) + L\sigma_w^2\right], \\ &\text{tr}\left[\mathbf{K}^2(n-1)\right] = \|\mathbf{k}(n-1)\|_2^2 = q^2\|\boldsymbol{\gamma}(n-1)\|_2^2. \end{aligned} \tag{26}$$

Consequently, the optimal step-size from (23) results in

$$\mu(n) = \frac{1}{q\left\{\sigma_v^2 + \sigma_x^2\left[m(n-1) + L\sigma_w^2\right]\right\}} \tag{27}$$

and the update of the parameter $m(n)$ from (24) becomes

$$m(n) = m(n-1) + L\sigma_w^2 - q\mu(n)\sigma_x^2\|\boldsymbol{\gamma}(n-1)\|_2^2. \tag{28}$$

Finally, the filter update from (16) is evaluated as

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + q\mu(n)\boldsymbol{\gamma}(n-1)\odot\mathbf{x}(n)\mathbf{e}(n), \tag{29}$$

where $\odot$ denotes the Hadamard product.

However, we still need to find the value of q and an update relation for $\boldsymbol{\gamma}(n)$. The parameter q is related to the value of $\text{tr}\left[\mathbf{K}(n-1)\right]$, as shown in (26). In the classical algorithm, the trace of this matrix is equal to L [since $\mathbf{K}(n-1)$ is replaced by $\mathbf{I}_L$]. In our case, we could impose $\text{tr}\left[\mathbf{K}(n-1)\right] = L$, but the gains will be individually distributed through the elements of $\boldsymbol{\gamma}(n-1)$. Consequently, we could evaluate

$$q = \frac{L}{m(n-1) + L\sigma_w^2}. \tag{30}$$

The vector $\boldsymbol{\gamma}(n)$ contains the diagonal elements of the matrix $\boldsymbol{\Gamma}(n)$. But this matrix depends on $\mathbf{R_c}(n)$. It is interesting to notice that the elements of $\boldsymbol{\gamma}(n)$, i.e., the individual control factors, depend on the coefficients' misalignment (as compared to the classical proportionate-type algorithms, where the adaptation is controlled depending on the coefficients' magnitude). Next, based on (17), we can write

$$\begin{aligned} \mathbf{R_c}(n) = \mathbf{R_c}(n-1) &+ \sigma_w^2\mathbf{I}_L \\ &- \mu E\left[\mathbf{w}(n)\mathbf{x}^T(n)\mathbf{K}(n-1)e(n)\right] \\ &- \mu E\left[\mathbf{K}(n-1)\mathbf{x}(n)\mathbf{w}^T(n)e(n)\right] \\ &- \mu E\left[\mathbf{c}(n-1)\mathbf{x}^T(n)\mathbf{K}(n-1)e(n)\right] \\ &- \mu E\left[\mathbf{K}(n-1)\mathbf{x}(n)\mathbf{c}^T(n-1)e(n)\right] \\ &+ \mu^2 E\left[e^2(n)\mathbf{K}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{K}(n-1)\right]. \end{aligned} \tag{31}$$

Following a similar approach we used to process (18), we could further develop (31), in order to finally obtain [also using the optimal step-size from (27)]:

$$\begin{aligned} \boldsymbol{\gamma}(n) = \boldsymbol{\gamma}(n-1) &+ \sigma_w^2\mathbf{1}_{L\times 1} \\ &+ q\mu(n)\sigma_x^2\left\{\sigma_v^2 + \sigma_x^2\left[m(n-1) + L\sigma_w^2\right] - 2q\mu(n)\right\} \\ &\times\boldsymbol{\gamma}(n-1)\odot\boldsymbol{\gamma}(n-1), \end{aligned} \tag{32}$$

where $\mathbf{1}_{L\times 1}$ denotes a column vector with all its L elements equal to one. Due to stability reasons, a normalization should be also performed in this step, i.e., $\overline{\gamma}_i(n) = \gamma_i(n)/\max\left[q\boldsymbol{\gamma}(n)\right]$, with $0 \leq i \leq L-1$. Summarizing, the resulting optimized LMS algorithm with individual control factors (OLMS-ICF) is defined by the relations (27)–(30) and (32) (including the normalization).

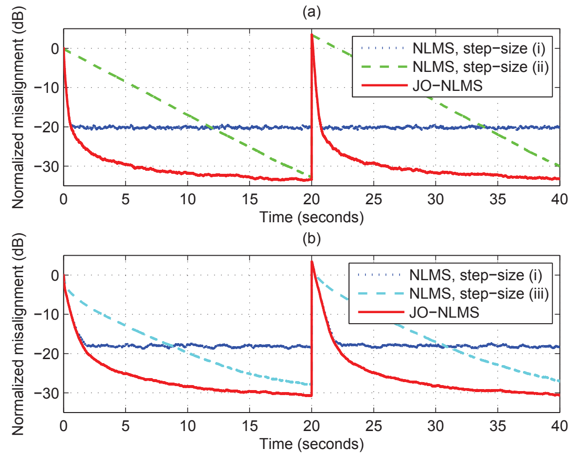

Figure 1: Misalignment of the NLMS algorithm using different step-sizes, i.e., (i) $1/[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]$, (ii) $0.025/[\delta+\mathbf{x}^T(n)\mathbf{x}(n)]$, and (iii) $0.15/[\delta+\mathbf{x}^T(n)\mathbf{x}(n)]$, and misalignment of the JO-NLMS algorithm. Echo path changes after 20 seconds, $L = 1000$, and SNR = 20 dB. The input signal is (a) white Gaussian and (b) an AR(1) process.

1.4 Practical Considerations

There are three main parameters that should be set or estimated within both JO-NLMS and OLMS-ICF algorithms. The first one is the variance of the input signal, which could be easily evaluated as in the NLMS algorithm, i.e., $\widehat{\sigma}_x^2(n) = \mathbf{x}^T(n)\mathbf{x}(n)/L$. The second parameter is the noise power, σ_v^2. This parameter can be estimated in different ways; for example, in echo cancellation, it can be estimated during silences of the near-end talker [10]. Also, other practical methods to estimate σ_v^2 can be found in [24], [25].

The third parameter to be found is σ_w^2, which is a specific one. In practice, we propose to estimate σ_w^2 as in [20], by taking the ℓ_2 norm on both sides of (2) and replacing $\mathbf{h}(n)$ by its estimate $\widehat{\mathbf{h}}(n)$, thus resulting

$$\widehat{\sigma}_w^2(n) = \frac{1}{L}\left\|\widehat{\mathbf{h}}(n) - \widehat{\mathbf{h}}(n-1)\right\|_2^2. \tag{33}$$

The difference from the right-hand side of (65) represents the update term of the algorithm, which can be used to save L subtractions associated with the direct evaluation of (33).

