A SPARSENESS CONTROLLED PROPORTIONATE ALGORITHM FOR ACOUSTIC ECHO CANCELLATION

Pradeep Loganathan [1], *Andy W.H. Khong* [2], *Patrick A. Naylor* [1]

[1] Electrical and Electronic Engineering,
Imperial College, London, United Kingdom
email: {pradeep.loganathan, p.naylor}@ic.ac.uk

[2] Electrical and Electronic Engineering,
Nanyang Technological University, Singapore
email: andykhong@ntu.edu.sg

ABSTRACT

Sparseness variation in acoustic impulse response arises due to changes in temperature, pressure, acoustic source movements and changes in the acoustic environment. Therefore, the algorithms employed in acoustic echo cancellation have to perform well for both sparse and dispersive unknown systems. The well-known algorithms, normalised least mean square (NLMS) or proportionate NLMS (PNLMS), are limited to perform well either in dispersive or sparse cases, respectively. The proposed sparseness-controlled PNLMS (SC-PNLMS) algorithm inherits the beneficial properties of both PNLMS and NLMS by employing the sparseness measure into the PNLMS algorithm. Simulation results presented show improved performance over the PNLMS algorithm even for dispersive impulse responses.

1. INTRODUCTION

Hands-free communication is regarded as an essential tool due to their flexibility. As the use for in-car telephony gain much popularity in recent years due to the rise in safety concerns, digital wireless subscribers are becoming more critical of the voice quality they receive from network providers. In the case of hands-free mobile telephony, acoustic echoes can seriously degrade user experience. For this reason, effective acoustic echo cancellation (AEC) is important to maintaining and improving the perceived voice quality of a call. Traditionally, adaptive filters have been deployed in acoustic echo cancellers, as illustrated in Fig. 1. These cancellers achieve echo cancellation by estimating the acoustic impulse responses (AIRs) using adaptive algorithms such as the normalized least-mean-square (NLMS) algorithm. As AIRs are time-varying in nature, these adaptive algorithms track variations in the impulse response to achieve sufficient level of echo cancellation.

The time variation of AIRs may arise due to, for example, a change in temperature [1], pressure, movement of the acoustic source [2][3] and changes in the acoustic environment. It is well known that the reverberation time of an AIR is proportional to the volume of the enclosed space and inversely proportional to the absorption area [4]. For an outdoor environment, the reverberation time is reduced significantly due to the lack of reflections from the walls. As a consequence, the AIR of an outdoor environment can be considered to be more sparse than that of an enclosed space. Hence, algorithms developed for mobile hands-free terminals have to be robust to changes in the sparseness of the acoustic path.

Variation in the sparseness of AIRs can also occur in AEC within an enclosed space. Consider an example case where the distance, a, between a fixed position loudspeaker
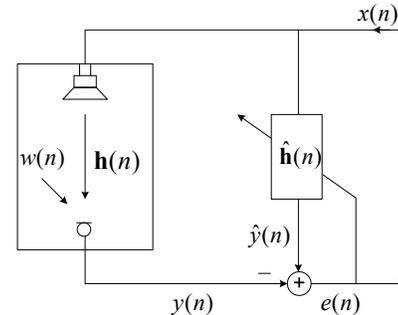

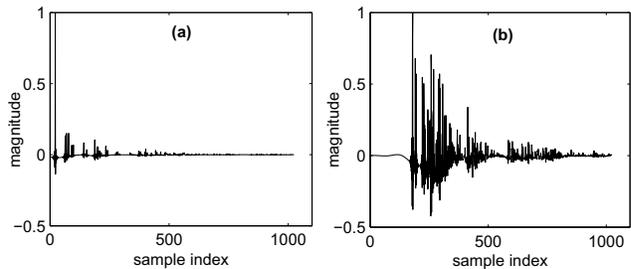Figure 1: Adaptive system for acoustic echo cancellation.

Figure 2: Acoustic impulse responses $\mathbf{h}(n)$ obtained from the image model [5] using room dimensions of $\{8 \times 10 \times 3\}$ m where the distances between the loudspeaker and microphone are (a) 0.9 m and (b) 7.7 m.

and the user using a wireless microphone is varying. Figure 2 shows two AIRs, generated using the method of images [5], for the cases when (a) $a = 0.9$ m and (b) $a = 7.7$ m. As can be seen, the sparseness of these AIRs vary with the loudspeaker-microphone distance. Hence, performance of the AECs are required to be robust to the variation in sparseness of the AIRs.

