AN IMPROVED PROPORTIONATE MULTI-DELAY BLOCK ADAPTIVE FILTER FOR PACKET-SWITCHED NETWORK ECHO CANCELLATION

Andy W.H. Khong[1]*, Jacob Benesty*[2] *and Patrick A. Naylor*[1]

[1]Imperial College London
Exhibition Road, London SW7 2AZ, UK
[2]Université du Québec, INRS-EMT
800 de la Gauchetière Ouest, Suite 6900
Montréal, Québec, H5A 1K6, Canada
email: {andy.khong, p.naylor}@imperial.ac.uk, benesty@emt.inrs.ca

ABSTRACT

We present an adaptive echo cancellation algorithm for sparse echo path impulse responses. This new approach exploits both the robustness of the improved proportionate normalized least-mean-square (IPNLMS) algorithm and the efficient implementation of the multi-delay (MDF) adaptive filtering algorithm inheriting the beneficial properties of both. Evaluation results are presented and the computational complexity is also discussed. Both speech and white Gaussian noise simulation results show that the IPMDF algorithm outperforms the MDF and IPNLMS algorithms for both sparse and dispersive echo path impulse responses.

1. INTRODUCTION

With the advent of IP voice telephony, research on network echo cancellation is increasingly important. In such systems where traditional telephony equipment is connected to the packet switched network, the echo path impulse response, which is typically of length 64-128 ms, exhibits an 'active' region in the range of 8-12 ms duration. As a result, the impulse response is dominated by regions where magnitudes are close to zero making the impulse response sparse. The 'inactive' region is due to the presence of bulk delay caused by network propagation, encoding and jitter buffer delays [1].

Classical adaptive algorithms such as the normalized least-mean-square (NLMS) algorithm have slow convergence in sparse network echo cancellation applications. The proportionate normalized least-mean-square algorithm (PNLMS) [2] was proposed which exploits the sparseness of the network impulse response such that each filter coefficient is updated independently of the others by a variable step-size proportional to the estimated filter coefficient. Subsequent improved versions such as the IPNLMS [3] and IIPNLMS [4] were proposed which achieve improved convergence by introducing a controlled mixture of proportionate (PNLMS) and non-proportionate (NLMS) adaptation. Consequently, these algorithms perform better than PNLMS for sparse, and in some cases for dispersive, impulse responses.

In recent years, frequency-domain adaptive algorithms have become popular due to their efficient implementation. These algorithms incorporate block updating strategies whereby the fast-Fourier transform (FFT) algorithm is used together with the overlap-save method [5]. One of the main drawbacks of these approaches is the delay introduced between the input and output which is equivalent to the length of the adaptive filter. Consequently, for long impulse responses, this delay can be considerable since the number of filter coefficients can be several thousands. To mitigate this problem, the authors in [6] proposed the multi-delay filter (MDF) which uses a block length N independent of the filter length L. Although it has been well-known, from the computational complexity point of view, that $N = L$ is the optimal choice, the MDF algorithm nevertheless is more efficient than time-domain implementation even for $N < L$.

In this paper, we propose the improved proportionate multi-delay filtering (IPMDF) algorithm for sparse impulse responses which exploits both the improvement in convergence brought about by the proportionality control of the IPNLMS algorithm and efficient implementation using the MDF structure. This paper is organized as follows: Section 2 reviews briefly the PNLMS, IPNLMS and MDF algorithms. We derive the proposed IPMDF algorithm in Section 3 while Section 4 presents the computational complexity. Section 5 shows simulation results and Section 6 concludes our work.

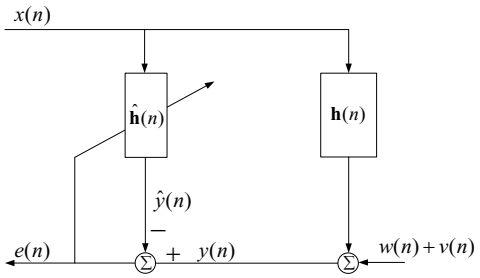

Figure 1: Schematic diagram of an echo canceller.

