

A NOISE-ROBUST DUAL FILTER APPROACH TO MULTICHANNEL BLIND SYSTEM IDENTIFICATION

Rehan Ahmad, Nikolay D. Gaubitch and Patrick A. Naylor

Department of Electrical and Electronic Engineering, Imperial College London
Exhibition Road, London, SW7 2AZ, United Kingdom

phone: + (44) 20759 46235, fax: + (44) 20759 46302, email: {rehan.ahmad02, nikolay.gaubitch, p.naylor}@imperial.ac.uk
web: www.commsp.ee.ic.ac.uk/sap

ABSTRACT

We present a novel frequency domain blind multichannel identification approach. This is a dual filter based approach comprising a background and a foreground filter. The foreground filter tracks the changes in the cost function of the background filter and employs an update decision criterion making it robust to observation noise and facilitating tracking of changes in the unknown system. Simulation results for both speech and white Gaussian noise input signals are presented to illustrate the algorithm's performance for acoustic systems.

1. INTRODUCTION

Blind channel identification (BCI) for single-input multiple-output (SIMO) systems is an important technique with a variety of potential applications in acoustics, wireless communications and other signal processing systems. Techniques for BCI based upon second order statistics [1] [2] and higher order statistics [3] have been studied. In addition, multichannel BCI techniques are becoming feasible for applications. The normalized multichannel frequency domain LMS (NM-CFLMS) algorithm [4] has been shown to be effective in identifying room impulse responses which are of particular interest in acoustic dereverberation. However, NMCFLMS lacks robustness to additive observation noise and can suffer misconvergence even in moderate noise conditions. This has been studied in [5] [6].

We propose a novel approach to blind multichannel identification which is robust to the presence of additive noise in the system. Dual-filter methods have been used successfully to achieve double-talk robustness in echo cancellation systems [7]. The dual-filter (DF) approach for multichannel BCI involves implementation of a dual-filter structure comprising a background and a foreground filter. A measure is defined on the basis of the cost function of the adaptive algorithm in the background filter. The value of the measure is decided on the basis of background filter adaptive estimation algorithm that in turn determines when to copy the estimates from the background filter into the foreground filter. Simulations with white gaussian noise (WGN) and speech signal input display the noise robustness of the DF approach over the NMCFLMS algorithm for multichannel BCI.

Consider a speech signal recorded inside a non-anechoic room using an array of microphones. The microphone signals represent the convolution of the speech signal and the impulse responses of the acoustic paths between source and microphones. With reference to Fig. 1 and defining $s(n)$ and $v_i(n)$ as the source signal and background noise respectively,

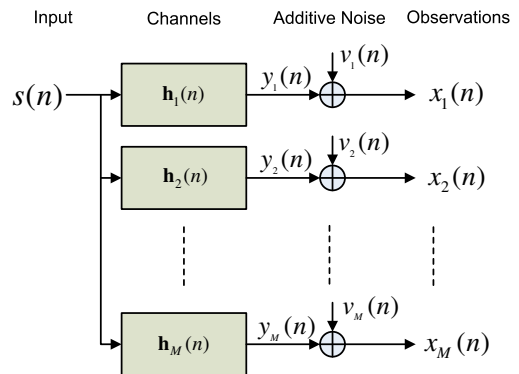


Figure 1: Relationship between input and output in a SIMO model.

the i^{th} channel output signal $x_i(n)$ is given by

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

$$y_i(n) = \mathbf{h}_i^T(n)\mathbf{s}(n), \quad (1b)$$

where M is the number of channels, $\mathbf{h}_i(n) = [h_{i,0}(n) \quad h_{i,1}(n) \quad \dots \quad h_{i,L-1}(n)]^T$, $\mathbf{s}(n) = [s(n) \quad s(n-1) \quad \dots \quad s(n-L+1)]^T$, $\mathbf{h}_i(n)$ is the i^{th} channel impulse response, L is the length of the longest channel impulse response and the superscript T denotes vector transposition. Defining $E\{\cdot\}$ as the expectation operator, we assume that the additive noise on M channels is uncorrelated, i.e., $E\{v_i(n)v_j(n)\} = 0$ for $i \neq j$ and $E\{v_i(n)v_i(n-n')\} = 0$ for $n \neq n'$ while additive noise $v_i(n)$ is uncorrelated with the input signal $s(n)$. For channel identifiability [4], we also assume that

1. The channel transfer functions do not contain any common zeros and
2. The autocorrelation matrix of the source signal, $\mathbf{R}_{ss} = E\{\mathbf{s}(n)\mathbf{s}^T(n)\}$, is of full rank.

