REGULARIZATION OF THE IMPROVED PROPORTIONATE AFFINE PROJECTION ALGORITHM

Constantin Paleologu[†], *Jacob Benesty*[‡], *and Felix Albu*

[†] University Politehnica of Bucharest, Romania, e-mail: pale@comm.pub.ro [‡] INRS-EMT, University of Quebec, Montreal, Canada, e-mail: benesty@emt.inrs.ca

ABSTRACT

In sparse adaptive filters, the adaptation gain is "proportionately" redistributed among all the coefficients, emphasizing the large ones in order to speed up their convergence. The improved proportionate affine projection algorithm (IPAPA) is a very attractive choice for echo cancellation, since it combines the good convergence features of the affine projection algorithm (APA) and the gain factors of the improved proportionate normalized least-mean-square (IPNLMS) algorithm. Similar to the APA, a matrix inversion is required within the IPAPA. For practical reasons, the matrix needs to be regularized before inversion, i.e., a positive constant is added to the elements of its main diagonal. In this paper, we propose a formula for choosing the regularization parameter of the IPAPA, aiming at attenuating the effects of the noise in the adaptive filter estimate. Simulation results indicate the validity of this approach in both network and acoustic echo cancellation scenarios.

Index Terms— Echo cancellation, adaptive filters, regularization, improved proportionate affine projection algorithm (IPAPA).

1. INTRODUCTION

The basic principle of an echo canceller is to build a model of the echo path impulse response that has to be identified with an adaptive filter, which provides at its output a replica of the echo (that is further subtracted from the reference signal) [1]. The echo paths (for both network and acoustic echo cancellation scenarios) are sparse in nature, i.e., a small percentage of the impulse response components have a significant magnitude while the rest are zero or small. The sparseness character of the echo paths inspired the idea to "proportionate" the algorithm behavior, i.e., to update each coefficient of the filter independently of the others, by adjusting the adaptation stepsize in proportion to the magnitude of the estimated filter coefficient [2]. In this manner, the adaptation gain is "proportionately" redistributed among all the coefficients to emphasize the large ones in order to speed up their convergence and, consequently, to increase the overall convergence rate.

Many interesting proportionate-type algorithms have been proposed in the last decade, e.g., see [3] and the references therein. Among these, the improved proportionate affine projection algorithm (IPAPA) [4] is one of the most attractive choice. This algorithm inherits the convergence features of the affine projection algorithm (APA) [5] and the robustness to the sparseness degree of the echo path specific to the improved proportionate normalized leastmean-square (IPNLMS) algorithm [6]. Similar to the APA, the IPAPA requires a matrix inversion within its update. In echo cancellation, due to the nature of the input signal (which is mainly speech), this matrix can be very ill-conditioned. Consequently, the matrix needs to be regularized before inversion by adding a positive constant to the elements of its main diagonal. In practice, it was found that the value of this regularization term is highly influenced by the level of the system noise (the near-end background noise), i.e., the more the noise, the larger the value of the regularization parameter [1].

In this paper, we propose a formula for choosing the constant regularization parameter of the IPAPA, based on a condition that intuitively makes sense, i.e., to attenuate the effects of the noise in the adaptive filter estimate. Simulations performed in the context of both network and acoustic echo cancellation supports the theoretical findings.

2. REGULARIZATION OF THE IPAPA

In echo cancellation, we basically deal with a system identification problem, having the reference (or desired) signal

$$d(n) = \mathbf{h}^T \mathbf{x}(n) + w(n)$$

= $y(n) + w(n),$ (1)

where n is the discrete-time index,

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

is the impulse response (of length L) of the echo path, superscript T denotes transpose of a vector or a matrix,

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

is a vector containing the most recent L samples of the far-end signal x(n) (i.e., the input signal), w(n) is a zero-mean additive noise signal (i.e., the near-end background noise in the single-talk scenario), which is independent of x(n), and the signal y(n) represents the echo. In order to cancel this echo, the main goal is to estimate or identify **h** with an adaptive filter

$$\hat{\mathbf{h}}(n) = \begin{bmatrix} \hat{h}_0(n) & \hat{h}_1(n) & \cdots & \hat{h}_{L-1}(n) \end{bmatrix}^T, \qquad (2)$$

which updates its coefficients with an adaptive algorithm.

