Improving Acoustic Echo Cancellation by Exploiting Prior Knowledge of RIR Energy Decay

Pedro de Carvalho Cayres Pinto, Roberto Esteban Campos Ruiz, Mariane Rembold Petraglia

***Abstract*— Acoustic echo cancellation (AEC) through adaptive filters requires a very large number of coefficients to establish good communication in reverberant environments. In order to increase the convergence rate of the adaptive algorithm, in this paper we explore the exponential decay rate of the room impulse response (RIR) by introducing an additional constraint, based on the energy behavior of sequential blocks of coefficients, into the optimization problem. The performances of the proposed algorithms are investigated in different scenarios and compared with the performance of traditional algorithms employed in the AEC application. We also examine the influence of the algorithms' parameters in the convergence rate and final solution.**

***Keywords*— Adaptive filters; Acoustic echo cancellation; Normalized least-mean-squares algorithm; Affine Projection Algorithm.**

I. INTRODUCTION

Adaptive filters have been applied to the acoustic echo cancellation (AEC) problem in order to identify the echo paths, present in hands-free communication systems. A good solution is generally not easy to obtain for acoustic environments with high reverberation, due to the non-stationary and coloring properties of the voice signals and the long filter length required to model the echo path.

Among the adaptive algorithms used in the AEC application, stand out the normalized least squares (NLMS), the proportionate normalized least mean square (PNLMS), and the affine projection algorithm (APA). The NLMS [1], [2] algorithm presents low computational complexity and robust convergence, but reduced learning rate for colored input signals. The PNLMS [3] algorithm has been proposed for modeling sparse systems, and also has low convergence rate for colored input signals. The APA [2], [4] updates weights based on current and previous input vectors to improve convergence speed for correlated input signals.

In recent years, different mechanisms have been introduced into the adaptive filters to increase their convergence rate and/or reduce their computational complexity, including frequency-domain techniques [5], subband processing [6], decorrelation [7], and block processing [8]. Especially in environments with high reverberation, obtaining a satisfactory solution in a short period of time remains a challenge [9].

In this paper we exploit the exponential behaviour of the room impulse response (RIR) to develop a novel adaptive filtering method for acoustic echo cancelation. We formulate the AEC as a constrained optimization problem and solve it using the Lagrange multipliers method. NLMS algorithm and APA are modified accordingly to generate new algorithms. The Aachen Impulse Response (AIR) [10] database of RIRs is employed in our experiments.

Pedro C. C. Pinto, PEE/COPPE, UFRJ, Rio de Janeiro-RJ, e-mail: pedrocayres@coppe.ufrj.br, Roberto E. C. Ruiz, PEE/COPPE, UFRJ, Rio de Janeiro-RJ, e-mail: roberto@pads.ufrj.br; Mariane R. Petraglia, PEE/COPPE, UFRJ, Rio de Janeiro-RJ, e-mail: mariane@poli.ufrj.br.

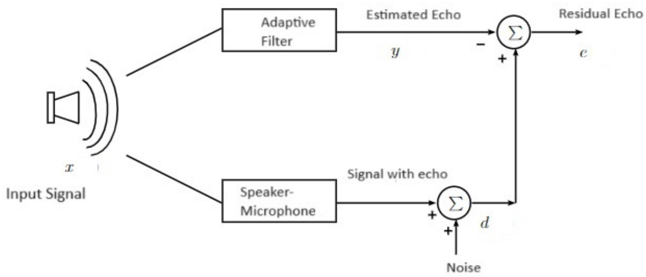

Fig. 1: Acoustic Echo Cancellation.

II. BACKGROUND

The purpose of acoustic echo cancellation is to remove an audio signal that reverberates around the room from another signal of interest. This problem is illustrated in Fig. 1. When there is no local voice signal, the microphone signal $d(k)$ can be modeled as

$$d(k) = x(k) * h(k) + v(k), \tag{1}$$

where $x(k)$ is the clean signal from the speaker, $h(k)$ is the RIR, and $v(k)$ is a random noise. We therefore wish to obtain the coefficients of a finite impulse response (FIR) filter $w(k)$ that approximates the RIR $h(k)$. The NLMS and APA are two iterative algorithms that can be used to obtain the coefficients of the adaptive filter.