Sparse system identification, such as occurs in packet-switched telephone networks for network echo cancellation (NEC), has been the focus of research in recent voice over IP (VoIP) applications. Network echo response in such systems is, typically of length 64-128 ms, characterised by an unknown bulk delay due to network loading, encoding and jitter buffer delays [6]. This results in an 'active' region in the range of 8-12 ms duration and consequently, the impulse response is dominated by 'inactive' regions where magnitudes are close to zero, making the impulse response sparse. For such applications, the NLMS algorithm suffers from slow convergence and therefore new algorithms have

been proposed for sparse adaptive filtering. One of the first algorithms for NEC is the proportionate NLMS (PNLMS) algorithm [7] where each filter coefficient is updated with an independent step-size that is proportional to the magnitude of that estimated filter coefficient. It is well known that PNLMS suffers from slow convergence when estimating dispersive AIRs. To address this, subsequent improved versions such as the improved PNLMS (IPNLMS) [8] algorithm was proposed. The IPNLMS achieves improved convergence by introducing a controlled mixture of proportionate (PNLMS) and non-proportionate (NLMS) adaptation. A sparseness controlled IPNLMS (SC-IPNLMS) algorithm was proposed in [3] to improve the robustness of IPNLMS to the sparseness variation in AIRs.

In this paper, we propose an algorithm that is robust to the sparseness variation of AIRs. This algorithm employs the PNLMS for the estimation of sparse AIRs. We then propose to improve the convergence of PNLMS by incorporating the sparseness control factor for dispersive AIRs. As will be shown, the proposed sparseness-controlled PNLMS (SC-PNLMS) algorithm achieves fast convergence for both sparse and dispersive AIRs which is effective for AEC.

2. ALGORITHMS FOR ECHO CANCELLATION

Figure 1 shows an AEC set up in a Loudspeaker-Room-Microphone system (LRMS) and an adaptive filter $\widehat{\mathbf{h}}(n)$ deployed to cancel acoustic echo. Defining the input signal $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-L+1)]^T$ and $\mathbf{h}(n) = [h_0(n) \ h_1(n) \ \dots \ h_{L-1}(n)]^T$ as the unknown impulse response, the output of the LRMS is given by

$$y(n) = \mathbf{h}^T(n)\mathbf{x}(n) + w(n), \tag{1}$$

where $[\cdot]^T$ is the transposition operator, $w(n)$ is the additive noise and L is the length of $\mathbf{h}(n)$. The AEC employs an adaptive filter with coefficients $\widehat{\mathbf{h}}(n) = [\widehat{h}_0(n) \ \widehat{h}_1(n) \ \dots \ \widehat{h}_{L-1}(n)]^T$ in order to estimate $\mathbf{h}(n)$ using the error signal

$$e(n) = y(n) - \widehat{\mathbf{h}}^T(n-1)\mathbf{x}(n). \tag{2}$$

Several adaptive algorithms such as those described below have been developed for AEC and NEC.

2.1 The NLMS algorithm

The NLMS algorithm is one of the most popular algorithm for AEC due to its simplicity in implementation and its relatively lower complexity compared to the better performing recursive least squares algorithm. The NLMS algorithm can be described by (2) and the following set of equations:

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \frac{\mu \mathbf{Q}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{Q}(n-1)\mathbf{x}(n) + \delta}, \tag{3}$$

$$\mathbf{Q}(n-1) = \text{diag}\{q_0(n-1) \ \dots \ q_{L-1}(n-1)\}, \tag{4}$$

where μ is the step-size and δ is the regularization parameter. The diagonal step-size control matrix $\mathbf{Q}(n)$ determines the step-size of each filter coefficient and is dependent on the specific algorithm. For NLMS, since the step-size is the same for all filter coefficients, $\mathbf{Q}(n) = \mathbf{I}_{L \times L}$ with $\mathbf{I}_{L \times L}$ being an $L \times L$ identity matrix.

One of the main drawbacks of the NLMS algorithm is that its convergence rate reduces significantly, compared to

an NLMS algorithm that updates only the active region of the impulse response, when the impulse response is sparse such as occur in network impulse response. This is due to the adaptation noise that occurs for the inactive region of the estimated impulse response.