1.5 Simulation Results LMS ICF

Simulations are performed in an acoustic echo cancellation scenario [16], [17]. In this context, an adaptive filter is used to estimate the impulse response (i.e., the acoustic echo path) between the loudspeaker and the microphone of a hands-free communication device. This is a very challenging system identification problem, due to the high length and time-varying nature of the acoustic echo paths.

The length of the acoustic impulse response used in the experiments is $L = 1000$ (the sampling rate is 8 kHz); the same length is set for the adaptive filter. The input signal, $x(n)$, is either a white Gaussian noise, an AR(1) process generated by filtering a white Gaussian noise through a first-order system $1/\left(1 - 0.8z^{-1}\right)$, or a speech sequence. The output of the echo path is corrupted by an independent white Gaussian noise $v(n)$ and the signal-to-noise ratio (SNR) is set to 20 dB. In our simulations, we assume that σ_v^2 is known; in practice, it can be estimated like in [10], [24], [25].

In the first set of experiments, we evaluate the VSS features of the JO-NLMS algorithm. An echo path change scenario is simulated, by shifting the impulse response to the right by 12 samples, in the middle of the simulation. The results are presented in Figs. 1-3, using a white Gaussian noise or an AR(1) process as the input. The NLMS algorithm (with different step-sizes and including a regularization parameter $\delta = 20\sigma_x^2$) is used for comparison.

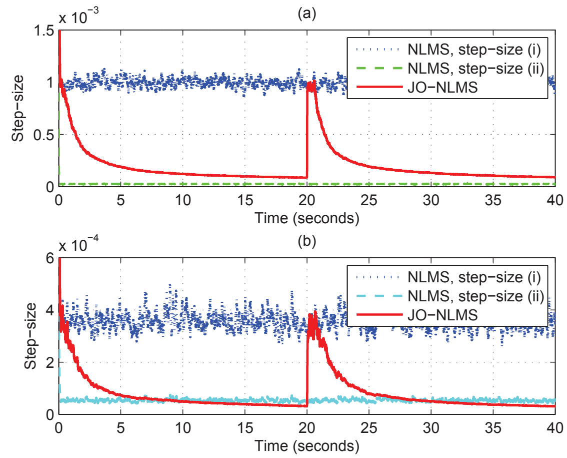

Figure 2: Evolution of the step-size parameters of the NLMS and JO-NLMS algorithms. Other conditions same as in Fig. 1.

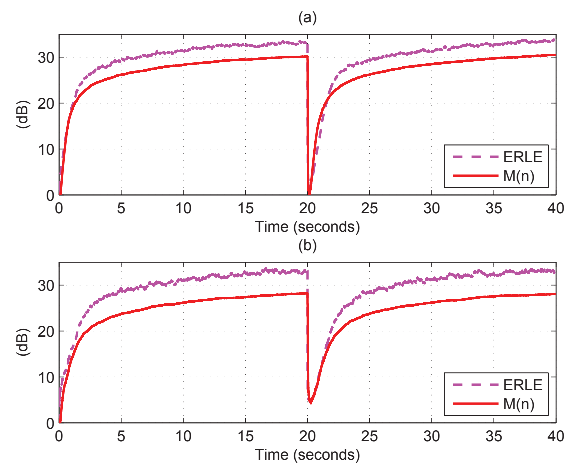

Figure 3: Evolution of the parameter $M(n) = \left\|\widehat{\mathbf{h}}(n)\right\|_2^2 / m(n)$ and the ERLE of the JO-NLMS algorithm. Other conditions same as in Fig. 1.

In Fig. 1, the performance measure is the normalized misalignment (in dB), defined as $20\log_{10}\left\|\mathbf{h}(n) - \widehat{\mathbf{h}}(n)\right\|_2 / \left\|\mathbf{h}(n)\right\|_2$. As we can notice from this figure, the JO-NLMS algorithm has a fast convergence rate and tracking, similar to the NLMS algorithm using the largest step-size, while achieving a lower misalignment level. This behavior is also supported in Fig. 2, where the evolution of the step-size parameters is presented. It can be noticed that the step-size of the JO-NLMS algorithm goes to the fastest convergence mode in the initial convergence phase and when the system changes, while its value decreases in the steady-state.

An interesting parameter of the JO-NLMS algorithm is $m(n)$. This can be used to evaluate (in an online manner) the overall performance of the algorithm. In echo cancellation, a well-known performance measure is the echo-return loss enhancement (ERLE) [16], which requires the a priori knowledge of the impulse response $\mathbf{h}(n)$ or the echo signal. In the case of the JO-NLMS algorithm, we propose to monitor the parameter $M(n) = \left\|\widehat{\mathbf{h}}(n)\right\|_2^2 / m(n)$, which should have a similar significance with the ERLE. This issue is supported in Fig. 3, where the evolution of the ERLE and $M(n)$ (both in dB) is depicted. However, contrary to the ERLE, the parameter $M(n)$ could be evaluated within the JO-NLMS algorithm.

In the second set of experiments, the input signal is a speech sequence. In Fig. 4, the JO-NLMS

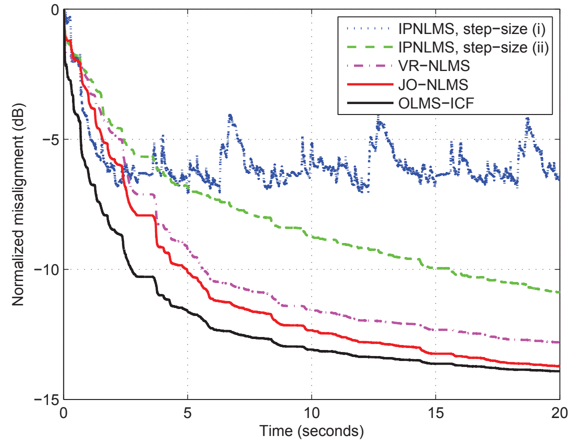

Figure 4: Misalignment of the IPNLMS algorithm using different step-sizes, i.e., (i) $1/[\delta' + \mathbf{x}^T(n)\mathbf{x}(n)]$ and (ii) $0.05/[\delta' + \mathbf{x}^T(n)\mathbf{x}(n)]$, and misalignment of the VR-NLMS [11], JO-NLMS, and OLMS-ICF algorithms. The input signal is speech, $L = 1000$, and SNR $= 20$ dB.