Defining $\mathbf{w}(k)$ as the vector containing the N coefficients of the adaptive filter at iteration k, the NLMS formulation as a constraint optimization problem is given by

$$\begin{aligned} \min_{\mathbf{w}(k+1)} \quad & \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\ \text{s.t.} \quad & e_p(k) = (1-\beta)e(k) \\ \text{with} \quad & 0 < \beta < 1, \end{aligned} \tag{2}$$

where

$$\mathbf{x}(k) = [x(k), x(k-1), \ldots, x(k-N+1)]^T \tag{3}$$

is the input vector,

$$e(k) = d(k) - \mathbf{w}(k)^T \mathbf{x}(k) \tag{4}$$

is the *a priori* error, and

$$e_p(k) = d(k) - \mathbf{w}(k+1)^T \mathbf{x}(k) \tag{5}$$

is the *a posteriori* error. The solution of the optimization problem (2) is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\beta e(k)}{\|\mathbf{x}(k)\|^2}\mathbf{x}(k). \tag{6}$$

Likewise, a formulation for APA is given by

$$\begin{aligned} \min_{\mathbf{w}(k+1)} \quad & \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\ \text{s.t.} \quad & \mathbf{e_p}(k) = (1-\beta)\mathbf{e}(k) \\ \text{with} \quad & 0 < \beta < 1, \end{aligned} \tag{7}$$

where

$$\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{X}(k)^T \mathbf{w}(k) \tag{8}$$

is the *a priori* error vector,

$$\mathbf{e_p}(k) = \mathbf{d}(k) - \mathbf{X}(k)^T \mathbf{w}(k+1) \tag{9}$$

is the *a posteriori* error vector obtained with the updated coefficient vector $\mathbf{w}(k+1)$, with

$$\mathbf{d}(k) = [d(k), d(k-1), \ldots, d(k-l+1)]^T \tag{10}$$

and

$$\mathbf{X}(k) = [\mathbf{x}(k), \mathbf{x}(k-1), \ldots, \mathbf{x}(k-l+1)] \tag{11}$$

for l restriction equations. The solution for the optimization problem (7) is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \mathbf{X}(k)[\mathbf{X}(k)^T \mathbf{X}(k)]^{-1}\mathbf{e}(k). \tag{12}$$

III. New Methods

Assuming that an estimate of the RIR is available, we develop variations of NLMS and APA by introducing into the objective function a term related to the ℓ_1-norm of the difference between the energies of each adaptive coefficient block and of the corresponding RIR block. This new term penalizes solutions for the coefficient vector that do not comply with the exponential decay of the RIR. In the case of NLMS, this extension leads to the optimization problem:

$$\begin{aligned} \min_{w(k+1)} \quad & \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \sum_{i=1}^{b} \alpha_i \left| \|\mathbf{w}_{B_i}(k+1)\|^2 - \gamma_i \right| \\ \text{s.t.} \quad & e_p(k) = (1-\beta)e(k) \\ \text{with} \quad & 0 < \beta < 1, \end{aligned} \tag{13}$$

where, for each block i, the vector $\mathbf{w}_{B_i}(k)$ contains n_b coefficients of $\mathbf{w}(k)$ whose indices are in the contiguous list $\{(i-1)n_b + 1, \cdots, in_b\}$, $\gamma_i = \|\mathbf{h}_{B_i}\|^2$ is the corresponding RIR block energy, and α_i is a multiplier parameter. Thus, the vector $\mathbf{w}(k)$ is given by

$$\mathbf{w}(k) = \begin{bmatrix} \mathbf{w}_{B_1}(k)^T & \mathbf{w}_{B_1}(k)^T & \cdots & \mathbf{w}_{B_b}(k)^T \end{bmatrix}^T \tag{14}$$

and has a total of $N = b \cdot n_b$ coefficients.

