IMPROVING ROBUSTNESS OF BLIND ADAPTIVE MULTICHANNEL IDENTIFICATION ALGORITHMS USING CONSTRAINTS

Md. Kamrul Hasan[1], *Jacob Benesty*[2], *Patrick A. Naylor*[1], *and Darren B. Ward*[1]

[1]Department of Electrical and Electronic Engineering, Imperial College London, UK
[2]INRS-EMT, University of Quebec, Montreal, Canada
e-mail: {mk.hasan,p.naylor,d.ward}@imperial.ac.uk, benesty@emt.inrs.ca

ABSTRACT

This paper shows that the robustness of the normalized multichannel frequency-domain LMS algorithm reported in [1] can be improved using constraints in the adaptation rule. In the identification of acoustic impulse responses with leading bulk zeros from noisy observations the proposed constraint shows significant performance improvement in terms of normalized projection misalignment. Experimental results for various simulated conditions are presented to justify our claim.

1. INTRODUCTION

Blind channel identification is a common issue in diverse fields of science and engineering. Signals transmitted from the source are adversely affected by the propagating medium/channel. The channel identification, therefore, is required to remove its detrimental effect from the received signal often by inversion. In communications, the problem is to equalize the channel effect on the received signal to obtain the transmitted signal [2]. In geophysics, the reflectivity of the earth layers is explored by extracting seismic wavelets from the sensor signals [3]. In speech processing, particularly in acoustic dereverberation, the problem is to separate the sound source from the received microphone signals [4].

Both single and multichannel identification schemes are reported in the literature by many researchers. Multichannel identification schemes, however, are increasingly becoming popular due to their suitability in removing the unknown channel effects more effectively than their single channel counterparts. Among the various techniques reported so far, e.g. least-squares approach [5], subspace method [6], [7], maximum-likelihood method [8], Newton algorithm [9], the LMS algorithm [9] is simple and efficient. Among all of its variants, it has been shown in [1] that the normalized multichannel frequency domain LMS (NMCFLMS) algorithm is more computationally efficient and effective for identifying long acoustic channels which are of particular interest for dereverberation. The convergence characteristic in Fig. 1, however, show that the NMCFLMS algorithm lacks robustness to additive noise (e.g., characteristic for 20 dB and 30 dB). With the increase in noise level, the algorithm has been found more and more prone to misconvergence. This characteristic of the NMCFLMS algorithm cannot be improved by lowering the value of μ, though the point of divergence may be delayed as shown for example in Fig. 1 comparing the characteristics for 20 and 30 dBs.

In this paper, we investigate how such a scheme is affected by additive noise and delay in the impulse response such as occurring for acoustic channels with leading bulk zeros.

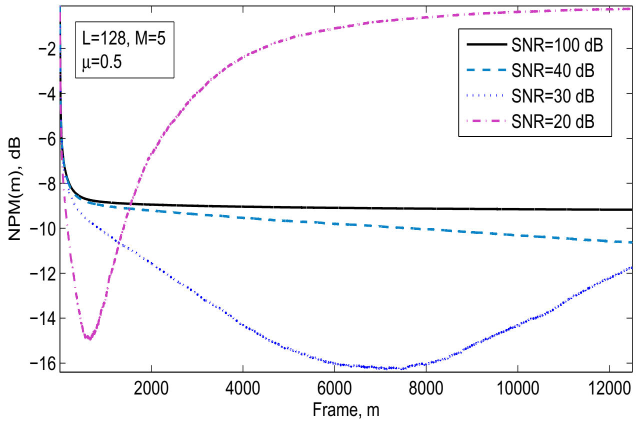

Figure 1: Variation of NPM with iteration at different SNRs for $\mu = 0.5$. The array consists of $M = 5$ microphones, source-mic separation $d = 1m$, and the length of the channel impulse response is $L = 128$.

The main objective of this paper is to show that the convergence rate as well as robustness of the NMCLMS algorithm to the identification of such AIRs in presence of noise can be improved significantly using certain constraints in the update rule based on the prior knowledge of one or more *significant coefficients* of the channel transfer function.