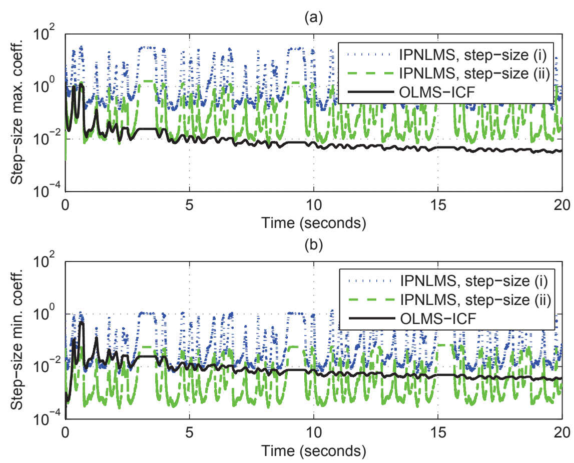

Figure 5: Evolution of the step-size parameters of the IPNLMS and OLMS-ICF algorithms, corresponding to (a) the maximum coefficient and (b) the minimum coefficient (in terms of magnitude). Other conditions same as in Fig. 4.

and OLMS-ICF algorithms are compared to the improved proportionate NLMS (IPNLMS) algorithm [26] (with different step-sizes and including a regularization parameter $\delta' = 20\sigma_x^2/L$). Also, we introduce for comparison the variable-regularized (VR) affine projection algorithm proposed in [11], using a projection order equal to one, which is equivalent to a VR-NLMS algorithm. As we can notice, the OLMS-ICF outperforms the other algorithms.

It would be interesting to observe the individual step-sizes of the OLMS-ICF algorithm. In Fig. 5, we depicted the evolution of the individual step-sizes of two coefficients, i.e., the maximum and the minimum one (in terms of magnitude), comparing the IPNLMS and OLMS-ICF algorithms. As expected, the IPNLMS algorithm allocates a significant gain to the largest coefficient. In the OLMS-ICF algorithm, this difference is valid only in the initial convergence phase, since the misalignments of the individual coefficients [i.e., the elements of $\boldsymbol{\gamma}(n)$] tend to become similar in the steady-state.

In this first section, we have presented a family of optimized LMS-based algorithms in the context of a state variable model for system identification. The optimization criterion is the minimization of the system misalignment, which is a natural approach in system identification problems. The resulting algorithms do not require additional features to monitor their behavior (e.g., system change detectors, stability thresholds, etc.), thus being reliable candidates for real-world applications. Simulations performed in the context of acoustic echo cancellation support the theoretical findings and indicate the good performance of these algorithms.

2 Optimized APA with Individual Control Factors

Nowadays, hands-free communication devices are involved in many popular applications, such as mobile telephony and teleconferencing systems. An important issue that has to be addressed when dealing with such devices is the acoustic coupling between the loudspeaker and the microphone. In other words, the microphone of the hands-free equipment captures a signal coming from its own loudspeaker, known as the acoustic echo. Depending of the environment's characteristics, this phenomenon can be very disturbing for the far-end speaker, which hears a replica of her/his own voice. From this point of view, there is a need to enhance the quality of the microphone signal by cancelling the unwanted acoustic echo.

The principal issue in acoustic echo cancellation (AEC) [16], [17] is to estimate the impulse response between the loudspeaker and the microphone of a hands-free communication device. The most reliable solution to this problem is the use of an adaptive filter that generates at its output a replica of the echo, which is further subtracted from the microphone signal. Basically, the adaptive filter has to model an unknown system, i.e., the acoustic echo path between the loudspeaker and microphone, like in a "system identification" scenario.

The affine projection algorithm (APA) [27], [28] is very popular in many system identification problems. As compared to the normalized least-mean-square (NLMS) algorithm, the APA usually offers an improved convergence rate, especially for correlated inputs. For example, in AEC, we deal with both high-length echo paths and highly-correlated signals like speech; consequently, the APA is frequently the algorithm of choice for this important application.

There are three main parameters that influence the performance of the APA, i.e., the step-size, the regularization term, and the projection order. Constant values of these parameters usually lead to a compromise between the main performance criteria, i.e., convergence rate versus misadjustment. In order to improve the overall performance of the APA, these parameters can be controlled (i.e., making them time dependent), thus resulting variable step-size (VSS), variable regularized (VR), and even variable projection (or evolving order) versions of the APA, e.g., see [29]–[37] and references therein. Most of these algorithms were developed in a system identification context, aiming to minimize the system misalignment, which is the natural approach in these scenarios.

In this section, we focus on a joint-optimization on both the step-size and regularization parameters of the APA. Moreover, we consider a more general framework, where the impulse response to be identified is modeled by a time-varying system following a first-order Markov model. This framework is valid in many real-world system identification problems, like AEC [17], [22]. Simulation results indicate that the optimized APA could be an appealing choice for such applications.

2.1 System Model and APA

Let us consider a system identification problem, like in AEC [16], [17]. In this context, the microphone or desired signal at the discrete-time index n is

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n), \tag{34}$$

where

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L+1) \end{bmatrix}^T$$

is a vector containing the L most recent time samples of the input (loudspeaker) signal $x(n)$, superscript ^{T} denotes transpose of a vector or a matrix,

$$\mathbf{h}(n) = \begin{bmatrix} h_0(n) & h_1(n) & \cdots & h_{L-1}(n) \end{bmatrix}^T$$

is the impulse response (of length L) of the system (from the loudspeaker to the microphone) that we need to identify, and $v(n)$ is a zero-mean white Gaussian noise signal. The variance of this additive noise is $\sigma_v^2 = E\left[v^2(n)\right]$, where $E(\cdot)$ denotes mathematical expectation.