2. REVIEW OF IPNLMS AND MDF ALGORITHMS

With reference to Fig. 1, we first define filter coefficients and tap-input vectors as $\hat{\mathbf{h}}(n) = [\hat{h}_0(n)\ \hat{h}_1(n) \ldots \hat{h}_{L-1}(n)]^T$ and $\mathbf{x}(n) = [x_0(n)\ x_1(n) \ldots x_{L-1}(n)]^T$ where L is the adaptive filter length. The adaptive filter will model the unknown impulse response $\mathbf{h}(n)$ using signal $y(n) = \mathbf{x}^T(n)\mathbf{h} + v(n) + w(n)$ where $v(n)$ and $w(n)$ are defined as the near-end speech signal and ambient noise respectively. For simplicity, we shall temporary ignore the effects of double talk and ambient noise, $v(n) = w(n) = 0$, in the description of algorithms.

2.1 PNLMS and IPNLMS algorithms

The PNLMS and IPNLMS have been proposed for echo cancellation of sparse systems. These algorithms are described by the following set of equations:

$$e(n) = y(n) - \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n) \quad (1)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mathbf{Q}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{Q}(n-1)\mathbf{x}(n) + } \quad (2)$$

$$\mathbf{Q}(n-1) = \text{diag}\{q_0(n-1)\ q_1(n-1)\dots q_{L-1}(n-1)\} \quad (3)$$

where is the adaptive step-size and is the regularization parameter. The diagonal control matrix $\mathbf{Q}(n)$ determines the step-size of each filter coefficient and is dependent on the specific algorithm as described below.

2.1.1 PNLMS

The elements of the control matrix $\mathbf{Q}(n)$ for PNLMS are [2]

$$q_l(n) = \frac{_l(n)}{\sum_{i=0}^{L-1} \ _i(n)}, \ 0 \le l \le L-1 \quad (4)$$

$$_l(n) = \max\{ \ \times \max\{ \ , |\hat{h}_0(n)| \dots |\hat{h}_{L-1}(n)|\}, |\hat{h}_l(n)|\}. \quad (5)$$

The parameter , with a typical value of 0.01, prevents $\hat{h}_l(n)$ from stalling during initialization stage where $\hat{\mathbf{h}}(0) = \mathbf{0}_{L\times 1}$ while prevents coefficients from stalling when they are much smaller than the largest coefficient. The regularization parameter in (2) for PNLMS should be taken as $_{\text{PNLMS}} = \ _{\text{NLMS}}/L$ [3] where $_{\text{NLMS}} = \ _x^2$ is the variance of the input signal. It can be seen that for ≥ 1, PNLMS is equivalent to NLMS.

2.1.2 IPNLMS

An enhancement of PNLMS is the IPNLMS algorithm [3] which is a combination of PNLMS and NLMS with the relative significance of each controlled by a factor . The elements of the control matrix $\mathbf{Q}(n)$ for IPNLMS are given by

$$q_l(n) = \frac{1-}{2L} + (1+ \ )\frac{|\hat{h}_l(n)|}{2\|\hat{\mathbf{h}}\|_1 + } \quad (6)$$

where is a small value and $\|.\|_1$ is the l_1-norm operator. The regularization parameter in (2) for IPNLMS should be taken as $_{\text{IPNLMS}} = \frac{1-}{2L} \ _{\text{NLMS}}$ [3]. It can be seen that IPNLMS is equivalent to NLMS when $= -1$ while for close to 1, IPNLMS behaves like PNLMS.

2.2 The MDF algorithm

We now summarize the MDF algorithm [6] by first assuming $L = KN$ where N is the frame-size and $K \in \mathbb{Z}^+$. Letting m be the frame index, we have the following quantities:

$$\mathbf{X}(m) = [\mathbf{x}(mN)\dots\mathbf{x}(mN+N-1)] \quad (7)$$

$$\mathbf{y}(m) = [y(mN)\dots y(mN+N-1]^T \quad (8)$$

$$\hat{\mathbf{y}}(m) = [\hat{y}(mN)\dots\hat{y}(mN+N-1)]^T = \mathbf{X}^T(m)\hat{\mathbf{h}} \quad (9)$$

$$\mathbf{e}(m) = \mathbf{y}(m) - \hat{\mathbf{y}}(m) = [e(mN)\dots e(mN+N-1)]^T. \quad (10)$$