2.2 The PNLMS algorithm

The PNLMS algorithm, originally developed in NEC for sparse system identification, achieves high convergence rate by allocating step-sizes that are proportional to the magnitude of the estimated impulse response coefficients. The PNLMS employs the following set of equations for computing elements of the time varying step-size control matrix $\mathbf{Q}(n)$ during adaptation:

$$q_l(n) = \frac{\kappa_l(n)}{\frac{1}{L}\sum_{i=0}^{L-1}\kappa_i(n)}, \qquad 0 \le l \le L-1, \tag{5}$$

$$k_l(n) = \max\{\rho \times \max\{\gamma, |\widehat{h}_0(n)| \ \dots \ |\widehat{h}_{L-1}(n)|\}, |\widehat{h}_l(n)|\}. \tag{6}$$

The parameter γ in (6), with a typical value of 0.01, prevents filter coefficients $\widehat{h}_l(n)$ from stalling when $\widehat{\mathbf{h}}(0) = \mathbf{0}_{L \times 1}$ at initialisation and ρ, with a typical value of 0.01, prevents coefficients from stalling when they are much smaller than the largest coefficient. As can be seen from (5) and (6), the PNLMS algorithm allocates larger step-sizes to the "active" coefficients and hence, PNLMS converges faster than NLMS for sparse impulse responses. However, the PNLMS experiences fast initial convergence, follow by a slower second phase convergence [9]. This slower phase adaptation is due to the slower convergence for the small magnitude filter coefficients.

It is important to note that PNLMS suffers from slow convergence when the unknown system $\mathbf{h}(n)$ is dispersive [10][11]. This is because when $\mathbf{h}(n)$ is dispersive, $\kappa_l(n)$ in (6) becomes significantly large for most $0 \le l \le L-1$. As a consequence, the denominator of $q_l(n)$ in (5) is large giving a small step-size for each large coefficient. This causes a significant degradation in convergence performance for PNLMS when the impulse response is dispersive.

3. A SPARSENESS CONTROLLED PROPORTIONATE ALGORITHM

We propose to improve the robustness of PNLMS to the sparseness of impulse response for AEC. As will be shown in the following, this is achieved by incorporating the sparseness measure of the estimated AIRs into the adaptation for PNLMS, in a different manner compared to SC-IPNLMS [3], since it employs single term in (5), instead of proportionate and NLMS terms for IPNLMS.

3.1 Variation of sparseness in AIRs

The degree of sparseness for an impulse response can be quantified by [3][12]

$$\xi(n) = \frac{L}{L - \sqrt{L}}\left\{1 - \frac{\|\mathbf{h}(n)\|_1}{\sqrt{L}\|\mathbf{h}(n)\|_2}\right\} \tag{7}$$

where $\|\cdot\|_1$ is defined as the l_1-norm while L is the length of the unknown filter $\mathbf{h}(n)$. It can be shown [3][12] that
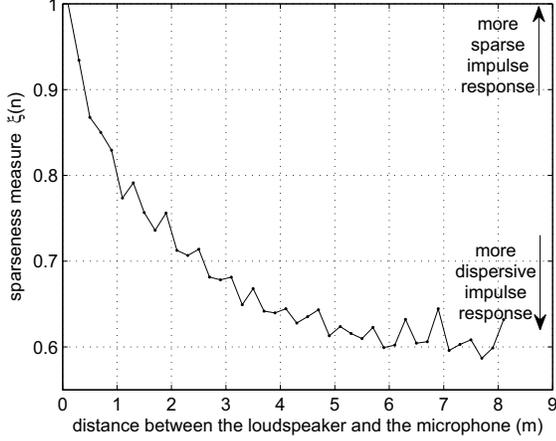
Figure 3: Sparseness measure against the distance between loudspeaker and microphone, a. The impulse responses are obtained from the image model proposed in [5] using a fixed room dimensions of $\{8 \times 10 \times 3\}$ m.

Figure 4: Magnitude of $q_l(n)$ for $0 \le l \le L-1$ against the magnitude of tap coefficients $\widehat{h}_l(n)$.

Figure 5: Variation of ρ against sparseness measure $\widehat{\xi}(n)$ of impulse response.