The IPAPA [4] is one of the most popular algorithms used for echo cancellation. It results as a straightforward combination of the APA [5] and the IPNLMS algorithm [6]. The IPAPA is defined by the following equations:

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^{T}(n)\hat{\mathbf{h}}(n-1), \qquad (3)$$

$$\mathbf{R}(n) = \delta \mathbf{I}_P + \mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n), \qquad (4)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \alpha \mathbf{G}(n-1)\mathbf{X}(n)\mathbf{R}^{-1}(n)\mathbf{e}(n), \quad (5)$$

This work was supported under the Grant POS-DRU/89/1.5/S/62557 and Grant CNCS-UEFISCDI PN-II-ID-PCE-2011-3-0097.

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

$$\mathbf{d}(n) = \begin{bmatrix} d(n) & d(n-1) & \cdots & d(n-P+1) \end{bmatrix}^T$$
(6)

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

$$\mathbf{X}(n) = \begin{bmatrix} \mathbf{x}(n) & \mathbf{x}(n-1) & \cdots & \mathbf{x}(n-P+1) \end{bmatrix}$$
(7)

is the input data matrix,

$$\mathbf{G}(n-1) = \operatorname{diag} \begin{bmatrix} g_0(n-1) & g_1(n-1) & \cdots & g_{L-1}(n-1) \end{bmatrix}$$
(8)

is a diagonal matrix containing the proportionate (or gain) factors, $\mathbf{R}(n)$ is the 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 stepsize parameter. The proportionate factors are evaluated as [6]

$$g_l(n-1) = \frac{1-\kappa}{2L} + (1+\kappa) \frac{|\hat{h}_l(n-1)|}{2\sum_{l=0}^{L-1} |\hat{h}_l(n-1)|}, \ 0 \le l \le L-1,$$
⁽⁹⁾

where κ ($-1 \le \kappa < 1$) is a parameter that controls the amount of proportionality. Looking of the equations that define the IPAPA, i.e., (3)–(5), it can be noticed that the classical APA [5] is obtained for $\mathbf{G}(n-1) = \mathbf{I}_L$ (where \mathbf{I}_L is the $L \times L$ identity matrix), while the IPNLMS algorithm [6] results when P = 1.

The regularization matrix $\delta \mathbf{I}_P$ from (4) prevents the problems associated with the inverse of the matrix $\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)$, which can be very ill-conditioned. However, the selection of the regularization parameter δ has a great impact in terms of the adaptive filter performance. If the value of δ is not chosen properly, the IPAPA may never converge, especially under low signal-to-noise ratio (SNR) conditions.

In this context, let us define the echo-to-noise ratio (ENR) [1] as

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

which is also the definition of the SNR, 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. It is known from practice that the value of the regularization parameter δ depends on the level of the noise that corrupts the output of the system that needs to be identified. Low ENRs require high values of the regularization parameter, while its importance becomes less apparent for high ENRs. The main question is how large or small should be chosen the value of δ as a function of the ENR? In order to provide an answer to this question, let us rewrite (5) as

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

where

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

and

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

Examining (11)–(13), we can see that the vector $\tilde{\mathbf{h}}(n)$ is the correction component of the IPAPA, since it depends on the new observation $\mathbf{d}(n)$. Also, it can be noticed that the matrix $\mathbf{P}(n)$ does not

depend on the noise signal or the desired signal, but only on the input signal. The correction term $\tilde{\mathbf{h}}(n)$ is obtained by solving

$$\min_{\tilde{\mathbf{h}}(n)} \left\{ \left[\mathbf{d}(n) - \mathbf{X}^{T}(n)\tilde{\mathbf{h}}(n) \right]^{T} \left[\mathbf{d}(n) - \mathbf{X}^{T}(n)\tilde{\mathbf{h}}(n) \right] + \delta \left\| \tilde{\mathbf{h}}(n) \right\|_{1} \right\}, \quad (14)$$

where $\|\cdot\|_1$ denotes the ℓ_1 norm. In fact, the previous optimization is the regularized version of the minimum ℓ_1 -norm solution of the linear system of P equations $\mathbf{d}(n) = \mathbf{X}^T(n)\tilde{\mathbf{h}}(n)$. Since the solution $\tilde{\mathbf{h}}(n)$ is not the optimal one, the other vector $\mathbf{P}(n)\hat{\mathbf{h}}(n-1)$ in (11) can be seen as a good initialization of the adaptive filter.