From the method of Lagrange multipliers, the solution must satisfy, for each block i, the equation:

$$\begin{aligned} &\nabla_{\mathbf{w}_{B_i}(k+1)} J(\mathbf{w}(k+1)) = \\ &2\mathbf{w}_{B_i}(k+1) - 2\mathbf{w}_{B_i}(k) + 2\alpha_i s_i \mathbf{w}_{B_i}(k+1) - \lambda \mathbf{x}_{B_i}(k) = \mathbf{0}, \end{aligned} \tag{15}$$

where $s_i \in \{-1, 1\}$ is the sign defined as

$$s_i = \text{sign}(\|\mathbf{w}_{B_i}(k+1)\|^2 - \gamma_i). \tag{16}$$

From Eq. (15), we obtain

$$(1 + \alpha_i s_i)\mathbf{w}_{B_i}(k+1) = \mathbf{w}_{B_i}(k) + \frac{\lambda}{2}\mathbf{x}_{B_i}(k). \tag{17}$$

Joining the equations from all blocks, we can write:

$$\text{diag}(1 + \alpha \mathbf{s})\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\lambda}{2}\mathbf{x}(k), \tag{18}$$

where $\mathbf{s} \in \{-1, 1\}^{bn_b}$ is the concatenation of the n_b vectors $[s_i, s_i, \ldots, s_i] \in \{-1, 1\}^b$ for all i, and $\alpha \in \mathbb{R}^{bn_b}$ is defined analogously. From the restriction of (13), we obtain

$$\beta e(k) - (\mathbf{w}(k+1) - \mathbf{w}(k))^T \mathbf{x}(k) = 0, \tag{19}$$

which yields, by substituting $\mathbf{w}(k+1)$ from Eq. (18),

$$\frac{\lambda}{2} = \frac{\beta e(k) + \mathbf{w}(k)^T \text{diag}\left(\frac{\alpha \mathbf{s}}{1+\alpha \mathbf{s}}\right)\mathbf{x}(k)}{\|\mathbf{x}(k)\|^2_{\text{diag}\left(\frac{1}{1+\alpha \mathbf{s}}\right)}}. \tag{20}$$

From Eqs. (18) and (20), we obtain the update equation for the new method, which we refer to as NLMS+BEO:

$$\begin{aligned} \mathbf{w}(k+1) &= \text{diag}\left(\frac{1}{1+\alpha \mathbf{s}}\right)\mathbf{w}(k) + \\ &\text{diag}\left(\frac{1}{1+\alpha \mathbf{s}}\right) \frac{\beta e(k) + \mathbf{w}(k)^T \text{diag}\left(\frac{\alpha \mathbf{s}}{1+\alpha \mathbf{s}}\right)\mathbf{x}(k)}{\|\mathbf{x}(k)\|^2_{\text{diag}\left(\frac{1}{1+\alpha \mathbf{s}}\right)}} \mathbf{x}(k). \end{aligned} \tag{21}$$

Similarly, by including the block energy term to APA, we obtain the update equation for the APA+BEO algorithm:

$$\begin{aligned} &\mathbf{w}(k+1) = \mathbf{D}_1 \mathbf{w}(k) + \\ &\mathbf{D}_1 \mathbf{X}(k)[\mathbf{X}(k)^T \mathbf{D}_1 \mathbf{X}(k)]^{-1}[\beta \mathbf{e}(k) + \mathbf{X}(k)^T \mathbf{D}_2 \mathbf{w}(k)], \end{aligned} \tag{22}$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are the diagonal matrices given by

$$\mathbf{D}_1 = \text{diag}\left(\frac{1}{1+\alpha \mathbf{s}}\right) \tag{23}$$

and

$$\mathbf{D}_2 = \text{diag}\left(\frac{\alpha \mathbf{s}}{1+\alpha \mathbf{s}}\right) \tag{24}$$

IV. Simulation Results

Experiments with simulated data were conducted in order to compare the performances of the traditional adaptive algorithms and the proposed new versions. The measured RIR of a lecture room from the Aachen Impulse Response (AIR) database [10], with reverberation time $T_{60} = 0.78$ s and sampling rate $f_s = 16$ kHz, was employed in all experiments. The number of adaptive coefficients was $N = 8000$.

The performances of the adaptive algorithms were evaluated using the Mean Square Deviation (MSD) metric, defined as

$$\text{MSD}(k) = \mathbb{E}[||\mathbf{w}_{opt} - \mathbf{w}(k))||^2], \tag{25}$$

where $\mathbb{E}[.]$ denotes the statistical expectation operator, $\mathbf{w}_{opt}$ is the impulse response of the plant and $||.||$ is the ℓ_2-norm operator. Mean values were approximated by calculating the average over 50 executions of the experiment.

