PRACTICAL REGULARIZATION OF THE AFFINE PROJECTION ALGORITHM

Constantin Paleologu†, *Jacob Benesty*‡, *and Silviu Ciochină*†

†Telecommunications Department, University Politehnica of Bucharest, Romania
email: {pale,silviu}@comm.pub.ro
‡INRS-EMT, University of Quebec, Montreal, Canada
email: benesty@emt.inrs.ca

ABSTRACT

It is well-known that a matrix inversion is required within the affine projection algorithm (APA). Depending on the character of the input signal, this matrix can be very ill-conditioned. Consequently, it needs to be regularized before inversion, i.e., a positive constant is added to the elements of its main diagonal. Also known as the regularization parameter, this constant plays a major role in practice; if it is not chosen properly, the APA may never converge, especially under low signal-to-noise ratio conditions. The contribution of this paper is twofold. First, we provide a formula for choosing an "optimal" regularization parameter, aiming at attenuating the effects of the noise in the adaptive filter estimate. Second, two practical ways for evaluating this parameter are proposed. Simulations performed in the context of acoustic echo cancellation (in different noisy environments) prove the validity of the proposed solutions.

1. INTRODUCTION

The main advantage of the affine projection algorithm (APA) [1] over the normalized least-mean-square (NLMS) algorithm [2] consists of a superior convergence rate, especially for correlated inputs. For this reason, the APA and different versions of it were found to be very attractive choices in many applications. For example, in echo cancellation [3], we deal with long adaptive filters and highly-correlated signals (like speech). Consequently, the APA is often the algorithm of choice for this very popular application.

A matrix inversion is required within the classical APA. In order to avoid any numerical problems (caused by an ill-conditioned matrix), a diagonal matrix is added to the matrix to be inverted. This diagonal matrix is often chosen as the identity matrix multiplied by a positive constant called the regularization parameter [2], [4].

It is known from practice that the value of the regularization parameter is highly influenced by the level of the system noise, i.e., the noise that corrupts the output of the system that we need to identify. In echo cancellation, this is the near-end background noise which corrupts the output of the echo path. The rule of thumb is: the more the noise, the larger the value of the regularization parameter [3]. However, it is not easy to tune the value of this constant in noisy environments [5]. As an alternative, several approaches have focused on finding a variable regularization parameter (instead of a constant one), thus resulting the so-called variable-regularized APAs (VR-APAs), e.g., [6], [7], [8]. The basic approach in most of these algorithms is to set the step-size parameter equal to one and to control the adaptation of the algorithm only by using a variable (i.e., time-dependent) regularization parameter.

The goal of this paper is twofold. First, we provide an "optimal" formula for choosing the constant regularization parameter of the APA, based on a condition that intuitively makes sense, i.e., to attenuate the effects of the noise in the adaptive filter estimate. Second, we propose two practical methods for evaluating the "optimal" regularization parameter in real-world applications. The validity of the proposed solutions are demonstrated in the context of acoustic echo cancellation.

2. OPTIMAL REGULARIZATION OF THE APA

In a system identification problem [2], which is the basic configuration for echo cancellation, the observed or desired signal is obtained as

$$\begin{aligned} d(n) &= \mathbf{h}^T\mathbf{x}(n) + w(n) \\ &= y(n) + w(n), \end{aligned} \quad (1)$$

where n is the discrete-time index,

$$\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{L-1} \end{bmatrix}^T \quad (2)$$

is the impulse response (of length L) of the system that we need to identify (i.e., the echo path), superscript ^{T} denotes transpose of a vector or a matrix,

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L+1) \end{bmatrix}^T \quad (3)$$

is a vector containing the most recent L samples of the zero-mean input signal $x(n)$ (i.e., the far-end signal), $w(n)$ is a zero-mean additive noise signal (i.e., the near-end background noise), which is independent of $x(n)$, and the output of the unknown system, $y(n)$, is the echo signal. In order to cancel this echo, the basic approach is to estimate or identify $\mathbf{h}$ with an adaptive filter

$$\hat{\mathbf{h}}(n) = \begin{bmatrix} \hat{h}_0(n) & \hat{h}_1(n) & \cdots & \hat{h}_{L-1}(n) \end{bmatrix}^T, \quad (4)$$

which updates its coefficients based on an adaptive algorithm. Due to its convergence features, the APA is one of the most popular algorithms used for echo cancellation. The classical version of this algorithm is defined by the following equations [1]:

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \quad (5)$$

$$\mathbf{R}(n) = \delta\mathbf{I}_P + \mathbf{X}^T(n)\mathbf{X}(n), \quad (6)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \alpha\mathbf{X}(n)\mathbf{R}^{-1}(n)\mathbf{e}(n), \quad (7)$$

where $\mathbf{e}(n)$ is the error signal vector of length P (with P denoting the projection order),

$$\mathbf{d}(n) = \begin{bmatrix} d(n) & d(n-1) & \cdots & d(n-P+1) \end{bmatrix}^T \quad (8)$$

is a vector containing the most recent P samples of the desired signal,

$$\mathbf{X}(n) = \begin{bmatrix} \mathbf{x}(n) & \mathbf{x}(n-1) & \cdots & \mathbf{x}(n-P+1) \end{bmatrix} \quad (9)$$

is the input data matrix (of size $L \times P$), $\mathbf{R}(n)$ is the correlation matrix to be inverted (of size $P \times P$), δ is the regularization parameter, $\mathbf{I}_P$ is the $P \times P$ identity matrix, and α is the step-size parameter. It can be easily noticed that the well-known NLMS algorithm is obtained when $P = 1$. When the projection order increases, the convergence rate of the APA also increases; of course, this also leads to an increased computational complexity.

Besides the projection order, the performance of the APA is controlled by two other important parameters. The first one is the step-size parameter α (usually, $0 < \alpha \leq 1$), which multiplies the update term in (7) in order to achieve a proper compromise between the convergence rate and the misadjustment [2]. The second one (of equal importance) is the regularization parameter δ. As we have mentioned in the previous section, the regularization matrix $\delta \mathbf{I}_P$ is introduced in order to prevent the problems associated with the inverse of the matrix $\mathbf{X}^T(n)\mathbf{X}(n)$, which can be very ill-conditioned. Despite its practical importance, the problem of choosing the step size was much often addressed in the literature as compared to the selection of the regularization parameter; however, the selection of δ has a great impact in terms of the adaptive filter performance.

It is known from practice that the value of the regularization parameter δ depends on the level of the noise [i.e., $w(n)$ in (1)] that corrupts the output of the system. From (1), we define the echo-to-noise ratio (ENR) [3], which is also the signal-to-noise ratio (SNR), as

$$\mathrm{ENR} = \frac{\sigma_y^2}{\sigma_w^2}, \tag{10}$$

where $\sigma_y^2 = E\left[y^2(n)\right]$ and $\sigma_w^2 = E\left[w^2(n)\right]$ are the variances of $y(n)$ and $w(n)$, respectively, with $E(\cdot)$ denoting the expectation. Low ENRs require high values of the regularization parameter, while its importance becomes less apparent for high ENRs. However, this is just a rule of thumb in practice; therefore, it would be very useful to know how large or small should be chosen the value of δ as a function of the ENR. In order to provide a solution for this problem, let us first rewrite (7) as

$$\hat{\mathbf{h}}(n) = \mathbf{P}(n)\hat{\mathbf{h}}(n-1) + \alpha\tilde{\mathbf{h}}(n), \tag{11}$$

where

$$\mathbf{P}(n) = \mathbf{I}_L - \alpha\mathbf{X}(n)\mathbf{R}^{-1}(n)\mathbf{X}^T(n), \tag{12}$$

$\mathbf{I}_L$ is the $L \times L$ identity matrix, and

$$\tilde{\mathbf{h}}(n) = \mathbf{X}(n)\mathbf{R}^{-1}(n)\mathbf{d}(n). \tag{13}$$