Let us assume that $\mathbf{h}(n)$ is a zero-mean random vector, which follows a simplified first-order Markov model, i.e.,

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{w}(n), \tag{35}$$

where $\mathbf{w}(n)$ is a zero-mean white Gaussian noise signal vector, which is uncorrelated with $\mathbf{h}(n-1)$. The correlation matrix of $\mathbf{w}(n)$ is assumed to be $\mathbf{R_w} = \sigma_w^2 \mathbf{I}_L$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix. The variance, σ_w^2, captures the uncertainties in $\mathbf{h}(n)$. Equations (34) and (35) define a state variable model, similar to Kalman filtering [38], [39]. The objective is to estimate or identify $\mathbf{h}(n)$ with an adaptive filter, defined by

$$\widehat{\mathbf{h}}(n) = \left[\begin{array}{cccc} \widehat{h}_0(n) & \widehat{h}_1(n) & \cdots & \widehat{h}_{L-1}(n) \end{array}\right]^T.$$

In this framework, let us consider the APA [27] with the update:

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \alpha \mathbf{X}(n)\left[\delta \mathbf{I}_P + \mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1}\mathbf{e}(n), \tag{36}$$

where α is the step-size of the algorithm, δ is the regularization parameter, $\mathbf{I}_P$ is the $P \times P$ identity matrix (with P denoting the projection order),

$$\mathbf{X}(n) = \left[\begin{array}{cccc} \mathbf{x}(n) & \mathbf{x}(n-1) & \cdots & \mathbf{x}(n-P+1) \end{array}\right]$$

is the input data matrix, and

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\widehat{\mathbf{h}}(n-1) \tag{37}$$

is the error signal vector of length P, where

$$\mathbf{d}(n) = \left[\begin{array}{cccc} d(n) & d(n-1) & \cdots & d(n-P+1) \end{array}\right]^T$$

is a vector containing the P most recent time samples of the desired signal. The celebrated NLMS algorithm is obtained for $P = 1$.

2.2 Convergence Analysis of the APA

In this framework, the a posteriori misalignment is defined as

$$\boldsymbol{\mu}(n) = \mathbf{h}(n) - \widehat{\mathbf{h}}(n). \tag{38}$$

Developing (36) based on (35) and (38), we obtain

$$\begin{aligned} \boldsymbol{\mu}(n) &= \boldsymbol{\mu}(n-1) + \mathbf{w}(n) \\ &\quad - \alpha \mathbf{X}(n)\left[\delta \mathbf{I}_P + \mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1}\mathbf{e}(n). \end{aligned} \tag{39}$$

Next, let us examine the update term of the APA, i.e., the second term on the right hand side of (36). It is clear that $E\left[\mathbf{X}^T(n)\mathbf{X}(n)\right] = L\mathbf{R}_{\mathbf{x},P}(n)$, where $\mathbf{R}_{\mathbf{x},P}(n)$ is the input signal correlation matrix of size $P \times P$. For large values of L, it holds that

$$\mathbf{X}^T(n)\mathbf{X}(n) \approx L\mathbf{R}_{\mathbf{x},P}(n). \tag{40}$$

Also, let us denote

$$\begin{aligned} \mathbf{Y}(n) &= \alpha\left[\delta \mathbf{I}_P + \mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1} \\ &\approx \alpha\left[\delta \mathbf{I}_P + L\mathbf{R}_{\mathbf{x},P}(n)\right]^{-1}. \end{aligned} \tag{41}$$

Obviously, $\mathbf{Y}^T(n) = \mathbf{Y}(n)$. As we can see, this matrix contains both the control parameters, i.e., α and δ, and also the statistical information of the input signal. Nevertheless, the existence of this inverse increases the difficulty in terms of computing the statistical expectations. However, for a large value of L and a certain stationarity degree of the input signal, we can treat $\mathbf{Y}(n)$ as a deterministic quantity.

Consequently, using the ℓ_2 norm in (39) and taking the expectations on both sides (also removing the uncorrelated products), we obtain

$$\begin{aligned} E\left[\|\boldsymbol{\mu}(n)\|_2^2\right] &= E\left[\|\boldsymbol{\mu}(n-1)\|_2^2\right] + E\left[\|\mathbf{w}(n)\|_2^2\right] \\ &- 2E\left[\mathbf{e}^T(n)\mathbf{Y}(n)\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\right] \\ &- 2E\left[\mathbf{e}^T(n)\mathbf{Y}(n)\mathbf{X}^T(n)\mathbf{w}(n)\right] \\ &+ E\left[\mathbf{e}^T(n)\mathbf{Y}(n)\mathbf{X}^T(n)\mathbf{X}(n)\mathbf{Y}(n)\mathbf{e}(n)\right] \end{aligned} \tag{42}$$

or, equivalently,

$$\begin{aligned} E\left[\|\boldsymbol{\mu}(n)\|_2^2\right] &= E\left[\|\boldsymbol{\mu}(n-1)\|_2^2\right] + E\left[\|\mathbf{w}(n)\|_2^2\right] \\ &- 2\text{tr}\left\{\mathbf{Y}(n)E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\mathbf{e}^T(n)\right]\right\} \\ &- 2\text{tr}\left\{\mathbf{Y}(n)E\left[\mathbf{e}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right]\right\} \\ &+ L\text{tr}\left\{\mathbf{Y}(n)\mathbf{R}_{\mathbf{x},P}(n)\mathbf{Y}(n)E\left[\mathbf{e}(n)\mathbf{e}^T(n)\right]\right\}, \end{aligned} \tag{43}$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix.

At this point, let us focus on the last three cross-correlation terms in (43). First, based on (34), (35), (38), and assuming that the impulse response $\mathbf{h}(n)$ slowly varies for P consecutive time indices, the error signal vector from (37) can be expressed as

$$\mathbf{e}(n) = \mathbf{X}^T(n)\boldsymbol{\mu}(n-1) + \mathbf{X}^T(n)\mathbf{w}(n) + \mathbf{v}(n), \tag{44}$$

where $\mathbf{v}(n) = \begin{bmatrix} v(n) & v(n-1) & \cdots & v(n-P+1) \end{bmatrix}^T$. Consequently, based on (44), the first cross-correlation term from (43) can be developed as

$$\begin{aligned} &E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\mathbf{e}^T(n)\right] \\ &= E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\mathbf{X}(n)\right] \\ &+ E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\mathbf{w}^T(n)\mathbf{X}(n)\right] \\ &+ E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\mathbf{v}^T(n)\right]. \end{aligned} \tag{45}$$

The second term in (45) can be assumed null, since $\mathbf{w}(n)$ is not correlated with $\mathbf{h}(n-1)$ and, consequently, with $\boldsymbol{\mu}(n-1)$. Also, the third term can be neglected [i.e., the correlation with $\mathbf{v}(n)$], even if some recent studies indicated a certain correlation between $\boldsymbol{\mu}(n-1)$ and the past samples of the noise [11]. Under these circumstances, we get

$$\begin{aligned} &E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\mathbf{e}^T(n)\right] \\ &\approx E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\mathbf{X}(n)\right]. \end{aligned} \tag{46}$$