$0 \le \xi(n) \le 1$ such that if the impulse response is sparse with $\mathbf{h}(n) = [k\ 0\ \ldots\ 0]^T$ and $k \in \Re$, then $\xi(n) = 1$. On the other hand, for a dispersive impulse response with $\mathbf{h}(n) = \pm k$, then $\xi(n) = 0$. In reality, as explained in Section 1, $\mathbf{h}(n)$ is time-varying and depends on factors such as temperature and pressure.

Consider an example case where the distance, a, between a fixed position loudspeaker and the talker using a wireless microphone is varying. Figure 3 illustrates how $\xi(n)$ varies with a for a room of dimension $\{8 \times 10 \times 3\}$ m and the loudspeaker is placed in $\{4 \times 9.1 \times 1.6\}$ m. For each loudspeaker-microphone distance, the microphone is directly in-front of the loudspeaker a. As can be seen, $\xi(n)$ reduces with increasing a and hence, we propose to incorporate $\xi(n)$ into PNLMS for fast convergence in AEC. Since $\mathbf{h}(n)$ is unknown during adaptation, we employ $\widehat{\xi}(n)$ to estimate the sparseness of an impulse response, where

$$\widehat{\xi}(n) = \frac{L}{L-\sqrt{L}}\left\{1 - \frac{\|\widehat{\mathbf{h}}(n)\|_1}{\sqrt{L}\,\|\widehat{\mathbf{h}}(n)\|_2}\right\}. \qquad (8)$$

3.2 Effect of ρ on step-size control matrix $\mathbf{Q}(n)$

As explained in Section 2.2, the parameter ρ in (6) was originally introduced to prevent freezing of the filter coefficients when they are much smaller than the largest coefficient. Figure 4 shows the effect of $\rho = 0.001$, 0.01 and 0.1 on the step-size control element $q_l(n)$. It can be seen from this illustration that a higher value of ρ will ensure that all filter coefficients will be updated with the same step-size while for a smaller ρ, the step-size gain is proportional to the magnitude of filter coefficients $|\widehat{h}_l(n)|$. Hence, for a dispersive unknown system $\mathbf{h}(n)$, we desire a high value of ρ while for a sparse $\mathbf{h}(n)$, we desire a low value of ρ.

3.3 The proposed SC-PNLMS algorithm

In order to address the problem of slow convergence in PNLMS for dispersive AIR, we require the step-size control elements $q_l(n)$ to 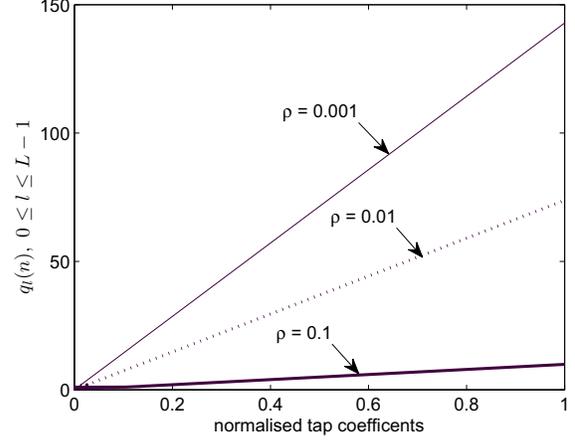be robust to the sparseness of the impulse response. We now propose to incorporate $\widehat{\xi}(n)$ into the computation of ρ for PNLMS. Several choices can be employed to obtain the desired effect of achieving a high ρ when $\widehat{\xi}(n)$ is small in estimating dispersive AIRs. We consider two example functions where
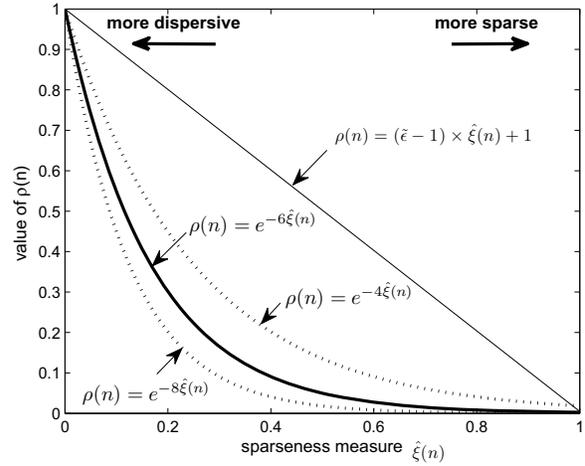
$$\rho(n) = (\tilde{\varepsilon} - 1)\widehat{\xi}(n) + 1, \qquad (9)$$

$$\rho(n) = e^{-\lambda\widehat{\xi}(n)}, \qquad \lambda \in \mathbb{R}^+, \qquad (10)$$

such that $\tilde{\varepsilon}$ in (9) is a small value to prevent $\rho(n) = 0$ when $\widehat{\xi}(n) = 1$. The variation of $\rho(n)$ for these functions are plotted in Fig. 5 for the cases where $\lambda = 4$, 6 and 8.