Based on the previous considerations, we can define a new error signal vector as

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

A reasonable way to attenuate the effects of the noise in the estimator $\tilde{\mathbf{h}}(n)$ is to find δ by imposing the condition [7]

$$E\left[\|\tilde{\mathbf{e}}(n)\|_{2}^{2}\right] = E\left[\|\mathbf{w}(n)\|_{2}^{2}\right],$$
(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 and $\|\cdot\|_2$ denotes the ℓ_2 norm.

In order to develop (16), we can use (13) in (15) to get

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

Next, let us use the eigenvalue decomposition:

$$\mathbf{X}^{T}(n)\mathbf{G}(n-1)\mathbf{X}(n) = \mathbf{V}(n)\mathbf{\Lambda}(n)\mathbf{V}^{T}(n),$$
(18)

where $\mathbf{V}(n)$ is an orthogonal matrix containing the eigenvectors of $\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)$ as columns and $\mathbf{\Lambda}(n)$ is a diagonal matrix containing the corresponding eigenvalues. Consequently, the inverse of the matrix from (4) is

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

and based on (17)–(19), we get

$$\|\tilde{\mathbf{e}}(n)\|_{2}^{2} = \mathbf{d}^{T}(n)\mathbf{V}(n) \\ \cdot \left\{\mathbf{I}_{P} - \mathbf{\Lambda}(n)\left[\delta\mathbf{I}_{P} + \mathbf{\Lambda}(n)\right]^{-1}\right\}^{2}\mathbf{V}^{T}(n)\mathbf{d}(n).$$
(20)

In order to further process (20), we need to know the eigenvalues of the matrix from the right-hand side of (18). Of course, these parameters depend on the character of the input signal, but also on the proportionate factors. Consequently, based on (9) and assuming that the input signal is white, we get

$$\mathbf{x}^{T}(n)\mathbf{G}(n-1)\mathbf{x}(n) = \frac{1-\kappa}{2L}\mathbf{x}^{T}(n)\mathbf{x}(n)$$

+
$$\frac{1+\kappa}{2\left\|\hat{\mathbf{h}}(n-1)\right\|_{1}}\sum_{l=0}^{L-1}x^{2}(n-l)\left|\hat{h}_{l}(n-1)\right|$$

$$\approx \frac{1-\kappa}{\sigma_{x}^{2}}\sigma_{x}^{2} + \frac{1+\kappa}{2}\sigma_{x}^{2}$$

$$\approx \sigma_{x}^{2},$$
 (21)

where $\sigma_x^2 = E[x^2(n)]$ is the variance of the input signal x(n). Hence, (20) simplifies to

$$\|\tilde{\mathbf{e}}(n)\|_{2}^{2} = \mathbf{d}^{T}(n)\mathbf{d}(n)\left(\frac{\delta}{\delta + \sigma_{x}^{2}}\right)^{2}.$$
 (22)

Therefore, taking the expectation in (22), the condition (16) becomes

$$E\left[\|\mathbf{d}(n)\|_{2}^{2}\right]\left(\frac{\delta}{\delta+\sigma_{x}^{2}}\right)^{2} = E\left[\|\mathbf{w}(n)\|_{2}^{2}\right].$$
 (23)