Similarly, using (44) and neglecting the uncorrelated products, the second cross-correlation term from (43) results as

$$E\left[\mathbf{e}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right] \approx E\left[\mathbf{X}^T(n)\mathbf{w}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right]. \tag{47}$$

The last cross-correlation term from (43) can be also developed based on (44). Using similar approach and assumptions, it results

$$\begin{aligned} &E\left[\mathbf{e}(n)\mathbf{e}^T(n)\right] \approx E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\mathbf{X}(n)\right] \\ &+ E\left[\mathbf{X}^T(n)\mathbf{w}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right] + E\left[\mathbf{v}(n)\mathbf{v}^T(n)\right]. \end{aligned} \tag{48}$$

Finally, introducing (46)–(48) in (43), denoting

$$m(n) = E\left[\|\boldsymbol{\mu}(n)\|_2^2\right], \tag{49}$$

and knowing that $E\left[\|\mathbf{w}(n)\|_2^2\right] = L\sigma_w^2$, we obtain

$$\begin{aligned} m(n) &= m(n-1) + L\sigma_w^2 \\ &+ \operatorname{tr}\left\{\mathbf{P}(n)\mathbf{Y}(n)E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\mathbf{X}(n)\right]\right\} \\ &+ \operatorname{tr}\left\{\mathbf{P}(n)\mathbf{Y}(n)E\left[\mathbf{X}^T(n)\mathbf{w}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right]\right\} \\ &+ L\operatorname{tr}\left\{\mathbf{Y}(n)\mathbf{R}_{\mathbf{x},P}(n)\mathbf{Y}(n)E\left[\mathbf{v}(n)\mathbf{v}^T(n)\right]\right\}, \end{aligned} \tag{50}$$

where

$$\mathbf{P}(n) = L\mathbf{Y}(n)\mathbf{R}_{\mathbf{x},P}(n) - 2\mathbf{I}_P. \tag{51}$$

It is difficult to further process (50) without any supporting assumptions on the character of the input signal. In order to facilitate the analysis, let us assume that the input signal is white (or weakly correlated), so that we can use the approximation:

$$\mathbf{R}_{\mathbf{x},P}(n) \approx \sigma_x^2\mathbf{I}_P, \tag{52}$$

which implies that

$$\mathbf{Y}(n) \approx \frac{\alpha}{L\sigma_x^2 + \delta}\mathbf{I}_P = \beta\mathbf{I}_P. \tag{53}$$

In this context, (50) can be simplified as

$$\begin{aligned} m(n) &= m(n-1) + L\sigma_w^2 \\ &+ \beta\left(\beta L\sigma_x^2 - 2\right)\operatorname{tr}\left\{E\left[\mathbf{X}^T(n)\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\mathbf{X}(n)\right]\right\} \\ &+ \beta\left(\beta L\sigma_x^2 - 2\right)\operatorname{tr}\left\{E\left[\mathbf{X}^T(n)\mathbf{w}(n)\mathbf{w}^T(n)\mathbf{X}(n)\right]\right\} \\ &+ \beta^2 L\sigma_x^2\operatorname{tr}\left\{E\left[\mathbf{v}(n)\mathbf{v}^T(n)\right]\right\}. \end{aligned} \tag{54}$$

Besides, we could consider that the a posteriori misalignment at time index $n-1$ is not correlated with the input data matrix at time index n. Also, the correlation matrix of $\mathbf{w}(n)$ was considered to be diagonal. Taking these two issues into account, (54) becomes

$$\begin{aligned} m(n) &= m(n-1) + L\sigma_w^2 + \beta\left(\beta L\sigma_x^2 - 2\right) \\ &\times \operatorname{tr}\left\{E\left[\mathbf{X}(n)\mathbf{X}^T(n)\right]E\left[\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\right]\right\} \\ &+ \beta\left(\beta L\sigma_x^2 - 2\right)\operatorname{tr}\left\{E\left[\mathbf{X}(n)\mathbf{X}^T(n)\right]E\left[\mathbf{w}(n)\mathbf{w}^T(n)\right]\right\} \\ &+ \beta^2 L\sigma_x^2\operatorname{tr}\left\{E\left[\mathbf{v}(n)\mathbf{v}^T(n)\right]\right\}. \end{aligned} \tag{55}$$

Furthermore, in (55), we can also use: i) $E\left[\mathbf{X}(n)\mathbf{X}^T(n)\right] = P\mathbf{R}_{\mathbf{x},L}(n) \approx P\sigma_x^2\mathbf{I}_L$, where $\mathbf{R}_{\mathbf{x},L}(n)$ is the input signal correlation matrix of size $L\times L$, ii) $E\left[\mathbf{w}(n)\mathbf{w}^T(n)\right] = \sigma_w^2\mathbf{I}_L$, and iii) $E\left[\mathbf{v}(n)\mathbf{v}^T(n)\right] = \sigma_v^2\mathbf{I}_P$. Besides, the correlation matrix of the a posteriori misalignment at time index $n-1$, i.e., $E\left[\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\right]$, tends to be a diagonal one when the algorithm converges (since the misalignment of the individual coefficients tend to become uncorrelated). Thus,

$$E\left[\boldsymbol{\mu}(n-1)\boldsymbol{\mu}^T(n-1)\right] \approx \frac{1}{L}m(n-1)\mathbf{I}_L. \tag{56}$$

Consequently, (55) can be finally expressed as

$$\begin{aligned} m(n) &= \left[m(n-1) + L\sigma_w^2\right]\left[1 + \beta P\sigma_x^2\left(\beta L\sigma_x^2 - 2\right)\right] \\ &+ \beta^2 LP\sigma_x^2\sigma_v^2. \end{aligned} \tag{57}$$

2.3 An Optimized APA

Since β contains both the control parameters (i.e., α and δ) we could follow an optimization problem on this parameter, i.e., a joint-optimization on both the step-size and regularization parameters. Therefore, we could evaluate [based on (57)]

$$\frac{\partial m(n)}{\partial \beta} = 0, \tag{58}$$

which leads to

$$\beta = \frac{m(n-1) + L\sigma_w^2}{L\sigma_x^2 \left[m(n-1) + L\sigma_w^2\right] + L\sigma_v^2}. \tag{59}$$

Next, following the idea of a VSS algorithm, we can use (59) in conjunction with (36), (41), and (53), thus obtaining