As can be seen from Fig. 5, the linear function as described by (9) will not achieve our desired effect of allocating large step-size to coefficients with large magnitudes when the AIR is sparse such as for $0.8 \le \widehat{\xi}(n) \le 1$. This is because, as can be seen from Fig. 4, the value of $\rho(n)$ is not low enough to achieve the desired proportionality control determined by $q_l(n)$.

Figure 5 also illustrates how $\rho(n)$ varies with $\widehat{\xi}(n)$ for different values of λ. Comparing to the linear function, lower values of $\rho(n)$ are allocated for a larger range of sparse impulse responses such as for $0.8 \leq \widehat{\xi}(n) \leq 1$. As a result, the proposed sparseness-controlled PNLMS algorithm (SC-PNLMS) using (10) inherits the proportionality step-size control over a larger range of sparse impulse response. When the impulse response is dispersive, such as for $\widehat{\xi}(n) < 0.4$, the proposed SC-PNLMS algorithm inherits the NLMS adaptation control with larger values of $\rho(n)$. This gives a more uniform step-size across $h_l(n)$ as can be seen from Fig. 4. Hence, the exponential function described by (10) will achieve our overall desired effect of the robustness to sparse and dispersive AIRs.

Although the exponential function described in (10) is favorable compared to the linear function, the choice of λ is important. As can be seen from Fig. 5, a larger choice of λ will cause the proposed algorithm to inherit more of PNLMS properties compared to NLMS. As a consequence, the performance of SC-PNLMS is reduced when the AIR is dispersive. A good compromise must be made and as will be shown in Section 5, a good choice of λ is 6. In addition, we note that when $n = 0$, $\|\widehat{\mathbf{h}}(0)\|_2 = 0$ and hence to prevent division by a small number or zero, $\widehat{\xi}(n)$ can be computed for $n \geq L$. The SC-PNLMS algorithm is thus described by (2)-(6) and (10) with $\lambda = 6$.

4. COMPUTATIONAL COMPLEXITY

The relative complexity of NLMS, PNLMS and SC-PNLMS in terms of the total number of additions, multiplications, divisions and comparisons per iteration is assessed in Table 1. The additional complexity of the SC-PNLMS algorithm, on top of the PNLMS, arises from the computation of the sparseness measure $\widehat{\xi}(n)$. Given that $L/(L - \sqrt{L})$ in (7) can be computed off-line, the remaining l-norms require an additional $2L$ additions and L multiplications. The SC-PNLMS algorithm additionally requires some computations for (10). Alternatively, a look-up table with values of $\rho(n)$ defined in (10) can be computed for $0 \leq \widehat{\xi}(n) \leq 1$. For an example case of $L = 1024$, SC-PNLMS requires only a modest 33% more computations than PNLMS.

5. SIMULATION RESULTS

We present simulation results to evaluate the performance of the proposed SC-PNLMS algorithm in the context of AEC. We assumed throughout our simulation that the length of the adaptive filter $L = 1024$ is equivalent to that of the unknown system. The receiving room impulse response $\mathbf{h}(n)$ is generated synthetically using the image model [5] with a room dimension of $\{8 \times 10 \times 3\}$ m and sampling frequency 8 kHz. A reflection coefficient of 0.57 is used. The loudspeaker is fixed at $\{4 \times 9.1 \times 1.6\}$ m in the LRMS while the microphone is positioned at $\{8 \times 8.2 \times 3\}$ m and $\{8 \times 1.4 \times 3\}$ m giving impulse responses shown in Fig. 2 (a) and (b) respectively. The sparseness measure of these AIRs are computed using (7) giving (a) $\xi(n) = 0.83$ and (b) $\xi(n) = 0.59$ respectively. The performance of each algorithm is then quantified using the normalised misalignment defined by