2. PROBLEM FORMULATION

Consider a speech signal recorded inside a non anechoic room using a linear array of microphones. The received signals at the microphones can be modeled as convolutional mixtures of the speech signal and the impulse responses of the acoustic paths between source and microphones. The channel outputs and observed signals are then given by

$$y_i(n) = s(n) * h_i(n) = \sum_{k=0}^{L-1} h_{i,k}(n)s(n-k) \quad (1)$$

$$x_i(n) = y_i(n) + v_i(n), \quad i = 1,2,\cdots,M \quad (2)$$

where M is the number of microphones, $s(n)$, $y_i(n)$, $x_i(n)$, $v_i(n)$ and $h_{i,k}(n)$ denote, respectively, the clean speech, reverberant speech, the reverberant speech corrupted by background noise, observation noise, and impulse response of the source to ith microphone. It is assumed that the additive noise on M channels is uncorrelated white random sequence, i.e., $E\{v_i(t)v_j(t)\} = 0$ for $i \neq j$ and $E\{v_i(t)v_i(t-t')\} = 0$ for $t \neq t'$. It is also assumed that $v_i(n)$ are uncorrelated with $s(n)$.

A blind channel identification algorithm estimates $\mathbf{h}_i$, $i = 1, 2, \cdots, M$, given by

$$\mathbf{h}_i = [h_{i,0}\ h_{i,1} \cdots h_{i,L-1}]^T \quad (3)$$

solely from the observations $x_i(n)$, $n = 1, 2, \cdots, N$. The identifiability conditions commonly stated are:

1. The channel transfer functions $H_i(z)$ don't contain any common zeros.
2. The autocorrelation matrix of the source signal, $R_{ss} = E\{\mathbf{s}(n)\mathbf{s}^T(n)\}$, is of full rank.

In this paper, we examine the robustness of the NMCFLMS algorithm reported in [1] to the blind identification of time-invariant $\mathbf{h}_i$ from the noise corrupted sequence $x_i(n)$ while $h_{i,k_i} \cong 0, \quad k_i = 0, 1, \cdots, K_i$, for the case with $K_i < L-1$. This situation might arise for impulse responses recorded inside a noise-free reverberant room. Thus for the problem of dereverberation taking care of this problem is of significant importance.

3. NMCFLMS ALGORITHM WITH CONSTRAINT

From (1), we deduce the following relationship:

$$\begin{aligned} y_i(n) * h_{j,k} - y_j(n) * h_{i,k} &= s(n) * [h_{i,k} * h_{j,k} \\ &\quad - h_{j,k} * h_{i,k}] = 0 \end{aligned} \quad (4)$$

However, in presence of noise an error function may be defined as

$$\begin{aligned} e_{ij}(n) &= x_i(n) * \widehat{h}_{j,k} - x_j(n) * \widehat{h}_{i,k} \\ &= [s(n) * h_{i,k} + v_i(n)] * \widehat{h}_{j,k} \\ &\quad - [s(n) * h_{j,k} + v_j(n)] * \widehat{h}_{i,k} \end{aligned} \quad (5)$$

The NMCFLMS algorithm reported in [1] is summarized below:

$$\begin{aligned} \widehat{\underline{\mathbf{h}}}_k^{10}(m+1) &= \widehat{\underline{\mathbf{h}}}_k^{10}(m) - \mu[\mathbf{p}_k(m) + \delta \mathbf{I}_{2L\times 2L}]^{-1} \\ &\quad \times \sum_{i=1}^{M} D_{x_i}^*(m)\underline{\mathbf{e}}_{ik}^{01}(m), \quad k = 1, 2, \cdots, M \end{aligned} \quad (6)$$

where

$$\widehat{\underline{\mathbf{h}}}_k^{10}(m) = \mathbf{F}_{2L\times 2L} \begin{bmatrix} \widehat{\mathbf{h}}_k(m) \\ \mathbf{0} \end{bmatrix} \quad (7)$$

$$\underline{\mathbf{e}}_{ik}^{01}(m) = \mathbf{F}_{2L\times 2L} \begin{bmatrix} \mathbf{0} \\ \mathbf{F}_{L\times L}^{-1}\underline{\mathbf{e}}_{ik}(m) \end{bmatrix} \quad (8)$$

$$\begin{aligned} \mathbf{p}_k(m) &= \lambda \mathbf{p}_k(m-1) + (1-\lambda) \sum_{i=1, i\neq k}^{M} D_{x_i}^*(m) D_{x_i}(m), \\ & k = 1, 2, \cdots, M. \end{aligned} \quad (9)$$