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mathbf{X}(n) \frac{\mathbf{e}(n)}{\dfrac{L\sigma_v^2}{m(n-1) + L\sigma_w^2} + L\sigma_x^2}. \tag{60}$$

In order to update the parameter $m(n)$, we need to use (59) in (57), and after some computations we get

$$m(n) = \left[1 - \frac{P\sigma_x^2}{\dfrac{L\sigma_v^2}{m(n-1) + L\sigma_w^2} + L\sigma_x^2}\right] \times \left[m(n-1) + L\sigma_w^2\right]. \tag{61}$$

At a very first glance, the algorithm defined by (37), (60), and (61) looks very appealing since it does not require any matrix inversion (as compared to the classical APA). However, let us remember that we used the white input assumption [see (52) and (53)] in the development, which limits the applicability of the algorithm for correlated inputs (or, alternatively, it would require orthogonalized inputs, e.g., [36]). Consequently, we should reconsider this approximation in the resulted algorithm, so that the update (60) becomes

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mathbf{X}(n)\mathbf{R}(n)\mathbf{e}(n), \tag{62}$$

where

$$\mathbf{R}(n) = \left[\frac{L\sigma_v^2}{m(n-1) + L\sigma_w^2}\mathbf{I}_P + \mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1}, \tag{63}$$

while (61) can be approximated as

$$m(n) \approx \left\{1 - \frac{\mathrm{tr}\left\{\mathbf{X}^T(n)\mathbf{X}(n)\mathbf{R}(n)\right\}}{PL}\right\} \left[m(n-1) + L\sigma_w^2\right]. \tag{64}$$

The resulting algorithm is referred as the joint-optimized APA (JO-APA). In practice, we propose to evaluate the parameter σ_w^2 based on (35) as

$$\sigma_w^2(n) = \frac{1}{PL} \left\| \widehat{\mathbf{h}}(n) - \widehat{\mathbf{h}}(n-1) \right\|_2^2, \tag{65}$$

while the parameter σ_v^2 can be also estimated in different practical ways within the algorithm, e.g., [30], [40].

Apparently, (62) looks like the update of a VR-APA. However, the update (62) was not obtained by focusing on the parameter δ only (like in most VR-APAs), but on a joint-optimization on both important parameters α and δ. The resulting algorithm is very similar to the simplified general Kalman filter [20], [22]. However, contrary to this one (which was obtained as an approximation of the general Kalman filter), the JO-APA was derived in a different manner following a specific optimization criterion. Overall, the JO-APA could be interpreted as an algorithm that achieves the maximum convergence rate to reach a certain misalignment level, in the context of a state variable model. This is an alternative way to obtain the same results as with the Kalman filter.

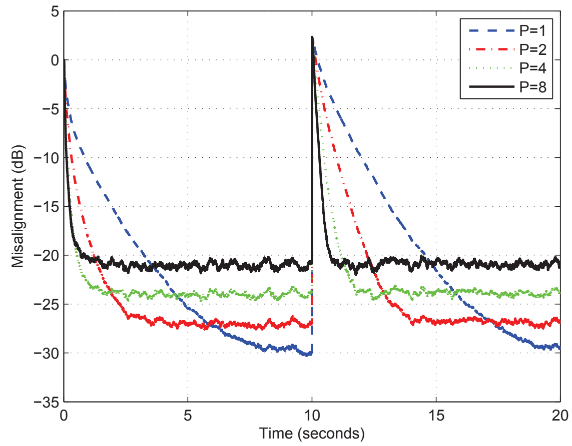

Figure 6: Misalignment of the JO-APA using different values of the projection order P and a constant value $\sigma_w^2 = 10^{-12}$. The input signal is an AR(1) process, $L = 512$, and SNR = 20 dB. Echo path changes after 10 seconds.

2.4 Simulation Results for Optimized APA

Experiments were performed in the context of AEC [16], [17]. The system to be identified is an acoustic impulse response of length $L = 512$ (the sampling rate is 8 kHz); its output (i.e., the echo signal) is corrupted by an independent white Gaussian noise signal with SNR = 20 dB. We assume that the power of this noise, σ_v^2, is known (for example, it can be estimated during silences, or like in [30], [40]). The input signal, $x(n)$, is either an AR(1) process generated by filtering a white Gaussian noise through a first-order system $1/\left(1 - 0.8z^{-1}\right)$ or a speech sequence. In order to evaluate the tracking capability of the algorithms, an echo path change scenario is simulated in the middle of each experiment, by shifting the impulse response to the right by 12 samples. The measure of performance is the normalized misalignment (in dB), defined as $20\log_{10}\left\|\mathbf{h}(n) - \widehat{\mathbf{h}}(n)\right\|_2 / \left\|\mathbf{h}(n)\right\|_2$.

In the first experiment, we compare the performance of the JO-APA when using different values of the projection order P and a constant value of the parameter σ_w^2. The results are presented in Fig. 6. As we can notice, similar to the APA, the convergence rate and tracking of the JO-APA (using a constant value of σ_w^2) improves when the value of the projection order P increases. However, this also leads to a higher misalignment.

Next, we evaluate the influence of the parameter σ_w^2 on the overall performance of the JO-APA. This parameter captures the uncertainties in the echo path and controls the behavior of the algorithm in terms of the compromise between fast convergence/tracking and low misadjustment. This issue is supported in Fig. 7, where the JO-APA using different constant values of σ_w^2 is compared with the JO-APA using $\sigma_w^2(n)$ estimated as in (65); the projection order is $P = 4$ and the input signal is an AR(1) process. It can be noticed that a lower value of σ_w^2 improves the misalignment but affects the tracking, while the estimated $\sigma_w^2(n)$ leads to a proper compromise between the performance criteria. Consequently, in the following experiments, only the JO-APA using the estimated $\sigma_w^2(n)$ from (65) will be used.

In Fig. 8, the JO-APA is compared to the APA using different values of the step-size parameter α. According to this figure, the JO-APA behaves like a VSS-APA, obtaining a convergence rate and tracking similar to the APA with the largest step-size, while achieving a much lower misalignment level.

In the second set of experiments, the input signal is speech. First, in Fig. 9, we compare the performance of the JO-APA when using different values of the projection order P. It can be noticed that the performance of the algorithm improves (in terms of both convergence/tracking and misalignment) when the projection order increases.