We note that $\mathbf{X}(m)$ is a Toeplitz matrix of dimension $L \times N$. Defining k as the block-index and $\mathbf{T}(m-k)$ as an $N \times N$ Toeplitz matrix such that

$$\mathbf{T}(m-k) = \begin{bmatrix} x(mN-kN) & \cdots & x(mN-kN-N+1) \\ x(mN-kN+1) & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ x(mN-kN+N-1) & \cdots & x(mN-kN) \end{bmatrix}, \quad (11)$$

it can be shown using (9) and (11) that

$$\hat{\mathbf{y}}(m) = \sum_{k=0}^{K-1} \mathbf{T}(m-k)\hat{\mathbf{h}}_k(m) \quad (12)$$

where

$$\hat{\mathbf{h}}_k(m) = [\hat{h}_{kN}(m) \ \ \hat{h}_{kN+1}(m)\dots\hat{h}_{kN+N-1}(m)]^T, \ \ k = 0,1\dots K-1 \quad (13)$$

is the k^{th} sub-filter of $\hat{\mathbf{h}}(m)$.

It can be shown that the Toeplitz matrix $\mathbf{T}(m-k)$ can be transformed, by doubling its size, to a circulant matrix

$$\mathbf{C}(m-k) = \begin{bmatrix} \mathbf{T}'(m-k) & \mathbf{T}(m-k) \\ \mathbf{T}(m-k) & \mathbf{T}'(m-k) \end{bmatrix} \quad (14)$$

with

$$\mathbf{T}'(m-k) = \begin{bmatrix} x(mN-kN+N) & \cdots & x(mN-kN+1) \\ x(mN-kN-N+1) & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ x(mN-kN-1) & \cdots & x(mN-kN+N) \end{bmatrix}. \quad (15)$$

The resultant circulant matrix $\mathbf{C}$ can then be decomposed as $\mathbf{C} = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}$ where $\mathbf{F}$ is a $2N \times 2N$ Fourier matrix and $\mathbf{D}$ is a diagonal matrix whose elements are the discrete Fourier transform of the first column of $\mathbf{C}$. Note that the diagonal of $\mathbf{T}'$ is arbitrary, but it is normally equal to the first sample of the previous block $k-1$ [7]. We now define the frequency domain quantities: $\underline{\mathbf{y}}(m) = \mathbf{F}\begin{bmatrix} \mathbf{0}_{N\times 1} \\ \mathbf{y}(m) \end{bmatrix}$, $\underline{\hat{\mathbf{h}}}_k(m) = \mathbf{F}\begin{bmatrix} \hat{\mathbf{h}}_k(m) \\ \mathbf{0}_{N\times 1} \end{bmatrix}$, $\underline{\mathbf{e}}(m) = \mathbf{F}\begin{bmatrix} \mathbf{0}_{N\times 1} \\ \mathbf{e}(m) \end{bmatrix}$, $\mathbf{G}^{01} = \mathbf{F}\mathbf{W}^{01}\mathbf{F}^{-1}$, $\mathbf{W}^{01} = \begin{bmatrix} \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{I}_{N\times N} \end{bmatrix}$, $\mathbf{G}^{10} = \mathbf{F}\mathbf{W}^{10}\mathbf{F}^{-1}$ and $\mathbf{W}^{10} = \begin{bmatrix} \mathbf{I}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \end{bmatrix}$. The MDF adaptive algorithm is then given by the following equations:

$$\underline{\mathbf{e}}(m) = \underline{\mathbf{y}}(m) - \mathbf{G}^{01}\sum_{k=0}^{K-1} \mathbf{D}(m-k)\underline{\hat{\mathbf{h}}}_k(m-1) \quad (16)$$

$$\mathbf{S}_{\text{MDF}}(m) = \ \mathbf{S}_{\text{MDF}}(m-1) + (1- \ )\mathbf{D}^*(m)\mathbf{D}(m) \quad (17)$$

$$\underline{\hat{\mathbf{h}}}_k(m) = \underline{\hat{\mathbf{h}}}_k(m-1) + \ \mathbf{G}^{10}\mathbf{D}^*(m-k)\times [\mathbf{S}_{\text{MDF}}(m) + \ _{\text{MDF}}]^{-1}\underline{\mathbf{e}}(m) \quad (18)$$

where * denotes complex conjugate, $0 \ll \ < 1$ is the forgetting factor and $= (1- \ )$ is the adaptive step-size with $0 < \ \le 1$. Letting $_x^2$ be the input signal variance, the initial regularization parameters are $\mathbf{S}_{\text{MDF}}(0) = \ _x^2/100$ and $_{\text{MDF}} = 20 \ _x^2 N/L$ [7].