A blind multichannel system can be identified adaptively by minimizing the cross-relation error given, for $i \neq j$, by

$$e_{ij}(n) = \mathbf{x}_i^T(n)\hat{\mathbf{h}}_j(n-1) - \mathbf{x}_j^T(n)\hat{\mathbf{h}}_i(n-1), \quad i, j = 1, \dots, M, \quad (2)$$

where $\hat{\mathbf{h}}_i(n)$ is the estimated i^{th} channel impulse response. Using (2), BCI algorithms such as NMCFLMS are derived by minimizing the cost function $J(n) = \sum_{i=1}^{M-1} \sum_{j=i+1}^M e_{ij}^2(n)$ with respect to the estimated impulse response $\hat{\mathbf{h}}_i(n)$ for $i = 1, \dots, M$. The NMCFLMS [4] algorithm is a frame-based frequency-domain BCI algorithm given, for each m^{th}

frame, by:

$$\hat{\mathbf{h}}_i^{10}(m) = \hat{\mathbf{h}}_i^{10}(m-1) - \rho [\mathcal{P}_i(m) + \sigma \mathbf{I}_{2L \times 2L}]^{-1} \times \sum_{j=1}^M \mathcal{D}_{x_j}^*(m) \underline{\mathcal{E}}_{ji}^{01}(m), \quad i = 1, \dots, M, \quad (3)$$

$$\mathcal{P}_i(m) = \lambda \mathcal{P}_i(m-1) + (1-\lambda) \sum_{j=1, j \neq i}^M \mathcal{D}_{x_j}^*(m) \mathcal{D}_{x_j}(m), \quad (4)$$

$$\underline{\mathcal{E}}_{ij}^{01}(m) = \mathbf{F}_{2L} \mathbf{W}_{2L \times 2L}^{01} \mathbf{e}_{ij}(m), \quad (5)$$

$$\mathbf{e}_{ij}(m) = [e_{ij}(mL-L) \dots e_{ij}(mL+L-1)]^T, \quad (6)$$

where * denotes complex conjugate, ρ is the step-size, λ is the forgetting factor and σ is the regularization constant. Defining $\mathbf{I}_{L \times L}$, $\mathbf{O}_{L \times L}$, and \mathbf{F}_L as the identity, null and Fourier matrices of dimension $L \times L$ respectively,

$$\mathbf{W}_{2L \times 2L}^{10} = [\mathbf{I}_{L \times L} \quad \mathbf{O}_{L \times L}]^T, \quad \mathbf{W}_{2L \times 2L}^{01} = \begin{bmatrix} \mathbf{O}_{L \times L} & \mathbf{O}_{L \times L} \\ \mathbf{O}_{L \times L} & \mathbf{I}_{L \times L} \end{bmatrix},$$

$$\hat{\mathbf{h}}_i(m) = \mathbf{F}_L \hat{\mathbf{h}}_i(m), \quad \hat{\mathbf{h}}_i^{10}(m) = \mathbf{F}_{2L} \begin{bmatrix} \hat{\mathbf{h}}_i(m) \\ \mathbf{O}_{L \times 1} \end{bmatrix} \text{ and } \mathcal{D}_{x_j}(m) = \text{diag}\{\mathbf{F}_{2L}[x_j(mL-L) \dots x_j(mL+L-1)]^T\}.$$

The accuracy of system identification can be quantified using the normalized projection misalignment (NPM) [8] given, for frame m , by

$$\text{NPM}(m) = 20 \log_{10} \left(\left\| \mathbf{h} - \frac{\mathbf{h}^T \hat{\mathbf{h}}(m)}{\hat{\mathbf{h}}^T(m) \hat{\mathbf{h}}(m)} \hat{\mathbf{h}}(m) \right\| / \|\mathbf{h}\| \right) \text{ dB}, \quad (7)$$

where $\|\cdot\|$ is the l_2 norm and $\hat{\mathbf{h}}(m) \stackrel{\text{def}}{=} [\hat{\mathbf{h}}_1^T(m) \hat{\mathbf{h}}_2^T(m) \dots \hat{\mathbf{h}}_M^T(m)]^T$.