Here m is the frame index and $\mathbf{F}$ denotes the discrete Fourier transform (DFT) matrix. The frequency-domain error function $\underline{\mathbf{e}}_{ik}(m)$ is given by

$$\underline{\mathbf{e}}_{ik}(m) = D_{x_i}(m)\widehat{\underline{\mathbf{h}}}_k(m) - D_{x_k}(m)\widehat{\underline{\mathbf{h}}}_i(m) \quad (10)$$

The diagonal matrix $D_{x_i}(m)$ is the DFT of the mth frame data block for the ith channel, i.e.,

$$\begin{aligned} D_{x_i}(m) &= diag\,(\mathbf{F}\{\mathbf{x}_i(m)_{2L\times 1}\}) \\ \mathbf{x}_i(m)_{2L\times 1} &= [x_i(mL-L)\ \ x_i(mL-L+1) \\ &\quad \cdots x_i(mL+L-1)]^T \end{aligned} \quad (11)$$

and the estimate of the kth channel coefficient vector is defined as

$$\widehat{\mathbf{h}}_k(m) = [\widehat{h}_{k,0}(m)\ \ \widehat{h}_{k,1}(m) \cdots \widehat{h}_{k,dp}(m) \cdots \widehat{h}_{k,L-1}(m)]^T \quad (12)$$

where $\widehat{h}_{k,dp}(m)$ denotes the estimate of the direct path component.

From (4), an error function related only to the AIRs can be defined as

$$\tilde{e}_{ij}(k) = [h_{i,k} * \widehat{h}_{j,k} - h_{j,k} * \widehat{h}_{i,k}] \quad (13)$$

As can be deduced from (4), the NMCFLMS algorithm essentially minimizes $\tilde{e}_{ij}(k)$ for a spectrally flat input signal. Taking the z-transform of (13), we obtain

$$\tilde{E}_{ij}(z) = [H_i(z)\widehat{H}_j(z) - H_j(z)\widehat{H}_i(z)] \quad (14)$$

AIRs recorded in noise-free rooms are usually headed with zeros due to the direct path propagation delay. If we assume that there are d_i and d_j zeros at the head of $h_{i,k}$ and $h_{j,k}$, respectively, and the estimates $\widehat{h}_{i,k}$ and $\widehat{h}_{j,k}$ have the same structure as their respective true values, then (14) may be rewritten as

$$\tilde{E}_{ij}(z) = z^{-(d_i+d_j)}[H_i'(z)\widehat{H}_j'(z) - H_j'(z)\widehat{H}_i'(z)] \quad (15)$$

where $H'(z)$ denotes a polynomial with no leading zeros. Now if we argue that the estimates $\widehat{h}_{i,k}$ and $\widehat{h}_{j,k}$ have structure with no leading zeros except the relative delay, for $d_j > d_i$ we obtain

$$\tilde{E}_{ij}(z) = z^{-d_j}[H_i'(z)\widehat{H}_j''(z) - H_j'(z)\widehat{H}_i''(z)] \quad (16)$$

where $\widehat{H}''(z)$ denotes a polynomial with no leading zeros. Comparing (15) and (16), it can be inferred that the NMCFLMS algorithm may converge with unknown likelihood to either of the two solutions differing by an amount of delay d_i unless subjected to some constraints to obtain the former solution.

In matrix-vector notation, (13) can be expressed as

$$\tilde{\mathbf{e}}_{ij} = [\ -\mathbf{H}_j \quad \mathbf{H}_i\ ] \begin{bmatrix} \widehat{\mathbf{h}}_i \\ \widehat{\mathbf{h}}_j \end{bmatrix} \quad (17)$$

where $\mathbf{H}_i$ and $\mathbf{H}_j$ are the convolution matrices formed from $h_i(n)$ and $h_j(n)$, respectively. The channel matrix $\mathbf{H}_{ij}$ defined as

$$\mathbf{H}_{ij} = [\ -\mathbf{H}_j \quad \mathbf{H}_i\ ] \quad (18)$$

plays, particularly in noise-free case, a critical role on the convergence of the adaptive algorithm. The eigenvalue spread of this matrix may be used to quantify the condition of the problem. However, comparing (4) and (5) it may be argued that the presence of additive noise would prevent formation of the error function as (17). Thus the presence of noise of an appropriate level improves the condition of the identification problem.