We can see that the vector $\tilde{\mathbf{h}}(n)$ is the correction component of the APA, since it depends on the new observation $\mathbf{d}(n)$. But it also can be noticed that the matrix $\mathbf{P}(n)$ does not depend on the noise signal or the desired signal (it depends only on the input signal). In fact, $\tilde{\mathbf{h}}(n)$ is obtained by solving

$$\min_{\tilde{\mathbf{h}}(n)} \left\{ \left[\mathbf{d}(n) - \mathbf{X}^T(n)\tilde{\mathbf{h}}(n)\right]^T \left[\mathbf{d}(n) - \mathbf{X}^T(n)\tilde{\mathbf{h}}(n)\right] + \delta \left\|\tilde{\mathbf{h}}(n)\right\|_2^2 \right\}, \tag{14}$$

where $\|\cdot\|_2$ denotes the ℓ_2 norm. The previous optimization is the regularized version of the minimum ℓ_2-norm solution of the linear system of P equations $\mathbf{d}(n) = \mathbf{X}^T(n)\tilde{\mathbf{h}}(n)$. Clearly, the solution $\tilde{\mathbf{h}}(n)$ is not the optimal one; thus, the other vector $\mathbf{P}(n)\hat{\mathbf{h}}(n-1)$ in (11) can be seen as a good initialization of the adaptive filter.

Taking the previous considerations into account, let us define the new error signal vector

$$\tilde{\mathbf{e}}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\tilde{\mathbf{h}}(n), \tag{15}$$

which is the difference between the desired signal vector and the estimated signal vector obtained from the filter optimized in (14). In order to attenuate the effects of the noise in the estimator $\tilde{\mathbf{h}}(n)$, it is reasonable to find δ in such a way that [9]

$$E\left[\|\tilde{\mathbf{e}}(n)\|_2^2\right] = E\left[\|\mathbf{w}(n)\|_2^2\right], \tag{16}$$

where $\mathbf{w}(n) = \begin{bmatrix} w(n) & w(n-1) & \cdots & w(n-P+1) \end{bmatrix}^T$ is a vector containing the most recent P samples of the system noise.

Following condition (16), we can use (13) in (15) to get

$$\tilde{\mathbf{e}}(n) = \left[\mathbf{I}_P - \mathbf{X}^T(n)\mathbf{X}(n)\mathbf{R}^{-1}(n)\right]\mathbf{d}(n). \tag{17}$$

We also have the eigenvalue decomposition:

$$\mathbf{X}^T(n)\mathbf{X}(n) = \mathbf{V}(n)\mathbf{\Lambda}(n)\mathbf{V}^T(n), \tag{18}$$

where $\mathbf{V}(n)$ is an orthogonal matrix containing the eigenvectors of $\mathbf{X}^T(n)\mathbf{X}(n)$ as columns and $\mathbf{\Lambda}(n)$ is a diagonal matrix containing the corresponding eigenvalues $\lambda_p(n)$, with $p = 1, 2, \ldots, P$. Hence, the inverse of the matrix from (6) is

$$\mathbf{R}^{-1}(n) = \mathbf{V}(n)\left[\delta\mathbf{I}_P + \mathbf{\Lambda}(n)\right]^{-1}\mathbf{V}^T(n). \tag{19}$$

Consequently, based on (17)–(19), we get

$$\begin{aligned} \|\tilde{\mathbf{e}}(n)\|_2^2 &= \mathbf{d}^T(n)\mathbf{V}(n) \\ &\cdot \left\{\mathbf{I}_P - \mathbf{\Lambda}(n)\left[\delta\mathbf{I}_P + \mathbf{\Lambda}(n)\right]^{-1}\right\}^2 \mathbf{V}^T(n)\mathbf{d}(n). \end{aligned} \tag{20}$$

At this point, we should further process (20). But this is difficult without any supporting assumptions on the character of the input signal. Therefore, in order to facilitate the analysis, let us assume that the input signal is white, so that the eigenvalues of the matrix $\mathbf{X}^T(n)\mathbf{X}(n)$ [see (18)] are

$$\lambda_p \approx L\sigma_x^2, \text{ for } p = 1, 2, \ldots, P, \tag{21}$$

where $\sigma_x^2 = E\left[x^2(n)\right]$ is the variance of the input signal $x(n)$. Thus, (20) can be expressed as

$$\|\tilde{\mathbf{e}}(n)\|_2^2 = \mathbf{d}^T(n)\mathbf{V}(n)\mathbf{V}^T(n)\mathbf{d}(n)\left(\frac{\delta}{\delta + L\sigma_x^2}\right)^2. \tag{22}$$