2. NMCFLMS UNDER NOISY CONDITIONS

The NMCFLMS algorithm lacks robustness to additive noise [5]. It can be seen from Fig. 2 that the NMCFLMS algorithm misconverges after achieving NPM of -20 dB, -23 dB, -25 dB, and -30 dB for corresponding signal-to-noise ratios (SNR) 25 dB, 30 dB, 35 dB, and 40 dB, respectively for a constant adaptation gain of $\mu = 0.6$. Hence, the problem of misconvergence in the NMCFLMS algorithms [4] increases with noise level.

Using (1a) and (1b), we can write

$$\mathbf{x}_i(n) = \mathbf{y}_i(n) + \mathbf{v}_i(n), \quad (8)$$

where $\mathbf{y}_i(n) = [y_i(n) \ y_i(n-1) \ \dots \ y_i(n-L+1)]^T$, and $\mathbf{v}_i(n) = [v_i(n) \ v_i(n-1) \ \dots \ v_i(n-L+1)]^T$.

Using (2) and (8), the error to be minimized can then be divided into a noise-free component e_{ij}^y and a noise-only component e_{ij}^v as

$$\begin{aligned} e_{ij}(n) &= [\mathbf{y}_i^T(n) \hat{\mathbf{h}}_j(n-1) - \mathbf{y}_j^T(n) \hat{\mathbf{h}}_i(n-1)] \\ &\quad + [\mathbf{v}_i^T(n) \hat{\mathbf{h}}_j(n-1) - \mathbf{v}_j^T(n) \hat{\mathbf{h}}_i(n-1)] \\ &= e_{ij}^y(n) + e_{ij}^v(n). \end{aligned} \quad (9)$$

In this noisy case, the cost function is described by

$$\begin{aligned} J(n) &= \frac{1}{\|\hat{\mathbf{h}}(n)\|^2} \sum_{i=1}^{M-1} \sum_{j=i+1}^M \left\{ [e_{ij}^y(n)]^2 + [e_{ij}^v(n)]^2 \right\} \\ &= J_y(n) + J_v(n). \end{aligned} \quad (10)$$

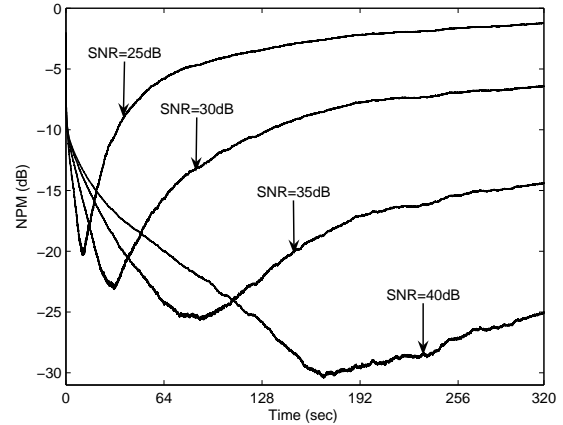


Figure 2: Effect of noise on normalized projection misalignment for the NMCFLMS algorithm with $\mu = 0.6$

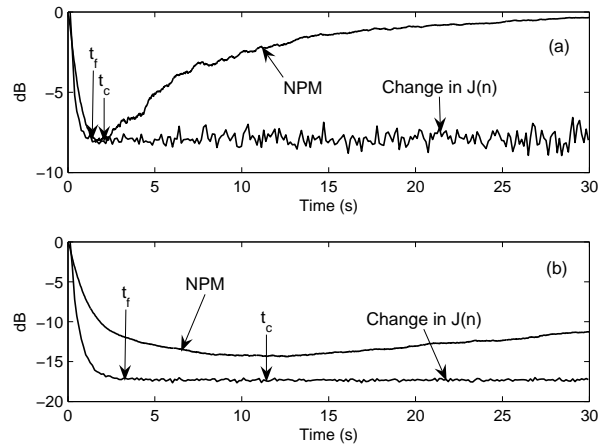


Figure 3: NPM and Cost function gradient plots at (a) SNR = 10dB (b) 20dB, t_c indicates the critical point, t_f indicates the flattening point.