We now consider modifying the adaptive algorithm such that the estimated direct path coefficients is constrained to match the true direct path coefficients in terms of both delay and magnitude. The aim of this constraint is to improve robustness. The constraint is of interest in practical applications since we can assume the existence of robust estimation of the direct path using such algorithms as [10]. At the mth iteration we substitute

$$\begin{aligned}\widehat{\widehat{\mathbf{h}}}_k(m) &= [\widehat{h}_{k,0}(m) \;\; \widehat{h}_{k,1}(m)\cdots h_{k,dp}\cdots\widehat{h}_{k,L-1}(m)]^T \\ &= \widehat{\mathbf{h}}_k(m)+\Delta\widehat{\mathbf{h}}_k(m) \qquad (19)\end{aligned}$$

where

$$\Delta\widehat{\mathbf{h}}_k(m) \;=\; [0 \;\; 0 \;\cdots\; h_{k,dp}-\widehat{h}_{k,dp}(m) \;\cdots\; 0 \;\; 0]^T \quad (20)$$

The parameter update equation in the frequency-domain, (7), for the proposed case can be written as

$$\begin{aligned}\widehat{\underline{\mathbf{h}}}_k^{10}(m+1) = \widehat{\widehat{\underline{\mathbf{h}}}}_k^{10}(m) &- \mu[\mathbf{p}_k(m)+\delta\mathbf{I}_{2L\times 2L}]^{-1} \qquad (21)\\ &\times\sum_{i=1}^{M} D_{x_i}^*(m)\underline{\mathbf{e}}_{ik}^{01}(m)-\mu[\mathbf{p}_k(m)+\delta\mathbf{I}_{2L\times 2L}]^{-1}\\ &\times\sum_{i=1}^{M} D_{x_i}^*(m)\Delta\underline{\mathbf{e}}_{ik}^{01}(m)\end{aligned}$$

where $\Delta\underline{\mathbf{e}}_{ik}^{01}(m) = D_{x_i}(m)\Delta\widehat{\widehat{\underline{\mathbf{h}}}}_k(m) - D_{x_k}(m)\Delta\widehat{\widehat{\underline{\mathbf{h}}}}_i(m)$. Thus the proposed substitution is equivalent to providing the adaptive algorithm in each iteration a better initial value for update of parameters in the next iteration. It is better in the sense that a possibly erroneous estimate resulting particularly due to ill conditioned channel matrix and/or noise is replaced by its true value. The effect of this initialization also propagates in all other directions of the transformed parameter vector which essentially changes all gradient directions. The proposed selective substitution in turn has the effect of giving rise to perturbations to the estimate of all unknown coefficients.

4. SIMULATION RESULTS

In this section, we present computer simulation results to investigate the effectiveness of the proposed substitution to blind channel estimation problems. The dimension of the room was taken to be $(5\times 4\times 3)$ m. A linear array consisting of $M=5$ microphones with uniform separation of $\tau=0.2$ m was used in the experiment. The first microphone and source were positioned at $(1.0, 1.5, 1.6)$ m and $(2.0, 1.2, 1.6)$ m, respectively. The positions of the other microphones can be obtained by adding $\tau=0.2$ m successively with the y-coordinate of the first microphone. The impulse responses were generated using the image model reported in [11] for reverberation time $T_{60}=0.1$ s and then truncated so as to make the length $L=128$. In all cases, the source signal was Gaussian white noise, and λ was fixed to $[1-1/(3L)]^L$.