Consequently, taking the expectation on both sides of (22), the condition (16) becomes

$$E\left[\|\mathbf{d}(n)\|_2^2\right]\left(\frac{\delta}{\delta + L\sigma_x^2}\right)^2 = E\left[\|\mathbf{w}(n)\|_2^2\right]. \tag{23}$$

We also know that the echo signal and the system noise are uncorrelated, so that (23) can be developed as

$$\begin{aligned} \left\{E\left[\|\mathbf{y}(n)\|_2^2\right] + E\left[\|\mathbf{w}(n)\|_2^2\right]\right\}\left(\frac{\delta}{\delta + L\sigma_x^2}\right)^2 \\ = E\left[\|\mathbf{w}(n)\|_2^2\right], \end{aligned} \tag{24}$$

where $\mathbf{y}(n) = \begin{bmatrix} y(n) & y(n-1) & \cdots & y(n-P+1) \end{bmatrix}^T$ contains the most recent P samples of the echo signal. Finally, knowing that $E\left[\|\mathbf{y}(n)\|_2^2\right] = P\sigma_y^2$ and $E\left[\|\mathbf{w}(n)\|_2^2\right] = P\sigma_w^2$, and taking (10) into account, the condition (24) becomes

$$\left(\frac{\delta}{\delta + L\sigma_x^2}\right)^2 = \frac{1}{1 + \mathrm{ENR}}, \tag{25}$$

which results in

$$\mathrm{ENR} \cdot \delta^2 - 2\left(L\sigma_x^2\right)\delta - \left(L\sigma_x^2\right)^2 = 0. \tag{26}$$

Solving the quadratic equation (26) we get the obvious solution

$$\begin{aligned} \delta &= \frac{L\left(1 + \sqrt{1 + \mathrm{ENR}}\right)}{\mathrm{ENR}}\sigma_x^2 \\ &= \beta_{\mathrm{APA}}\sigma_x^2, \end{aligned} \tag{27}$$

where

$$\beta_{\mathrm{APA}} = \frac{L\left(1+\sqrt{1+\mathrm{ENR}}\right)}{\mathrm{ENR}} \tag{28}$$

is the normalized (with respect to the variance of the input signal) regularization parameter of the APA.

The result in (27) justifies a well-known practical rule, i.e., the parameter δ is usually taken as

$$\delta = \beta\sigma_x^2, \tag{29}$$

where β is a positive constant (i.e., the normalized regularization parameter). Also, we know that $\beta = 20$ is an "ad-hoc" choice in many echo cancellation scenarios. However, as the simulations will show, this value is not valid any more when the system noise is high (i.e., corresponding to a low ENR). Therefore, (28) provides a more rigorous way to choose the "optimal" normalized regularization parameter as a function of the ENR.

It is important to notice that (27) is in consistence with the dependence between the regularization parameter and the level of the system noise, i.e., $\lim_{\mathrm{ENR}\to\infty}\delta = 0$ and $\lim_{\mathrm{ENR}\to 0}\delta = \infty$. Also, since β_{APA} does not depend on the projection order P, the "optimal" regularization parameter is also valid for the NLMS algorithm [10].

3. PRACTICAL EVALUATION OF THE OPTIMAL REGULARIZATION PARAMETER

It can be noticed from (27) that δ depends on three elements: the length, L, of the adaptive filter, the variance, σ_x^2, of the input signal, and the ENR. In both network and acoustic echo cancellation, the first two elements (L and σ_x^2) are known, while the ENR could be estimated since it depends on the power of the system noise. Consequently, (27) provides an "optimal" value for the regularization constant, when some information about the system noise power is available.