Hence the cost function $J(n)$ is a sum of a noise-free element $J_y(n)$ and a noise-only element $J_v(n)$. It can be deduced that the minimization of the cost function $J(n)$ under noisy conditions does not necessarily minimize the noise-free cost function element $J_y(n)$, which is desired so as to identify the unknown system correctly. This is illustrated for two different SNR conditions in Fig. 3, showing that minimization of $J(n)$ does not result in correct estimates of the unknown system in the presence of observation noise.

3. THE DUAL FILTER APPROACH FOR MULTICHANNEL BCI

We now develop the DF approach for multichannel BCI. The approach comprises two sets of filters; background filters $\mathbf{B}(z)$ and foreground filters $\mathbf{F}(z)$ as shown in the block diagram in Fig. 4. The $\mathbf{B}(z)$ are adapted by the NMCFLMS algorithm. $\mathbf{B}(z)$ estimates are copied into $\mathbf{F}(z)$ at different times during the adaptation following an update decision criterion.

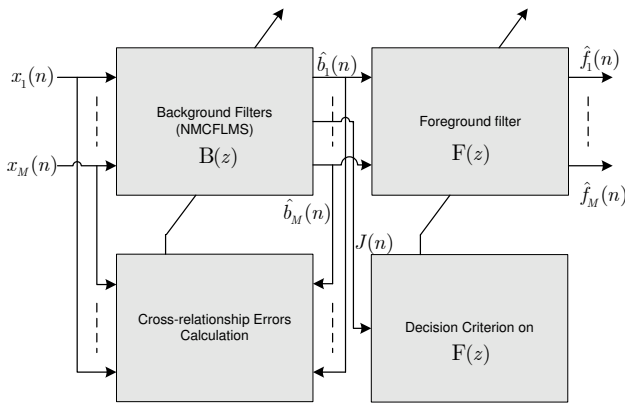


Figure 4: Block diagram of the multichannel DF approach for the BCI.

The characteristics of the convergence of $\mathbf{B}(z)$ for a noisy system are illustrated in Fig. 3 and exhibit a critical point. We define the critical point as the instant in time when the NPM is minimum; this being after initial convergence and before misconvergence. The critical point is denoted by t_c in Fig. 3.

3.1 Estimation of the critical point

The critical point of $\mathbf{B}(z)$ in the DF approach is determined by smoothing the cost function $J(n)$ using a 1024 tap rectangular window moving average filter. Smoothed $J(n)$ initially decreases then flattens to a certain value which is specific to the SNR of the system. Denoting the initial value of the smoothed $J(n)$ in dB as J_i and its value when it flattens in dB as J_f , we define

$$\delta_J = |J_f - J_i| \text{ dB}. \quad (11)$$

The value of δ_J has been found to be proportional to the SNR of the system but independent of the number of channels, M , as well as the length, L , of the system and the specific value of $\mathbf{h}(n)$ which is shown in the Fig. 5. The flattening time, t_f , is the instant in time when $J(n)$ decreases to J_f as shown in Fig. 3. The critical point t_c is found empirically from the flattening time t_f using the following relationship

$$t_c = t_f \times 2^{\frac{\delta_J}{10}}. \quad (12)$$

3.2 The Dual filter Approach

The DF approach aims to avoid misconvergence in the presence of observation noise. The procedure is to copy the coefficients of $\mathbf{B}(z)$ to $\mathbf{F}(z)$ according to a decision criterion. γ is used as a flag to implement the decision criterion. $\mathbf{B}(z)$ is initialized after every time interval T_{BG} and adapts using the NMCFLMS algorithm during the interval. Interaction between $\mathbf{B}(z)$ and $\mathbf{F}(z)$ is controlled by the value of γ as shown in Fig. 6 with γ initialized to 1. When $\gamma = 1$, $\mathbf{B}(z)$ estimates are copied into $\mathbf{F}(z)$ after each iteration until $\mathbf{B}(z)$ reach the critical point t_c then estimates are not copied. At the end of T_{BG} interval, γ is set to 0. In the next T_{BG} interval $\gamma = 0$ so $\mathbf{B}(z)$ estimates are not copied after each iteration, instead just once after reaching the critical point t_c . During $\mathbf{B}(z)$ adaptation $J(n)$ is analyzed at each iteration for a change equal