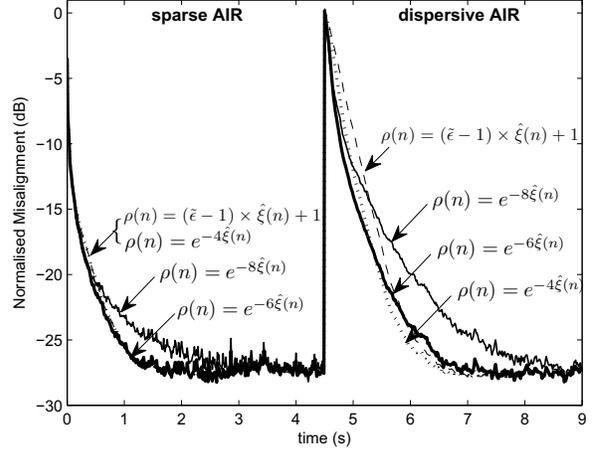$$\eta(n) = \frac{\|\mathbf{h}(n) - \widehat{\mathbf{h}}(n)\|_2^2}{\|\mathbf{h}(n)\|_2^2}. \tag{11}$$

Figure 6: Convergence of the SC-PNLMS for the exponential function with different values of λ and the linear function using WGN input signal with echo path changes at 4.5 s. Impulse response is changed from Fig. 2 (a) to (b) and $\mu_{\text{SC-PNLMS}} = 0.3$, SNR = 20 dB.

For the first result shown in Fig. 6, linear function (9) and different λ values in (10) were employed to illustrate the performance of SC-PNLMS. A zero mean white Gaussian noise (WGN) sequence is used as the input signal while another WGN sequence is added to give an SNR of 20 dB. An echo path change was introduced using impulse responses shown from Fig. 2 (a) to (b) at 4.5 s. As explained in Section 3.3, the exponential function achieves better convergence performance than the linear function. Moreover, the proposed algorithm inherits properties of the NLMS for a small λ value. As a result, a faster rate of convergence can be seen after the echo path is changed to a dispersive AIR. For a high λ, the SC-PNLMS inherits properties of the PNLMS giving good performance for sparse AIR before the echo path change. As can be seen, a good compromise of λ is given by $\lambda = 6$.

Figure 7 illustrates the performance of NLMS, PNLMS and SC-PNLMS using WGN as the input signal. The step-size parameters for all algorithms are adjusted so that they reach the same steady state performance. This corresponds to $\mu_{\text{NLMS}} = \mu_{\text{PNLMS}} = \mu_{\text{SC-PNLMS}} = 0.3$. As before an echo path change was introduced using AIRs as shown from Fig. 2(a) to 2(b). It can be seen from Fig. 7 that the convergence rate of SC-PNLMS is the highest for both sparse and dispersive AIRs. More importantly, the SC-PNLMS inherits the beneficial properties of both PNLMS and NLMS. This can be seen from the result that SC-PNLMS achieves high rate of convergence similar to PNLMS giving approximately 5 dB improvement in normalised misalignment during initial convergence compared to NLMS for a sparse AIR. After the echo path change, for a dispersive AIR, the SC-PNLMS is close to the performance of NLMS giving approximately 4 dB improvement in normalised misalignment compared to PNLMS. The ability to achieve good convergence performance for both sparse and dispersive AIR for SC-PNLMS is due to the beneficial properties of both PNLMS and NLMS.

Figure 8 shows simulation results for a male speech signal where we used the same parameters as before. As can be seen, the proposed SC-PNLMS algorithm achieves the high-

Table 1: Complexity of algorithms

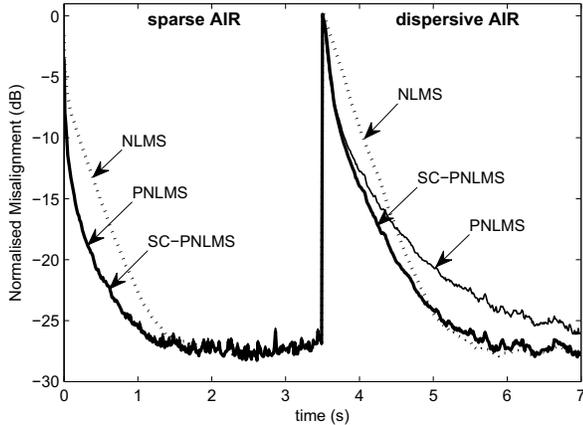| Algorithm	Addition	Multiplication	Division	Comparison
NLMS	$L+3$	$2L+3$	1	0
PNLMS	$2L+3$	$5L+4$	$L+1$	$2L$
SC-PNLMS	$4L+6$	$6L+8$	$L+3$	$2L$