Finally, the JO-APA is compared to the VR-APA proposed in [11], which represents one of the benchmarks VR algorithms. Also, the APA is included for comparison, using different values of the step-size parameter. The results are presented in Figs. 10 and 11, using $P = 2$ and $P = 8$, respectively. According

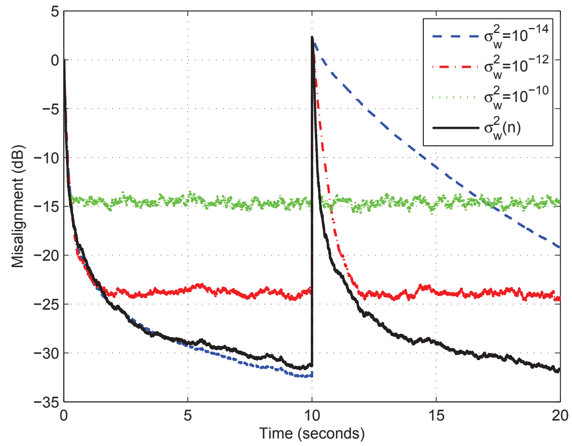

Figure 7: Misalignment of the JO-APA with different constant values of σ_w^2 and estimated $\sigma_w^2(n)$ from (65); the projection order is $P = 4$. The input signal is an AR(1) process, $L = 512$, and SNR = 20 dB. Echo path changes after 10 seconds.

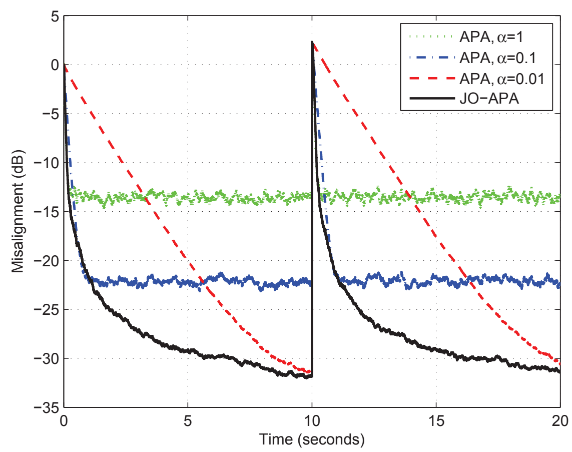

Figure 8: Misalignment of the APA (with different step-sizes) and JO-APA; other conditions same as in Fig. 7.

to these figures, the JO-APA outperforms both the APA and the VR-APA, especially for a larger value of the projection order.

In this section, we developed an optimized APA, in the framework of a state variable model (like in Kalman filtering) for system identification. Following a convergence analysis of the APA, the proposed JO-APA results based on a joint-optimization on both the step-size and regularization parameters, aiming to minimize the system misalignment. Simulations performed in the context of AEC indicate that the proposed algorithm achieves both fast convergence/tracking and low misadjustment, thus representing an attractive choice for this type of applications.

References

[1] L. M. Dogariu, J. Benesty, C. Paleologu, and S. Ciochină, "Identification of room acoustic impulse responses via Kronecker product decompositions," *IEEE/ACM Trans. Audio, Speech, Language Processing*, vol. 30, pp. 2828–2841, Sept. 2022.

[2] G. Huang, J. Benesty, J. Chen, C. Paleologu, S Ciochină, W. Kellermann, and I. Cohen, "Acoustic system identification with partially time-varying models based on tensor decompositions," in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2022, pp. 1–5.

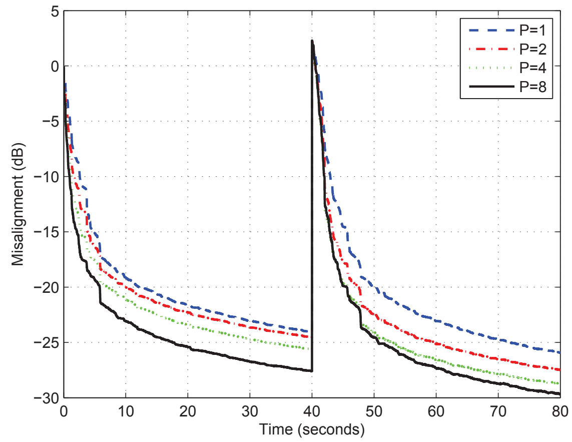

Figure 9: Misalignment of the JO-APA with different values of the projection order P. The input signal is speech, $L = 512$, and SNR = 20 dB. Echo path changes after 40 seconds.

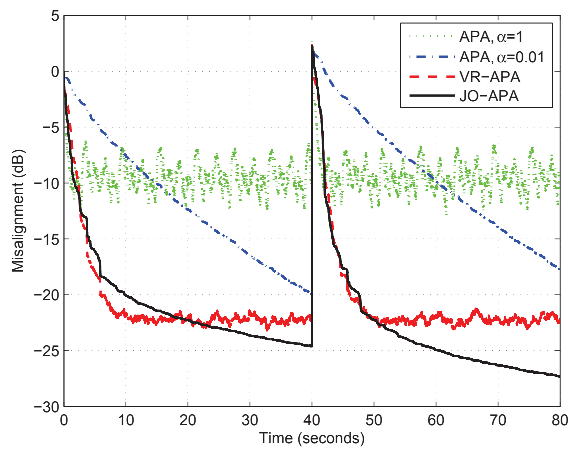

Figure 10: Misalignment of the APA (with different step-sizes), VR-APA [11], and JO-APA; the projection order is $P = 2$. The input signal is speech, $L = 512$, and SNR = 20 dB. Echo path changes after 40 seconds.

[3] B. Moroşanu, V. Popa, C. Negrescu, and I. D. Fîciu, "Control room design for subjective audio critical listening,", in *Proc. International Conference on Advanced Topics in Optoelectronics, Microelectronics and Nanotechnologies (ATOM-N)*, 2022, pp. 1–8.

[4] J. Benesty, L. M. Dogariu, C. Paleologu, and S. Ciochină, "An iterative Wiener filter for stereophonic acoustic echo cancellation," in *Proc. IEEE Asilomar Conference on Signals, Systems, and Computers*, 2022, pp. 1–5.

[5] I. D. Fîciu, L. M. Dogariu, R. L. Costea, C. Paleologu, J. Benesty, and S. Ciochină, "A Kalman filter for the identification of impulse responses with particular symmetry features," *in Proc. International Symposium on Electronics and Telecommunications (ISETC)*, 2022, pp. 1–4.

[6] I. D. Fîciu, C. L. Stanciu, C. Elisei-Iliescu, C. Anghel, and C. Paleologu, "A data-reuse approach for the RLS-DCD algorithm," in *Proc. IARIA Congress*, 2022, pp. 85–86.