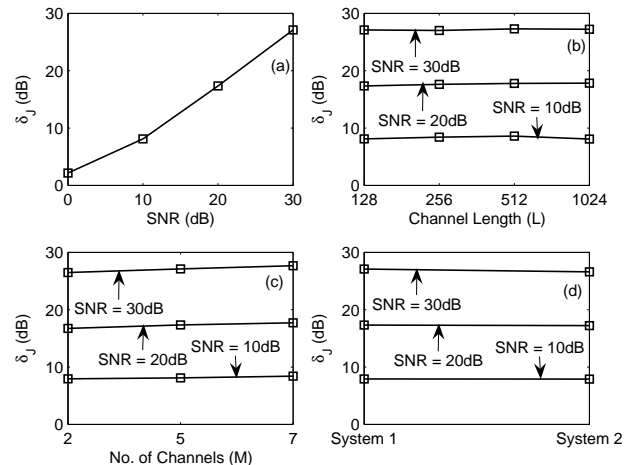


Figure 5: Relationship of δ_J with (a) SNR (b) Length of the system (c) Number of channels (d) Two different systems.

to δ_J after its flattening. This δ_J change in $J(n)$ reflects a change in the system. In this case, γ is reset to 1 and the new value of γ is used in the following T_{BG} . Hence the DF approach always keeps the best estimates in $\mathbf{F}(z)$, avoiding the misconvergence which is otherwise observed in multichannel BCI in the presence of observation noise. Fast update of $\mathbf{F}(z)$ could be achieved by adapting multiple instances of $\mathbf{B}(z)$ with a small T_{BG} . Hence $\mathbf{F}(z)$ is more frequently updated but with a price of higher computation.

4. SIMULATIONS AND RESULTS

We now present simulation results to compare the performance of the DF approach against the NMCFLMS algorithm [4] in the context of acoustic room impulse response identification. The dimensions of the room are $(5 \times 4 \times 3)$ m and impulse responses are generated using the method of images [9] with reverberation time $T_{60} = 0.1$ s which are then truncated to length $L = 128$. A linear microphone array containing $M = 5$ microphones with uniform separation $d = 0.2$ m is used. The source and the first microphone are placed at $(1.0, 1.5, 1.6)$ m and $(2.0, 1.2, 1.6)$ m, respectively and a source at $(1.0, 3.0, 1.6)$ m is used to simulate change in the system. Input signals are white Gaussian noise (WGN) and a male speech while uncorrelated zero-mean WGN is added to achieve the SNR specified for each experiment. The sampling frequency is 8 kHz and the SNR is 20 dB unless otherwise specified. Defining $\mathbf{h} = [\mathbf{h}_1^T \mathbf{h}_2^T \dots \mathbf{h}_M^T]^T$, the SNR for this BCI application is given [4] as $\text{SNR} \triangleq 10 \log_{10}[\sigma_s^2 \|\mathbf{h}\|^2 / (M\sigma_b^2)]$ where σ_s^2 and σ_b^2 are the signal and noise powers, respectively, while the following parameters are chosen for all simulations: $\gamma = 1$ at the start, $\lambda = [1 - 1/(3L)]^N$, $\hat{\mathbf{h}}_i(0) = [1 \ 0 \dots 0]^T / \sqrt{M}$ for $\mathbf{B}(z)$ filters.

Figure 7 shows the comparison of convergence of the DF approach with the NMCFLMS algorithm using a WGN input signal at an SNR = 20 dB. The step-size for both algorithms is $\mu = 0.6$. T_{BG} for $\mathbf{B}(z)$ of the DF approach is 20 s. In Fig. 7(a) NMCFLMS clearly misconverges after initial convergence upto t_c . The NPM of $\mathbf{B}(z)$ is shown with a dotted

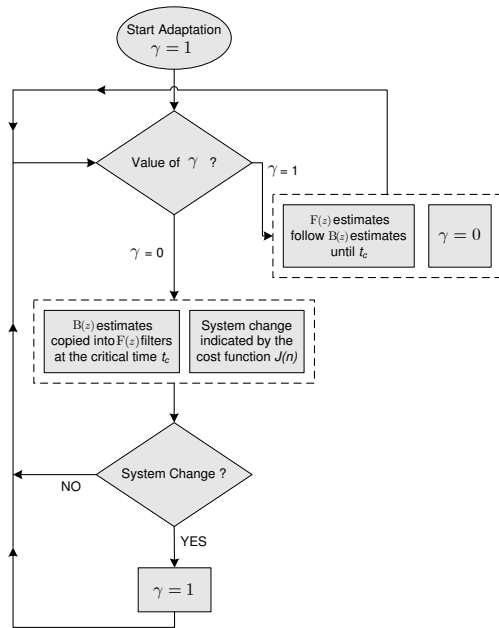


Figure 6: Flow Diagram of the DF approach for multichannel BCI.