3. IPMDF ALGORITHM

We now propose the IPMDF algorithm which incorporates proportional updates as in IPNLMS whilst achieving efficient implementation by exploiting the MDF algorithm described in Section 2.2.

We note that direct use of $\mathbf{Q}(n)$, with elements as described by (6), into the weight update equation in (18) is inappropriate since the former is in the time-domain whereas the latter is in the frequency-domain. Thus our proposed method will be to update the filter coefficients in the time-domain. This is achieved by first defining the matrix

$$\widetilde{\mathbf{G}}^{10} = \mathbf{W}^{10}\mathbf{F}^{-1}. \tag{19}$$

We next define

$$\mathbf{q}_k(m) = [q_{kN}(m)\ \ q_{kN+1}(m)\ldots q_{kN+N-1}(m)], \quad k = 0,\ 1,\ldots,K-1 \tag{20}$$

as the partitioned control elements of the k^{th} block such that each element in this block is now determined by

$$q_{kN+j}(m) = \frac{1-}{2L} + (1+\ \ )\frac{|\hat{h}_{kN+j}(m)|}{2\|\hat{\mathbf{h}}\|_1 +} \tag{21}$$

where $k = 0,\ 1,\ldots,K-1$ is the block-index while $j = 0,\ 1,\ldots,N-1$ is the tap-index of each k^{th} block. The IPMDF filter update equation is then given by

$$\begin{aligned}\hat{\mathbf{h}}_k(m) \quad &= \quad \hat{\mathbf{h}}_k(m-1) + L\ \ \mathbf{Q}_k(m)\widetilde{\mathbf{G}}^{10}\mathbf{D}^*(m-k) \\ &\quad \times[\mathbf{S}_{\text{IPMDF}}(m) + \ \ _{\text{IPMDF}}]^{-1}\underline{\mathbf{e}}(m)\end{aligned} \tag{22}$$

where the diagonal control matrix $\mathbf{Q}_k(m) = \text{diag}\{\mathbf{q}_k(m)\}$. The IPMDF algorithm can be summarized as follows:

IPMDF Algorithm

$$\begin{aligned}
_{\text{IPMDF}} &= \tfrac{(1-\ \ )\ \ _x^2 20N}{2L} \\
&= [1-\tfrac{1}{3L}]^N \\
&= \quad (1-\ \ ),\ 0 < \quad \le 1 \\
\mathbf{S}_{\text{IPMDF}}(0) &= \tfrac{(1-\ \ )\ \ _x^2}{2\times 100} \\
\hat{\mathbf{h}}(0) &= \mathbf{0}_{L\times 1} \\
\hat{\mathbf{h}}_k(m) &= [\hat{h}_{kN}(m)\ \ \hat{h}_{kN+1}(m)\ldots\hat{h}_{kN+N-1}(m)]^T \\
q_{kN+j}(m) &= \tfrac{1-}{2L} + (1+\ \ )\tfrac{|\hat{h}_{kN+j}(m)|}{2\|\hat{\mathbf{h}}\|_1 +},\quad j = 0,1,\ldots,N-1 \\
\mathbf{q}_k(m) &= [q_{kN}(m)\ \ q_{kN+1}(m)\ldots q_{kN-N+1}(m)] \\
\mathbf{Q}_k(m) &= \text{diag}\{\mathbf{q}_k(m)\} \\
\underline{\mathbf{e}}(m) &= \underline{\mathbf{y}}(m) - \mathbf{G}^{01}\ \ \sum_{k=0}^{K-1}\mathbf{D}(m-k)\underline{\hat{\mathbf{h}}}_k(m-1) \\
\mathbf{S}_{\text{IPMDF}}(m) &= \ \ \mathbf{S}_{\text{IPMDF}}(m-1) + (1-\ \ )\mathbf{D}^*(m)\mathbf{D}(m) \\
\hat{\mathbf{h}}_k(m) &= \hat{\mathbf{h}}_k(m-1) + L\ \ \mathbf{Q}_k(m)\widetilde{\mathbf{G}}^{10}\mathbf{D}^*(m-k)\times \\
&\quad [\mathbf{S}_{\text{IPMDF}}(m) + \ \ _{\text{IPMDF}}]^{-1}\underline{\mathbf{e}}(\text{m}).
\end{aligned}$$