Based on these considerations, it would be useful to find some practical methods for estimating the ENR [which is the key parameter in (27)], in order to provide the value of δ for real-world applications. Since the echo signal and the system noise are uncorrelated, (10) can be rewritten as

$$\begin{aligned}\mathrm{ENR} &= \frac{\sigma_d^2 - \sigma_w^2}{\sigma_w^2} \\ &= \frac{\sigma_d^2}{\sigma_w^2} - 1,\end{aligned} \tag{30}$$

where $\sigma_d^2 = E\left[d^2(n)\right]$ is the variance of $d(n)$. In echo cancellation, the power of the background noise can be estimated during silence, while the variance of $d(n)$ can be recursively evaluated as

$$\hat{\sigma}_d^2(n) = \gamma\hat{\sigma}_d^2(n-1) + (1-\gamma)d^2(n), \tag{31}$$

where $\gamma = 1 - 1/(KL)$ (with $K > 1$) is an exponential window and the initial value is $\hat{\sigma}_d^2(0) = 0$. Consequently, using (31), the parameter from (30) can be estimated as

$$\widehat{\mathrm{ENR}}(n) = \frac{\hat{\sigma}_d^2(n)}{\sigma_w^2} - 1. \tag{32}$$

Using (32) in (27) results in a first practical regularized APA (PR-APA-1).

However, it is not always easy to estimate the parameter σ_w^2 in practice. For example, in echo cancellation, the signal $w(n)$ is in fact the near-end signal, which could contain both the background noise and the near-end speech. Therefore, the power of the near-end signal should be estimated. In this case, we could use a recently proposed method to evaluate this parameter [11]. Assuming that the adaptive filter has converged to a certain degree, it can be considered that

$$E\left[y^2(n)\right] \approx E\left[\hat{y}^2(n)\right], \tag{33}$$

where $\hat{y}(n) = \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n)$ is the output of the adaptive filter. Consequently, in terms of power estimates,

$$\hat{\sigma}_w^2(n) \approx \hat{\sigma}_d^2(n) - \hat{\sigma}_{\hat{y}}^2(n), \tag{34}$$

where $\hat{\sigma}_{\hat{y}}^2(n)$ can be evaluated in a similar way to (31). Using (34) in (30), we get a second relation for the estimation of the ENR, i.e.,

$$\widehat{\mathrm{ENR}}(n) = \frac{\hat{\sigma}_{\hat{y}}^2(n)}{\hat{\sigma}_d^2(n) - \hat{\sigma}_{\hat{y}}^2(n)}. \tag{35}$$

Therefore, using (35) in (27) results in a second practical regularized APA (PR-APA-2).

Finally, some practical issues should be addressed. First, for numerical reasons (since power estimates are used), it is recommended to use the absolute values in both (32) and (35); besides, a very small positive number should be added to the denominator of (35), in order to avoid division by zero. Second, it is recommended to use a constant regularization (e.g., the "classical" $\beta = 20$) in the first N iterations of the algorithms (with $N \geq L$), in order to obtain some valid values for the power estimates and also to let the adaptive filter converge to a certain degree [see assumption (33)].

It is clear now that the regularization parameters of both PR-APA-1 and PR-APA-2 are not constants anymore, but they are time-dependent; hence, these algorithms look similar to the VR-APAs discussed in the introduction. However, the role of the step-size parameter α is not omitted in our approach, as assumed in most of the VR-APAs. For this reason, we prefer not to include the proposed PR-APAs in the family of VR-APAs.

4. SIMULATION RESULTS

In the context acoustic echo cancellation, an adaptive filter is used to identify the acoustic echo path between the loudspeaker and the microphone of a hands-free device [3]. In this case, the level of the background noise can be high, so that the influence of the regularization parameter is critical. This is the framework of our experimental results.

The measured acoustic impulse response used in simulations has 512 coefficients and the same length is used for the adaptive filter, i.e., $L = 512$; the sampling rate is 8 kHz. The far-end (input) signal, $x(n)$, is either an AR(1) process generated by filtering a white Gaussian noise through a first-order system $1/(1-0.8z^{-1})$, or a speech sequence. Even if (27) was derived based on the white input assumption, simulation results show the validity of this formula for different types of input signals.