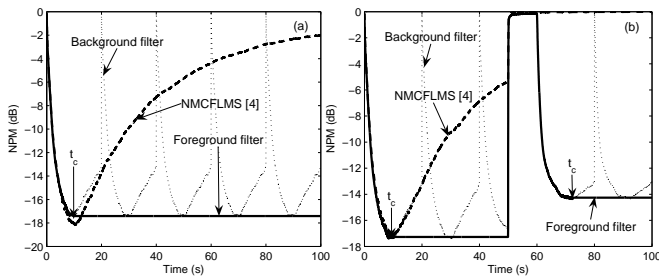


Figure 7: Comparison of convergence of NMCFLMS, $B(z)$ and $F(z)$ filters with WGN input for (a) Single system (b) Step change in the system.

line and misconverge within the T_{BG} interval. $B(z)$ are re-initialized at the end of every T_{BG} interval. $F(z)$ initially converge with $B(z)$ but after t_c they keep the best estimates of the unknown channels following the update decision criterion hence do not misconverge. Fig. 7(b) shows the comparison of convergence with a step change in system after 50 s. $F(z)$ has the best estimate of the first system for 50 s but has wrong estimate after system change until the following T_{BG} starts. Then it has the best estimate of second system copied from $B(z)$ following the update decision criterion.

Figure 8 shows the comparison of convergence of the DF approach with the NMCFLMS algorithm with speech signal input at an SNR = 40 dB. The step-size for the algorithms is $\mu = 0.08$ and T_{BG} for $B(z)$ is 32.88 s. Fig. 8(a) shows the DF approach does not misconverge while NMCFLMS does and Fig. 8(b) shows the DF approach tracks changes in the system successfully.

5. CONCLUSION

The NMCFLMS is an effective algorithm for adaptive blind system identification but suffers from misconvergence in the

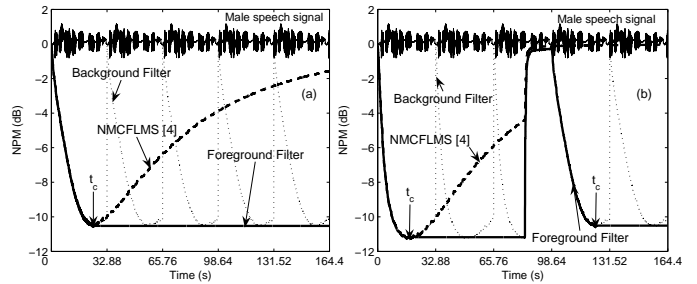


Figure 8: Comparison of convergence of NMCFLMS, $B(z)$ and $F(z)$ filters with speech input for (a) Single system (b) Step change in the system.

presence of noise. To alleviate this problem, we proposed a dual filter approach. An expression, which is proportional to the SNR but invariant to channel coefficient values, channel length and number of channels was given, relating the smoothed cost function of an adaptive background filter with the critical point of misconvergence. We then employed this relation in a decision criterion for the update of a foreground filter. In this way, misconvergence is avoided and system variations can be tracked. Finally, simulation results demonstrated the performance improvement over the standard NMCFLMS using the proposed approach.

REFERENCES

- [1] 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.
- [2] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.
- [3] G. Giannakis and J. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 7360–7377, Mar. 1989.
- [4] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Signal Processing*, vol. 51, no. 1, pp. 11–24, Jan. 2003.
- [5] M. K. Hasan, J. Benesty, P. A. Naylor, and D. B. Ward, "Improving robustness of blind adaptive multichannel identification algorithms using constraints," in *Proc. 13th European Signal Processing Conf.*, 2005.
- [6] R. Ahmad, A. W. H. Khong, M. K. Hasan, and P. A. Naylor, "The extended normalized multichannel FLMS algorithm for blind channel identification," in *Proc. 14th European Signal Processing Conf.*, 2006.
- [7] E. J. Diethorn, "Improved decision logic for two-path echo cancellers," in *Proc. Int. Workshop Acoust. Echo Noise Control*, 2001.
- [8] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Processing Lett.*, vol. 5, no. 7, pp. 174–176, July 1998.
- [9] 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.