The first three experiments were conducted with white Gaussian noise input signal (3×10^5 samples) with variance $\sigma_x^2 = 1$. In the colored noise experiment, the white noise was passed through a filter with transfer function $H(z) = 1+0.8z^{-1}-0.2z^{-2}$, and in the last experiment the input signal was a male or a female voice recording. In all experiments the additive noise $v(k)$ is a white Gaussian noise with variance $\sigma_v^2 = 10^{-3}$, $\beta = 1$, and the number of restriction equations for APA is $l = 4$. The values $\alpha = 0.001$ and $n_b = 100$ were used in all simulations, except when these parameters were varied to observe their influence on the performance of the proposed algorithms. The values of γ_i were computed exactly from the RIR. In practice, their values would be obtained from the blind estimate of T_{60} [11], [12].

A. White Noise Input

The MSD results of all four methods for white noise input are shown in Fig. 2. The modified algorithms present better results, converging faster and to lower MSD values. The results for APA and NLMS are almost identical, which is expected, since there is no correlation between input samples from different time instants.

B. Block Size Analysis

The MSD results of the NLMS+BEO algorithm for different block sizes are shown in Fig. 3. From this figure, it can be seen that the performance of the algorithm is not sensitive to the number of blocks. Smaller block sizes produce slightly better results.

It is worth mentioning that this result was obtained with exact values of the energies of the RIR blocks. It is expected that when estimated values for the γ_i parameters are used, the optimal block size will depend on the accuracy of the estimation method. In practice, a blind estimate of the reverberation time [12] would be used, the accuracy of which is influenced by the characteristics of the room (such as size and number of people inside it) and the quality of the audio equipment, among other factors. For scenarios with low accuracy estimates of T_{60}, it would be recommended to use a reduced number of blocks.

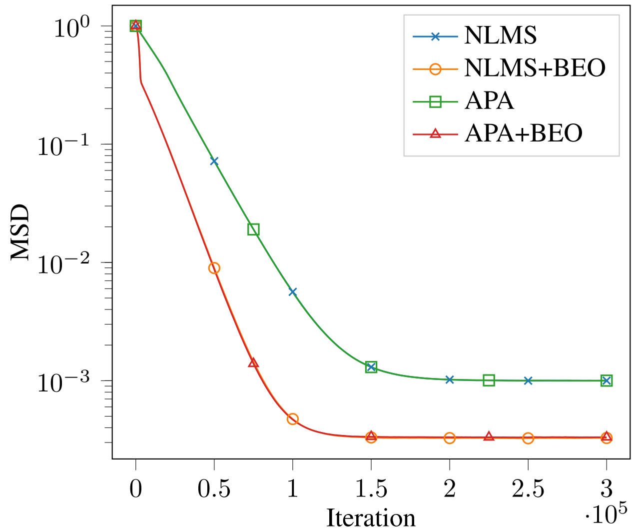

Fig. 2: MSD evolution for all methods with white noise input.

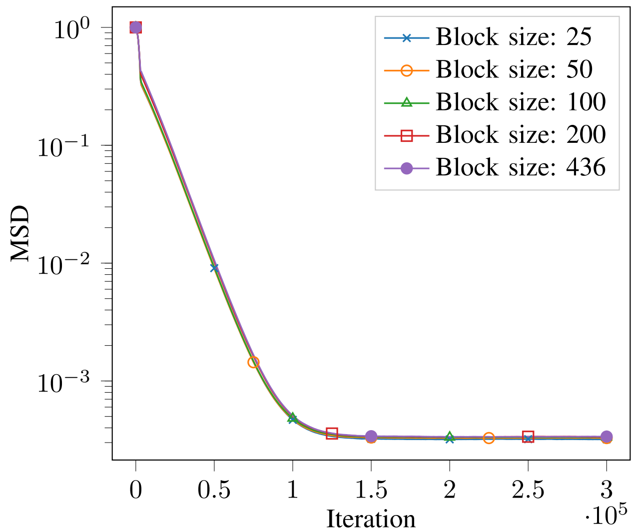

Fig. 3: Analysis of block size influence on MSD evolution for NLMS+BEO with white noise input.