Figure 7: Relative convergence of NLMS, PNLMS and SC-PNLMS using WGN input signal with an echo path change at 3.5 s. Impulse response is changed from that shown from Fig. 2 (a) to (b) and $\mu_{\text{NLMS}} = \mu_{\text{PNLMS}} = \mu_{\text{SC-PNLMS}} = 0.3$, SNR = 20 dB.
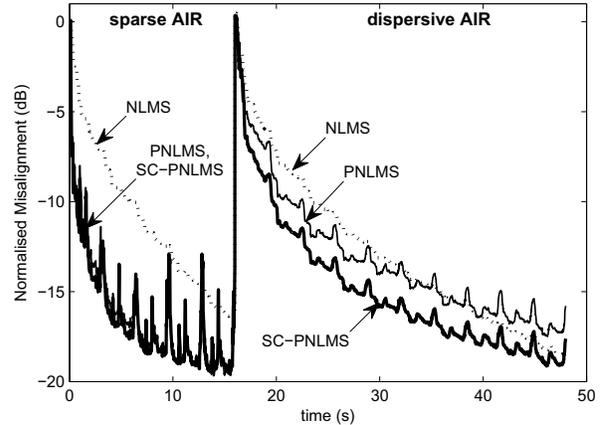

Figure 8: Relative convergence of NLMS, PNLMS and SC-PNLMS using speech input signal with echo path changes at 16 s. Impulse response is changed from that shown in Fig. 2 (a) to (b) and $\mu_{\text{NLMS}} = \mu_{\text{PNLMS}} = \mu_{\text{SC-PNLMS}} = 0.3$, SNR = 20 dB.

est rate of convergence giving approximately 7 dB improvement in normalised misalignment during initial convergence compared to NLMS for the sparse AIR and approximately 2 dB improvement for dispersive AIR compared to PNLMS.

6. CONCLUSION

The NLMS algorithm achieves good convergence in dispersive AIRs, whereas PNLMS performs well in sparse impulse response. We propose to incorporate the sparseness measure into PNLMS for AEC to achieve fast convergence that is robust to sparse and dispersive impulse response. The proposed SC-PNLMS algorithm takes into account the sparseness measure via the coefficient update function. Simulation results show that the SC-PNLMS algorithm exhibits robustness to sparse and dispersive AIRs than PNLMS and NLMS for a modest increase in computational complexity.

REFERENCES

[1] G. W. Elko, E. Diethorn, and T. Gänsler, "Room impulse response variation due to thermal fluctuation and its impact on acoustic echo cancellation," in *Proc. Int. Workshop on Acoustic Echo and Noise Control*, 2003, pp. 67–70.

[2] E. Hänsler, "The hands-free telephone problem- an annotated bibliography," *Signal Processing*, vol. 27, no. 3, pp. 259–271, Jun. 1992.

[3] A. W. H. Khong and P. A. Naylor, "Efficient use of sparse adaptive filters," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference*, Oct 2006.

[4] H. Sabine, "Room acoustics," *Transactions of the IRE Professional Group on Room Acoustics*, vol. 1, no. 4, pp. 4–12, Jul. 1953.

[5] 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.

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

[7] 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.

[8] 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.

[9] H. Deng and M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Processing Lett.*, vol. 12, no. 3, pp. 181–184, Mar. 2005.

[10] S. L. Gay, "An efficent, fast converging adaptive filter for network echo cancellation," vol. 1, November 1998, pp. 394–398.

[11] A. Deshpande and S. L. Grant, "A new multi-algorithm approach to sparse system adaptation," in *Proc. European Signal Process. Conf.*, 2005.

[12] J. Benesty, Y. A. Huang, J. Chen, and P. A. Naylor, "Adaptive algorithms for the identification of sparse impulse responses," in *Selected methods for acoustic echo and noise control*, E. Hänsler and G. Schmidt, Eds. Springer, 2006, ch. 5, pp. 125–153.