4. COMPUTATIONAL COMPLEXITY

We note that although the IPMDF algorithm is updated in the time-domain, the error $\underline{\mathbf{e}}(m)$ is generated using frequency domain coefficients and hence five FFT-blocks are required. Since a $2N$ point FFT requires $2N\log_2 N$ real multiplications, the number of multiplications required per output sample for each algorithm is described by the following relations:

$$\begin{aligned}
\text{IPNLMS} &: \ 4L \\
\text{FLMS} &: \ 8+10\log_2 L \\
\text{MDF} &: \ 8K+(4K+6)\log_2 N \\
\text{IPMDF} &: \ 10K+(4K+6)\log_2 N.
\end{aligned}$$

It can be seen that the complexity of IPMDF is only modestly higher than MDF. However, as we shall see in Section 5, the performance of IPMDF far exceeds that of MDF for both speech and white Gaussian noise (WGN) inputs.

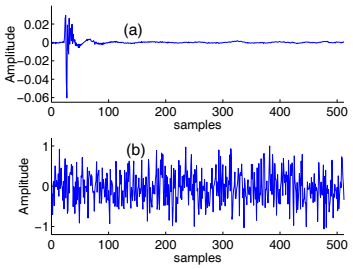

Figure 2: Impulse responses used in simulations: (a) sparse and (b) dispersive.

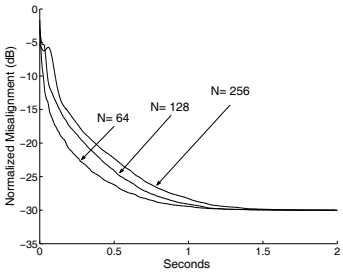

Figure 3: IPMDF convergence for different N with sparse impulse response.

5. SIMULATION RESULTS

We now wish to compare, by way of simulation, the performance of MDF and IPMDF in the context of network echo cancellation. We illustrate the robustness of IPMDF to 'sparseness' of the impulse response using two different echo paths $\mathbf{h}$ each of length $L = 512$ as shown in Fig. 2. The adaptive filter $\hat{\mathbf{h}}(m)$ was chosen to be of the same length as $\mathbf{h}$ and we define the normalized misalignment as

$$\|\mathbf{h} - \hat{\mathbf{h}}(m)\|^2/\|\mathbf{h}\|^2. \tag{23}$$

In all our simulations, the sampling frequency is 8 kHz and the signal-to-noise ratio is 30 dB while the following parameters were chosen for all simulations:
$\ = -0.75,\quad \ = [1-1/(3L)]^N,\quad \ = 1\times(1-\ \ ),$
$\mathbf{S}_{\text{MDF}}(0) = \ \ _x^2/100,\quad _{\text{MDF}} = \ \ _x^2 20N/L,$
$\mathbf{S}_{\text{IPMDF}}(0) = (1-\ \ )\ \ _x^2/200$ and $\ _{\text{IPMDF}} = 20(1-\ \ )\ \ _x^2 N/(2L).$

5.1 Sparse impulse response simulations

In the first experiment, a sparse impulse response as shown in Fig. 2(a) was used. Figure 3 shows the convergence with various frame-sizes N for IPMDF. It can be seen that the convergence is faster for smaller N since the adaptive filter coefficients are being updated more frequently. The normalized misalignment of the IPNLMS, IPMDF and MDF algorithms is compared in Fig. 4 using a WGN input sequence. The frame-size for IPMDF and MDF were $N = 64$ while the step-size of IPNLMS was adjusted so that its final misalignment is the same as IPMDF and MDF. This corresponds to $\ _{\text{IPNLMS}} = 0.15$. We can see that there is a significant improvement in normalized misalignment of approximately 5 dB for the IPMDF compared to MDF and IPNLMS.

The tracking performance of the proposed IPMDF algorithm is now compared with MDF and IPNLMS for WGN input sequence as shown in Fig. 5. In this simulation an echo

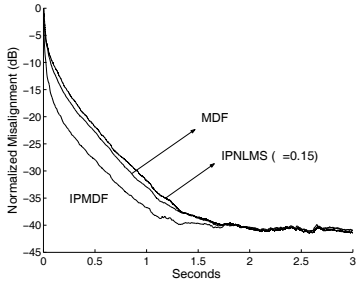

Figure 4: IPMDF, MDF and IPNLMS algorithms for a sparse impulse response.