C. Alpha Values Analysis

The MSD results of the NLMS+BEO algorithm for different values of α are shown in Fig. 4. The best α value evaluated for the simulated RIR was 0.001. From the curves, it can be seen that the proposed algorithm is quite insensitive to the choice of α over a wide range (from 0.01 to 0.0001). Additional investigations, including theoretical analyses, are necessary to obtain the best value of the α parameter for other RIRs.

Again, it is expected that when estimated values for the γ_i parameters are used, the optimal value of the α parameter will be influenced by the estimation error. If a very accurate estimate is not likely, the weight given to the additive term should be less than when a good estimate is obtained.

D. Colored Noise Input

The MSD results for all four methods with colored noise input are shown in Fig. 5. The modified algorithms exhibit

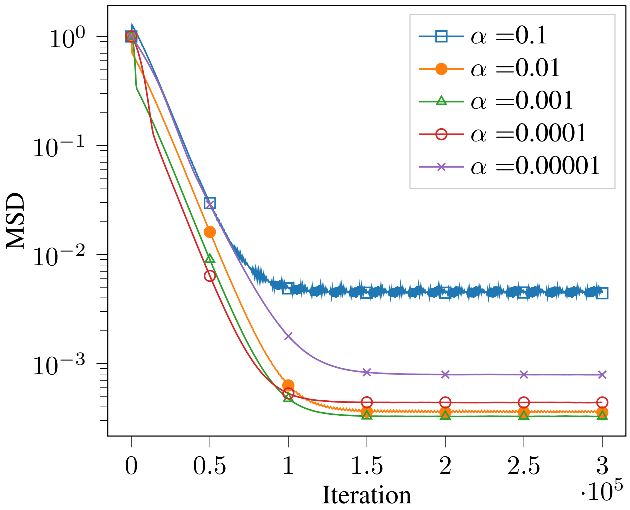

Fig. 4: Analysis of α value influence on MSD evolution for NLMS+BEO with white noise input.

better results, converging faster than their counterparts.

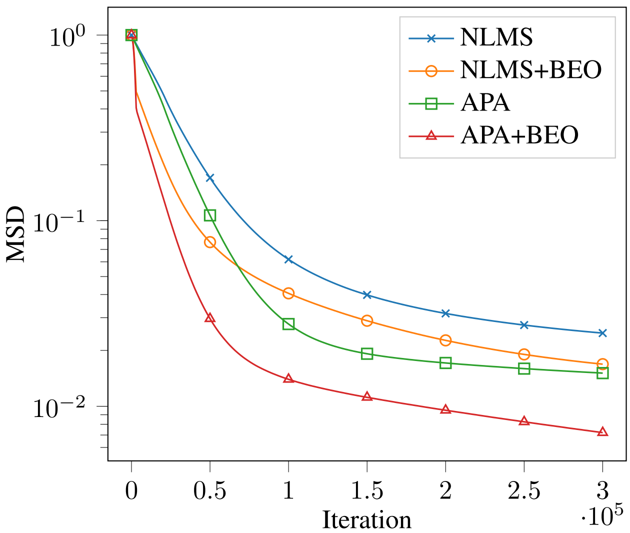

Fig. 5: MSD evolution for all methods with colored noise input.

E. Speech Signal Input

Figures 6 and 7 show the MSD results of all methods for female and male voice signals, respectively. The APA+BEO approach presented the best results for both voice signals.

It can be observed that the weights of the modified methods deviate from the optimal solution in the first iterations, which indicates that the initialization of the new algorithms still need some adjustments.

V. Conclusions

In this paper we develop variations of two adaptive algorithms for acoustic echo cancellation applications by taking into consideration some available estimate of the exponential energy decay of room impulse responses. Experiments were

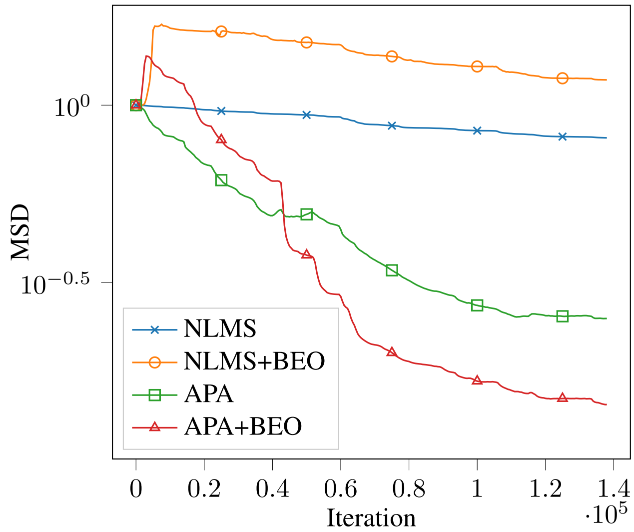

Fig. 6: MSD evolution for all methods with a female voice signal.