The performance index used for measurement of improvement is the modified normalized projection misalignment defined as

$$\mathrm{NPM}(m) = 20\log_{10}\left(\left[\underline{\mathbf{h}} - \frac{\underline{\mathbf{h}}^T\widehat{\underline{\mathbf{h}}}(m)}{\widehat{\underline{\mathbf{h}}}^T(m)\widehat{\underline{\mathbf{h}}}(m)}\widehat{\underline{\mathbf{h}}}(m)\right]/||\underline{\mathbf{h}}||\right) \quad (22)$$

where $||\cdot||$ is the l_2 norm, $\underline{\mathbf{h}} = [\mathbf{h}_{l+1:L-1}^T \quad \mathbf{0}_{1\times(l+1)}]^T$, $\widehat{\underline{\mathbf{h}}} = [\widehat{\mathbf{h}}_{l+1:L-1}^T \quad \mathbf{0}_{1\times(l+1)}]^T$ and l denotes the position of the direct path coefficient. Note that only the coefficients that appear after the direct path component are considered. However, the channel vector length is made equal to L by inserting the required number of zeros at the tail. This modification is made for fair comparison with the conventional NMCFLMS algorithm as the NPM(m) could be better in the proposed case due to the use of the true direct path component if it is not excluded in the computation. The coefficients $\mathbf{h}_{0:l-1}$ prior to the direct path component are neglected as they are essentially zero.

The proposed constraint is applied to the NMCFLMS algorithm from the second iteration, i.e. (19) is activated for $m>1$. This allows both the original and constrained NMCFLMS algorithms to start from the same point of NPM in Fig. 3. The positions and amplitudes of the direct path components are assumed to be known *a priori*. In practical cases, as these quantities are unknown, they can be estimated using the reported robust techniques for time difference of arrival (TDOA) [10].

The results of channel estimation using the conventional and constrained NMCFLMS algorithms for SNR=20 dB are presented in Fig. 2 where Fig. 2 (a) is the true impulse response, and Figs. 2 (b) and (c) are the estimated impulse responses. The parameter μ was set different in the two algorithms so as to make them reach almost the same value of final NPM. As can be seen from Figs. 2 (a) and (b), there exists some delay between the true and estimated impulse responses while using the conventional NMCFLMS algorithm. As demonstrated by (15) and (16), this delay is equal to the minimum delay among all the direct path components of the AIRs. It is also known that the LMS type blind identification algorithms can only estimate coefficients up to a scaling factor. The results in Fig. 2 (a) and (c) show that there exists no delay between the true and estimated impulse responses due to the proposed constraint on the direct path coefficients. Comparative results on the convergence rate of the NMCFLMS algorithm with and without the proposed substitution are depicted in Fig. 3 for SNR=20 dB. As shown, the convergence of the conventional NMCFLMS algorithm under noisy condition is more dependent on the choice of the μ parameter. It is also clear that the convergence rate as well as the asymptotic performance are significantly better in the proposed case for the same number of iterations.

In Fig. 4, the effect of the proposed substitution is demonstrated for the case when the source is moving away from the microphone array. The amplitudes of the direct-path components decrease as the source-microphones separation increases. Therefore, it is expected that the impact of the proposed substitution will be less as the separation increases. Indeed, this can be observed from the asymptotic NPM values obtained for each source position and SNR=20 dB. To investigate the effect of inaccuracy in amplitude estimation, simulations are performed introducing uncertainty into these amplitude estimates. It can be observed that results with 30 dB error are very similar to those of the case with no error, whilst for 20 dB the performance is noticeably degraded. We have not considered fractional delays for substitution as this will be undertaken as future work.

5. CONCLUSIONS

In this paper, we have investigated the performance of the NMCFLMS algorithm in the identification of AIRs with constraints on the direct path coefficients. The results of our experiment have demonstrated that the presence of additive

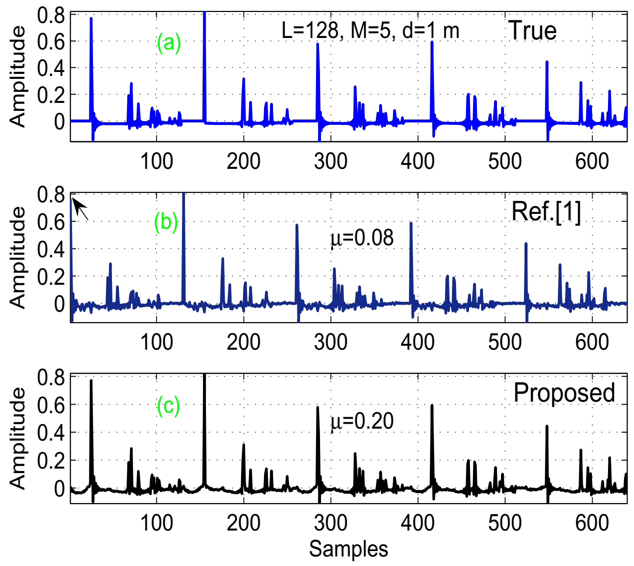

Figure 2: Results on channel estimation at SNR=20 dB using the NMCFLMS.

noise beyond a certain limit leads to the misconvergence of the conventional NMCFLMS algorithm. It has been also shown in our tests that the correct convergence of the adaptive algorithm can be restored with improved estimation accuracy when constraints on the direct path coefficients are imposed at an SNR of 20 dB. This improvement is maintained even when the direct path constraint contains some inaccuracy.