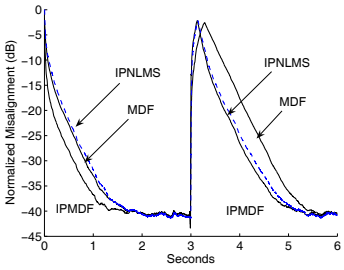

Figure 5: IPMDF and MDF and IPNLMS algorithms for a sparse impulse response with echo path change at 3 s.

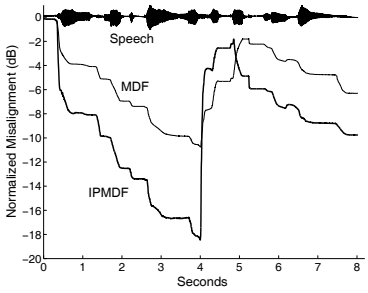

Figure 6: IPMDF and MDF algorithms for a sparse impulse response with speech input and echo path change at 4 s.

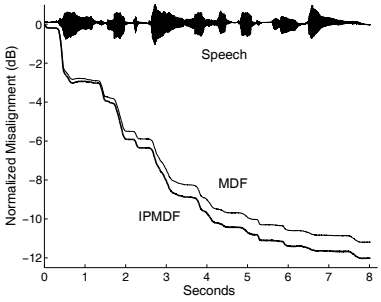

Figure 7: IPMDF and MDF algorithms for a dispersive impulse response using speech.

path change, comprising an additional 12 samples delay, was introduced after 3 s, and $N = 64$. We see that the IPMDF algorithm achieves the fastest initial convergence. When compared with MDF, IPMDF has higher tracking capability achieving approximately 8 dB improvement in normalized misalignment following the echo path change at 3 s. When compared to IPNLMS, the improvement in tracking capability of IPMDF is approximately 2 dB. This modest improvement is limited by the inherent delay that exists in frequency domain algorithms.

We have also tested using speech from a male talker as shown in Fig. 6. The sparse impulse response was delayed by an additional 12 samples at 4 s. It can be seen that the normalized misalignment of IPMDF is approximately 8 dB lower than that of MDF before the echo path change was introduced. The IPMDF algorithm exhibits a 4 dB improvement in normalized misalignment as compared to MDF after an echo path change was introduced.

5.2 Dispersive impulse response with speech input

In this last simulation, we illustrate that the performance of the IPMDF is better than that of the MDF algorithm even for a dispersive impulse response such as shown in Fig 2(b) using speech input from a male talker. Figure 7 shows the normalized misalignment plot of IPMDF and MDF with $N = 64$ for both algorithms. It can be seen that the convergence performance for IPMDF is approaching 1 dB higher than MDF even for this dispersive impulse response.

6. CONCLUSION

We have proposed the IPMDF algorithm for echo cancellation with sparse impulse responses. This algorithm exploits both the improvement in convergence brought about by the proportionality control of IPNLMS and the efficient implementation in the frequency-domain of MDF. Simulation results, using both WGN and speech inputs, have shown that the improvement in initial convergence and tracking of IPMDF over MDF for both sparse and dispersive impulse responses far outweighs the modest increase in computational cost.

REFERENCES

[1] J. Radecki, Z. Zilic, and K. Radecka, "Echo cancellation in IP networks," in *Proc. 45th Midwest Symposium on Circuits and Systems*, vol. 2, 2002, pp. 219–222.

[2] D. L. Duttweiler, "Proportionate normalized least mean square adaptation in echo cancellers," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 5, pp. 508–518, Sep. 2000.

[3] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, vol. 2, 2002, pp. 1881–1884.

[4] J. Cui, P. A. Naylor, and D. T. Brown, "An improved IPNLMS algorithm for echo cancellation in packet-switched networks," in *Proc. IEEE Int. Conf. on Signal Processing*, vol. 4, May 2004, pp. 141–144.

[5] S. Haykin, *Adaptive Filter Theory*, 4th ed., ser. Information and System Science. Prentice Hall, 2002.

[6] J. S. Soo and K. K. Pang, "Multidelay block frequency domain adaptive filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 2, pp. 373–376, Feb. 1990.

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