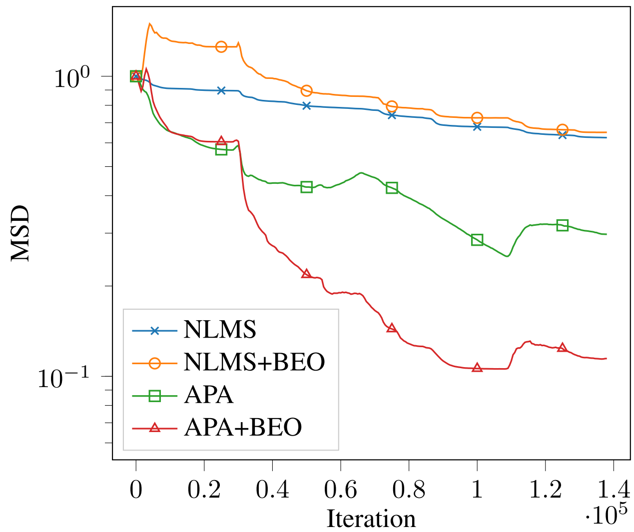

Fig. 7: MSD evolution for all methods with a male voice signal.

conducted employing a measured impulse response of a lecture room, using white Gaussian noise, colored noise and speech as input signals. Results demonstrated a faster convergence of the proposed algorithms compared to the classical NLMS and APA. The proposed NLMS+BEO algorithm is faster than NLMS in simulated experiments with stationary input signals. The proposed APA+BEO is faster than APA in all cases. Future work includes theoretical and experimental analyzes to determine the optimal value ranges for the parameters α and n_b, taking into account inaccuracies in the estimates of the values of T_{60} and γ_i.

Acknowledgements

This work was partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and by Fundação Carlos

Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

REFERENCES

[1] J.-I. Nagumo and A. Noda, "A learning method for system identification," *IEEE Transactions on Automatic Control*, vol. 12, no. 3, pp. 282–287, June 1967.

[2] A. H. Sayed, *Adaptive filters*. John Wiley & Sons, 2011.

[3] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 5, pp. 508–518, Sep. 2000.

[4] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan (Part I: Communications)*, vol. 67, no. 5, pp. 19–27, May 1984.

[5] F. Yang, G. Enzner, and J. Yang, "On the convergence behavior of partitioned-block frequency-domain adaptive filters," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4906–4920, 2021.

[6] M. R. Petraglia, D. B. Haddad, and E. L. Marques, "Affine projection subband adaptive filter with low computational complexity," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 10, pp. 989–993, 2016.

[7] S. Zhang, J. Zhang, and H. C. So, "Low-complexity decorrelation nlms algorithms: Performance analysis and aec application," *IEEE Transactions on Signal Processing*, vol. 68, pp. 6621–6632, 2020.

[8] S. Zhang, H. Chen, and A. H. Sayed, "Design of a tap-amplitude-based block proportional adaptive filtering algorithm," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2448–2459, 2023.

[9] T. van Waterschoot, G. Rombouts, and M. Moonen, "Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement," *Signal Processing*, vol. 88, no. 3, pp. 594–611, 2008.

[10] M. Jeub, M. Schafer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *2009 16th International Conference on Digital Signal Processing*, July, Santorini, Greece, 2009.

[11] J. Eaton, N. D. Gaubitch, and P. A. Naylor, "Noise-robust reverberation time estimation using spectral decay distributions with reduced computational cost," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 161–165.

[12] J. Eaton, N. D. Gaubitch, A. H. Moore, and P. A. Naylor, "Estimation of room acoustic parameters: The ace challenge," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 10, pp. 1681–1693, 2016.