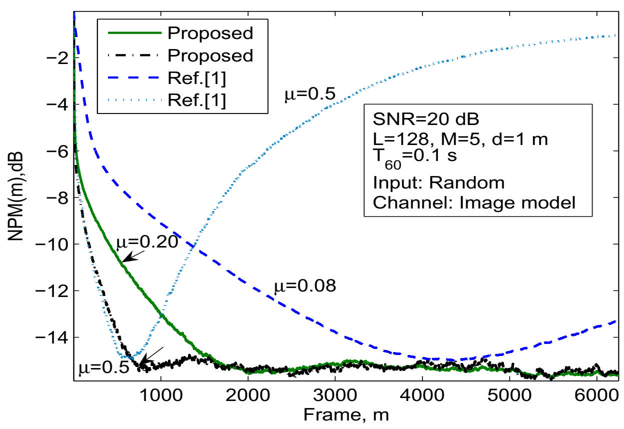

Figure 3: Variation of NPM with iteration at SNR=20 dB. The array consists of $M = 5$ microphones, source-mic separation $d = 1m$, and the length of the channel impulse response is $L = 128$.

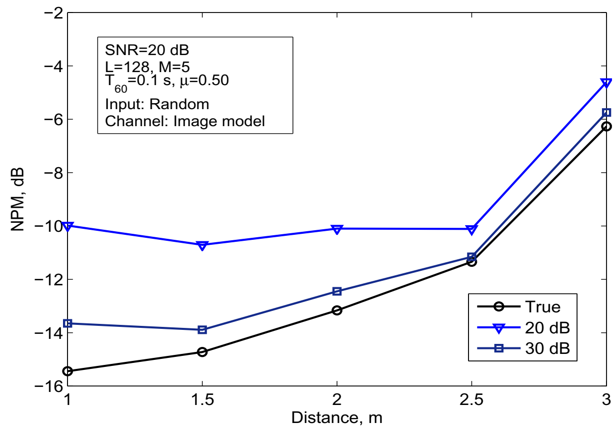

Figure 4: Final value of NPM as the source moves away from the microphone array at SNR=20 dB. The array consists of $M = 5$ microphones and the length of the channel impulse response is $L = 128$.

REFERENCES

[1] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Speech Audio Processing*, vol. 51, no. 1, pp. 11–24, Jan. 2003.

[2] Z. Xu and M. K. Tsatsanis, "Blind channel estimation for long code multiuser cdma systems," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1919–1930, Apr. 2000.

[3] K. F. Kaaresen and T. Taxt, "Multichannel blind deconvolution of seismic signals," *GEOPHYSICS*, vol. 63, no. 6, pp. 2093–2107, Nov. 1998.

[4] S. Subramaniam, A. P. Petropulu, and C. Wendt, "Cepstrum-based deconvolution for speech dereverberation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 4, no. 5, pp. 392–396, Sept. 1996.

[5] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[6] H. H. Zeng and L. Tong, "Blind channel estimation using the second-order statistics: algorithms," *IEEE Trans. Signal Processing*, vol. 45, no. 8, pp. 1919–1930, Aug. 1997.

[7] S. Gannot and M. Moonen, "Subspace methods for multimircophone speech dereverberation," *EURASIP Journal Applied Signal Processing*, no. 11, pp. 1074–1090, 2003.

[8] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 661–672, Mar. 1996.

[9] Y. Huang and J. Benesty, "Adaptive multi-channel least mean square and newton algorithms for blind channel identification," *Signal Process.*, vol. 82, no. 8, pp. 1127–1138, Aug. 2002.

[10] J. Benesty, T. Gansler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Springer, 2001.

[11] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.