An independent white Gaussian noise, $w(n)$, is added to the echo signal $y(n)$, with different values of the ENR, i.e., 30 dB, 10 dB, and 5 dB. In most of the presented experiments, an echo path change scenario is simulated by shifting the acoustic impulse response to the right by 12 samples, in order to evaluate the tracking capabilities of the adaptive filter. The normalized misalignment (in dB), defined as

$$\mu(n) = 20\log_{10}\frac{\left\|\mathbf{h} - \hat{\mathbf{h}}(n)\right\|_2}{\|\mathbf{h}\|_2} \tag{36}$$

is used as the performance measure.

We choose to compare the APA using the "classical" $\beta = 20$ with the APA using the "optimal" β_{APA} (assuming that the true value of the ENR is known), the PR-APA-1 (assuming that the power of the system noise, σ_w^2, is known), and the PR-APA-2. Due

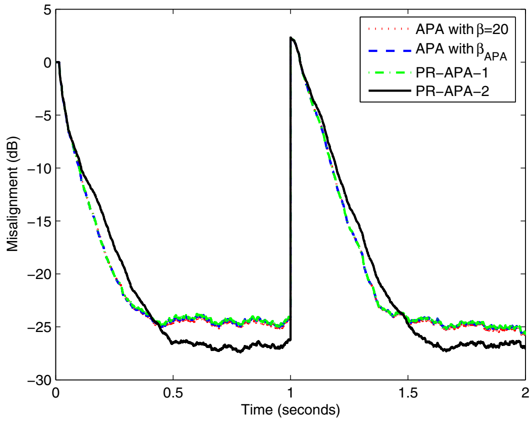

Figure 1: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. The input signal is an AR(1), $P = 2$, $\alpha = 1$, $L = 512$, and ENR = 30 dB. Echo path changes at time 1 second.

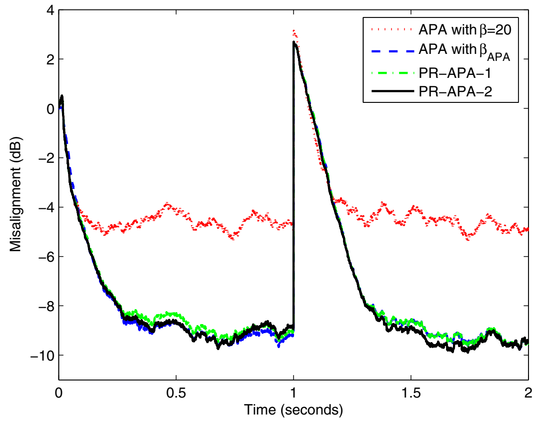

Figure 3: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. The input signal is an AR(1), $P = 2$, $\alpha = 1$, $L = 512$, and ENR = 10 dB. Echo path changes at time 1 second.

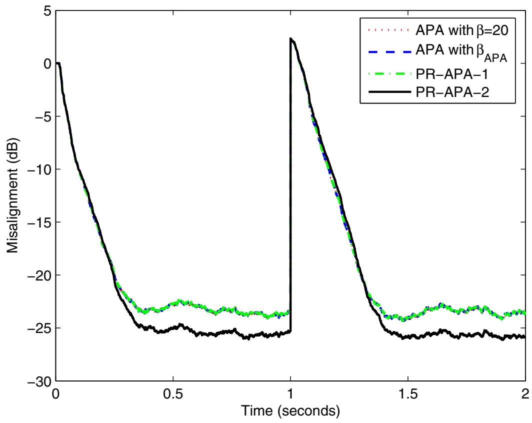

Figure 2: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. The input signal is an AR(1), $P = 4$, $\alpha = 1$, $L = 512$, and ENR = 30 dB. Echo path changes at time 1 second.