Since the echo signal and the system noise are uncorrelated, (23) can be rewritten as

$$\left\{ E\left[\|\mathbf{y}(n)\|_{2}^{2} \right] + E\left[\|\mathbf{w}(n)\|_{2}^{2} \right] \right\} \left(\frac{\delta}{\delta + \sigma_{x}^{2}} \right)^{2}$$

$$= E\left[\|\mathbf{w}(n)\|_{2}^{2} \right],$$

$$(24)$$

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

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

which results in the quadratic equation

$$\delta^2 - 2\frac{\sigma_x^2}{\text{ENR}}\delta - \frac{\left(\sigma_x^2\right)^2}{\text{ENR}} = 0.$$
 (26)

The obvious solution of the quadratic equation (26) is

$$\delta = \frac{1 + \sqrt{1 + \text{ENR}}}{\text{ENR}} \sigma_x^2$$
$$= \beta_{\text{IPAPA}} \sigma_x^2, \qquad (27)$$

where

$$\beta_{\rm IPAPA} = \frac{1 + \sqrt{1 + \rm ENR}}{\rm ENR}$$
(28)

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

According to (27), the regularization parameter δ depends on the variance of the input signal (σ_x^2) and the ENR. In both network and acoustic echo cancellation, the first parameter is known, while the ENR could be estimated since it depends on the power of the system noise [8]. Therefore, (28) provides a more rigorous way to choose the normalized regularization parameter as a function of the ENR.

We can notice that the regularization does not depend on the parameter κ . In fact, the regularization of the IPAPA is equivalent to the regularization recently proposed for the APA [7] up to the scaling factor L, which is due to the definition of $g_l(n-1)$ [see (9)]. Also, it can be noticed that the regularization parameter of the IPAPA does not depend on the projection order P and is identical to the regularization parameter of the IPNLMS algorithm when we assume that the input signal is white [9]. In the general case, the expression of the regularization parameter needs not to be accurate; an approximate value gives, usually, good performances.

3. SIMULATION RESULTS

Simulations were performed in the context of echo cancellation, since this is the main application of sparse adaptive filters. Two echo paths were used (see Fig. 1), having different sparseness degree. The first one [Fig. 1(a)] is a network echo path from G168 Recommendation [10]; its impulse response can be considered to be very sparse. The second one [Fig. 1(b)] is a measured acoustic



Fig. 1. Impulse responses used in simulations. (a) Network echo path (G168). (b) Acoustic echo path.

echo path, which is less sparse. Both impulse responses have 512 coefficients, using a sampling rate of 8 kHz. All adaptive filters used in the experiments have the same length, i.e., L = 512.

The far-end signal (i.e., the input signal) is a speech sequence. The output of the echo path is corrupted by an independent white Gaussian noise (i.e., the background noise at the near-end) with different ENRs, i.e., 30 dB, 10 dB, and 0 dB. All the simulations are performed in the absence of the near-end talker (i.e., the single-talk case). In order to evaluate the tracking capabilities of the algorithms, an echo path change scenario is simulated in all the experiments, by shifting the impulse response to the right by 12 samples. The performance measure is the normalized misalignment (in dB), which is defined as $20\log_{10} \left[\left\| \mathbf{h} - \hat{\mathbf{h}}(n) \right\|_2 / \left\| \mathbf{h} \right\|_2 \right]$.

We choose to compare the APA with the IPAPA, both using two types of regularization. The first type is the "classical" choice $\delta = \beta \sigma_x^2$, where β is the normalized regularization parameter. In many simulation scenarios, a frequently used value for this parameter is $\beta = 20$ in the case of APA [1], [3]; equivalent, it was intuitively shown in [6] that this corresponds to $\beta = (1 - \kappa)20/(2L)$ in the case of IPAPA. The second type of regularization is the "optimal" one, which was recently proposed in [7] for the APA [i.e., $\beta_{APA} = L(1 + \sqrt{1 + ENR})/ENR$] and is given in (28) for the IPAPA (i.e., β_{IPAPA}). For all the algorithms, the stepsize parameter is set to $\alpha =$ 0.2 and the projection order is P = 2. The proportionality parameter of IPAPA is chosen as $\kappa = 0$.