[7] A. H. Sayed, *Adaptive Filters*. New York, NY: Wiley, 2008.

[8] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Fourth Edition, New York, NY: Springer-Verlag, 2013.

[9] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Processing Lett.*, vol. 11, pp. 132–135, Feb. 2004.

[10] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A nonparametric VSS-NLMS algorithm," *IEEE Signal Processing Lett.*, vol. 13, pp. 581–584, Oct. 2006.

[11] H. Rey, L. Rey Vega, S. Tressens, and J. Benesty, "Variable explicit regularization in affine projection algorithm: robustness issues and optimal choice," *IEEE Trans. Signal Processing*, vol. 55, pp. 2096–2108, May 2007.

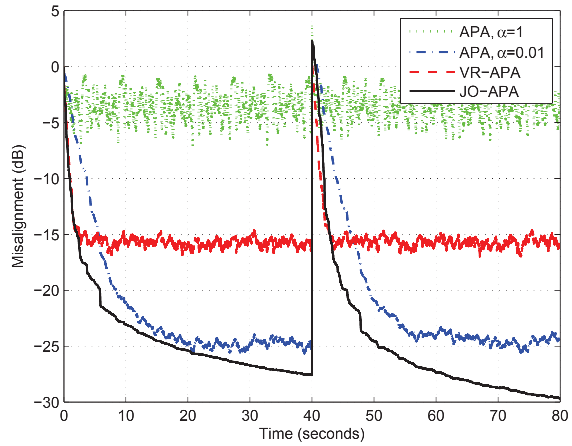

Figure 11: Misalignment of the APA (with different step-sizes), VR-APA [11], and JO-APA; the projection order is $P = 8$ and other conditions same as in Fig. 10.

[12] S. Werner, J. A. Apolinário Jr., and P. S. R. Diniz, "Set-membership proportionate affine projection algorithms," *EURASIP J. Audio, Speech, Music Processing*, vol. 2007, no. 1, pp. 1–10, 2007.

[13] S. Ciochină, C. Paleologu, J. Benesty, S. L. Grant, and A. Anghel, "A family of optimized LMS-based algorithms for system identification," in *Proc. EUSIPCO*, 2016, pp. 1803–1807.

[14] S. Ciochină, C. Paleologu, and J. Benesty, "An optimized NLMS algorithm for system identification," *Signal Processing*, vol. 118, pp. 115–121, Jan. 2016.

[15] C. Paleologu, J. Benesty, and S. Ciochină, *Sparse Adaptive Filters for Echo Cancellation*. Morgan & Claypool Publishers, Synthesis Lectures on Speech and Audio Processing, 2010.

[16] J. Benesty, T. Gaensler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.

[17] G. Enzner, H. Buchner, A. Favrot, and F. Kuech, "Acoustic echo control," in *Academic Press Library in Signal Processing*, vol. 4, ch. 30, pp. 807–877, Academic Press 2014.

[18] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Engineering*, vol. 82, pp. 35–45, Mar. 1960.

[19] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, pp. 18–60, July 1994.

[20] C. Paleologu, J. Benesty, and S. Ciochină, "Study of the general Kalman filter for echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, pp. 1539–1549, Aug. 2013.

[21] P. A. C. Lopes and J. B. Gerald, "New normalized LMS algorithms based on the Kalman filter," in *Proc. IEEE ISCAS*, 2007, pp. 117–120.

[22] G. Enzner, "Bayesian inference model for applications of time-varying acoustic system identification," in *Proc. EUSIPCO*, 2010, pp. 2126–2130.

[23] S. Ciochină, C. Paleologu, J. Benesty, and S. L. Grant, "An optimized proportionate adaptive algorithm for sparse system identification," in Proc. *IEEE Asilomar*, 2015, 5 pages.

[24] C. Paleologu, S. Ciochină, and J. Benesty, "Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation," *IEEE Signal Processing Lett.*, vol. 15, pp. 5–8, 2008.

[25] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithms for echo cancellation," in *Proc. IEEE ICASSP*, 2008, pp. 241–244.

[26] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE ICASSP*, 2002, pp. II-1881–II-1884.

[27] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn.*, vol. 67-A, pp. 19–27, May 1984.

[28] S. L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proc. IEEE ICASSP*, 1995, pp. 3023–3026.

[29] Y.-S. Choi, H.-C. Shin, and W.-J. Song, "Adaptive regularization matrix for affine projection algorithm," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 54, pp. 1087–1091, Dec. 2007.

[30] C. Paleologu, J. Benesty, and S. Ciochină, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 16, pp. 1466–1478, Nov. 2008.

[31] S.-E. Kim, S.-J. Kong, and W.-J. Song, "An affine projection algorithm with evolving order," *IEEE Signal Processing Lett.*, vol. 16, pp. 937–940, Nov. 2009.

[32] W. Yin and A. S. Mehr, "A variable regularization method for affine projection algorithm," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 57, pp. 476–480, June 2010.

[33] F. Albu, C. Paleologu, and J. Benesty, "A variable step size evolutionary affine projection algorithm," in *Proc. IEEE ICASSP*, 2011, pp. 429–432.

[34] C. H. Lee and P. Park, "Optimal step-size affine projection algorithm," *IEEE Signal Processing Lett.*, vol. 19, pp. 431–434, July 2012.

[35] C. H. Lee and P. Park, "Scheduled-step-size affine projection algorithm," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 59, pp. 2034–2043, Sept. 2012.

[36] P. Park, J.-H. Seo, and N. Kong, "Variable matrix-type step-size affine projection algorithm with orthogonalized input vectors," *Signal Processing*, vol. 98, pp. 135–142, May 2014.

[37] S. Ciochină, C. Paleologu, J. Benesty, and C. Anghel, "An optimized affine projection algorithm for acoustic echo cancellation," in *Proc. IEEE International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, 2015, pp. 159–164.

[38] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo cancellation," *Signal Processing*, vol. 86, pp. 1140–1156, 2006.

[39] G. Enzner, *A Model-Based Optimum Filtering Approach to Acoustic Echo Control: Theory and Practice*. Dissertation, RWTH Aachen, Aachener Beitrge zu digitalen Nachrichtensystemen, Vary P. (ed.), Wissenschaftsverlag Mainz, Aachen, June 2006.

[40] S. Malik and G. Enzner, "Online maximum-likelihood learning of time-varying dynamical models in block-frequency-domain," in *Proc. IEEE ICASSP*, 2010, pp. 3822–3825.