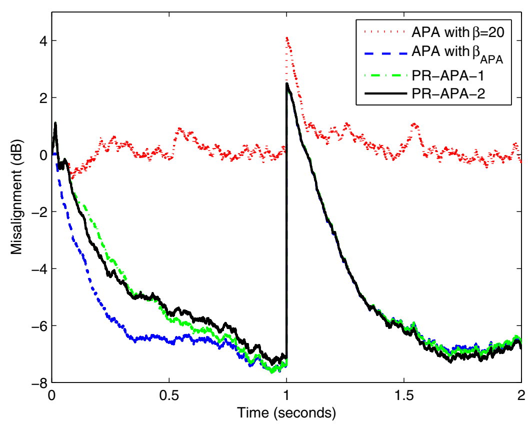

Figure 4: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. The input signal is an AR(1), $P = 2$, $\alpha = 1$, $L = 512$, and ENR = 5 dB. Echo path changes at time 1 second.

to the lack of space, we do not include here comparisons with other constant values of β. However, extensive simulations proving the validity of the "optimal" regularization parameter can be found in [9]. As explained in the previous section, both PR-APAs use a constant normalized regularization parameter $\beta = 20$ in the first $N = L = 512$ iterations; also, the exponential windows in the PR-APAs use $K = 6$.

For practical reasons, in terms of computational complexity, the projection order is fixed to $P = 2$ in most of the simulations (except one experiment using $P = 4$). The step-size parameter of the algorithms is set to $\alpha = 1$ when the AR(1) process is used as input (in order to provide the fastest convergence rate for the adaptive filter, thus outlining the influence of the regularization parameter) and to $\alpha = 0.5$ for speech input. Two scenarios are evaluated, i.e., 1) single-talk, when only the background noise is considered at the near-end and 2) double-talk, when the near-end speech is also present. For the second scenario (reported in the last experiment), we do not use a double-talk detector (DTD) [3] to inhibit the adaptation process; in this way, we aim to evaluate the robustness of the algorithms during double-talk situations.

In the first experiment, the input signal is an AR(1), $P = 2$, and ENR = 30 dB. The results are given in Fig. 1. First, it can be noticed that the performance of the APA using the "optimal" normalized regularization is similar to the "classical" $\beta = 20$. This is expected, since the the value of β_{APA} is very close to 20 in this case (i.e., $\beta_{\mathrm{APA}}|_{\mathrm{ENR=30dB}} \approx 16.7$). Also, the PR-APA-1 matches very well the performance of the APA using the "optimal" β_{APA}. The PR-APA-2 achieves a slightly slower convergence rate and tracking due to the assumption (33); however, this leads to a lower misalignment level.

The previous limitation of the PR-APA-2 can be overcome by using a higher projection order. In Fig. 2, the previous experiment is repeated using $P = 4$. It can be noticed that the convergence rate and tracking capability of the PR-APA-2 is clearly improved, while keeping the lowest misalignment level.

The importance of the regularization parameter becomes more apparent when the level of the background noise increases. The first experiment is repeated in Fig. 3, but using a lower value of the ENR, i.e., 10 dB. In this case, the classical choice $\beta = 20$ does not lead to a satisfactory performance. Also, it can be noticed that the performance obtained using the "optimal" normalized regularization is similar to the case of both PR-APAs.

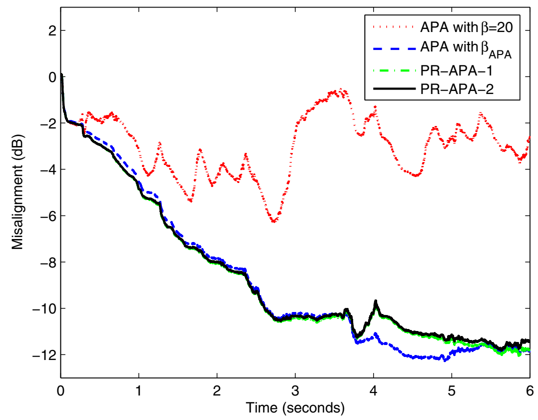

Figure 5: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. Single-talk scenario, the input signal is speech, $P = 2$, $\alpha = 0.5$, $L = 512$, and ENR = 10 dB.