Figure 2 presents the misalignment of the algorithms when ENR = 30 dB. First, it is clear that the IPAPA outperforms APA for both echo paths, which is an expectable result; of course, the gain is more apparent for the network impulse response [Fig. 2(a)], which is very sparse. Second, we can notice that the performance obtained with the "classical" regularization [i.e., $\beta = 20$ for the APA and $\beta = (1 - \kappa)20/(2L) \approx 0.02$ for the IPAPA] are very similar with the "optimal" case (i.e., β_{APA} and β_{IPAPA} , respectively). This is also expected, because if we consider L = 512 and ENR = 30 dB, we easily get $\beta_{APA} \approx 16.7$ and $\beta_{IPAPA} \approx 0.03$, which are very close to the "classical" values.

However, the importance of the regularization parameter becomes more apparent in noisy environments. The previous experiment is repeated in Fig. 3 but using ENR = 10 dB. According to



Fig. 2. Misalignment of the APA and IPAPA using different values of the normalized regularization parameter, corresponding to (a) the network echo path from Fig. 1(a) and (b) the acoustic echo path from Fig. 1(b). The echo path changes at time 5 seconds. The input signal is speech, L = 512, P = 2, $\alpha = 0.2$, $\kappa = 0$, and ENR = 30 dB.



Fig. 3. Misalignment of the APA and IPAPA using different values of the normalized regularization parameter, corresponding to (a) the network echo path from Fig. 1(a) and (b) the acoustic echo path from Fig. 1(b). ENR = 10 dB; other conditions as in Fig. 2.

these results, it is clear that the IPAPA using $\beta_{\rm IPAPA}$ outperforms by far the other algorithms. In this case, a much higher value of the normalized regularization constant is required; according to (28), this value is $\beta_{\rm IPAPA} \approx 0.43$, which is much higher as compared to the "classical" choice for IPAPA, i.e., $\beta \approx 0.02$.

Finally, in Fig. 4 the value of the ENR is set to 0 dB. It can be noticed that the regularization process is critical in this case. For an improper value of the normalized regularization constant the misalignment of the adaptive filter fluctuates much and never converges. It is clear that the IPAPA using $\beta_{\rm IPAPA}$ (≈ 2.43 in this case) performs much better in this scenario. The advantage of the proper regularization is also clearly visible in the case of the APA, even if this algorithm converges slower as compared to the IPAPA.



Fig. 4. Misalignment of the APA and IPAPA using different values of the normalized regularization parameter, corresponding to (a) the network echo path from Fig. 1(a) and (b) the acoustic echo path from Fig. 1(b). ENR = 0 dB; other conditions as in Fig. 2.

4. CONCLUSIONS

In this paper, we have proposed a more rigorous way to choose the regularization parameter of the IPAPA as a function of the ENR. The basic condition was to attenuate the effects of the system noise in the adaptive filter estimate. Simulations performed in the context of both network and acoustic echo cancellation prove the validity of this approach in different noisy environments.

5. REFERENCES

- 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.
- [2] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech, Audio Process.*, vol. 8, pp. 508–518, Sept. 2000.
- [3] C. Paleologu, J. Benesty, and S. Ciochină, Sparse Adaptive Filters for Echo Cancellation. Morgan & Claypool Publishers, 2010.
- [4] O. Hoshuyama, R. A. Goubran, and A. Sugiyama, "A generalized proportionate variable step-size algorithm for fast changing acoustic environments," in *Proc. IEEE ICASSP*, 2004, pp. IV-161–IV-164.
- [5] 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.
- [6] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in Proc. IEEE ICASSP, 2002, pp. 1881–1884.
- [7] C. Paleologu, J. Benesty, and S. Ciochină, "Regularization of the affine projection algorithm," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 58, pp. 366–370, June 2011.
- [8] C. Paleologu, J. Benesty, and S. Ciochină, "Practical regularization of the affine projection algorithm," in *Proc. EUSIPCO*, 2011, pp. 704– 708.
- [9] J. Benesty, C. Paleologu, and S. Ciochină, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, pp. 1734–1742, Aug. 2011.
- [10] Digital Network Echo Cancellers, ITU-T Rec. G.168, 2002.