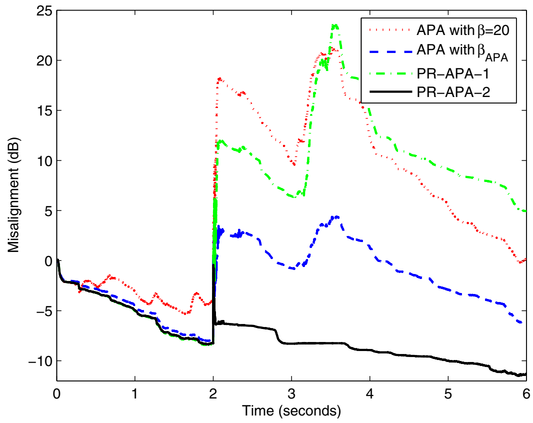

Figure 6: Misalignment of the APA using $\beta = 20$, APA using the "optimal" β_{APA}, PR-APA-1, and PR-APA-2. Double-talk scenario without DTD (the near-end speech appears between time 2 seconds and time 4 seconds), other conditions as in Fig. 5.

For the experiment provided in Fig. 4 the ENR is further decreased, i.e., it is set to 5 dB. In this case, the APA using the "classical" $\beta = 20$ is outperformed by the other algorithms based on the "optimal" regularization. The slower initial convergence rate of both PR-APAs is due to the use of an improper value of the constant regularization (i.e., $\delta = 20\sigma_x^2$) in the first L iterations; however, they perform very well in terms of tracking.

It is known that the character of the input signal significantly influences the performance of the adaptive filters. However, the APA is recognized for its good performance even for highly-correlated and non-stationary inputs like speech. Figure 5 presents the behavior of the algorithms with the speech sequence as input, ENR = 10 dB, $P = 2$, and $\alpha = 0.5$. It can be noticed that the regularization process is critical in this case. For an improper value of the normalized regularization constant (e.g., the "classical" $\beta = 20$), the misalignment of the adaptive filter fluctuates a lot. It is clear that the APA using the "optimal" value, β_{APA}, but also the PR-APAs perform much better in this scenario.

Finally, the previous experiment is repeated in a double-talk scenario (Fig. 6). The near-end speech appears between time 2 seconds and time 4 seconds. In this case, the advantage of the PR-APA-2 (in terms of robustness during double-talk) becomes more apparent. This is due to the estimation of the near-end signal power given in (34), which incorporates not only the contribution of the background noise (as in the case of the other algorithms) but also the influence of the near-end speech.

5. CONCLUSIONS

The "ad-hoc" (classical) regularization of the APA, i.e., a constant that multiplies the variance of the input signal, is difficult to tune at low ENRs. We know from practice that it is critical to properly regularize adaptive filters especially in noisy environments; otherwise, their performance are far from satisfactory.

The contribution of this paper was two fold. First, we have developed a formula for choosing the constant value of the regularization parameter of the APA. The basic condition behind this approach was to attenuate the effects of the system noise in the adaptive filter estimate. The resulted formula provides a more rigorous way to choose the regularization parameter as a function of the ENR. Second, we have proposed two practical ways to evaluate this "optimal" regularization parameter, by estimating the ENR within the algorithm. Simulations performed in the context of acoustic echo cancellation prove the validity of the approach for different ENR levels and variations (like double-talk).

Acknowledgment

This work was supported under the Grant POSDRU/89/1.5/S/62557.

REFERENCES

[1] 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.

[2] S. Haykin, *Adaptive Filter Theory*. Fourth Edition, Upper Saddle River, NJ: Prentice-Hall, 2002.

[3] 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.

[4] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ: John Wiley & Sons, Inc., 2003.

[5] T. van Waterschoot, G. Rombouts, and M. Moonen, "Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement," *Signal Processing*, vol. 88, pp. 594–611, Mar. 2008.

[6] 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 Process.*, vol. 55, pp. 2096–2108, May 2007.

[7] 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.

[8] 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.

[9] C. Paleologu, J. Benesty, and S. Ciochină, "Regularization of the affine projection algorithm," submitted to *IEEE Trans. Circuits and Systems II: Express Briefs*.

[10] J. Benesty, C. Paleologu, and S. Ciochină, "On the regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Language Processing*, to appear.

[11] 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.