

Sparsity Promoting LMS for Adaptive Feedback Cancellation

Ching-Hua Lee, Bhaskar D. Rao, and Harinath Garudadri

Department of Electrical and Computer Engineering
University of California, San Diego
{chl438, brao, hgarudadri}@ucsd.edu

Abstract—In hearing aids (HAs), the acoustic coupling between the microphone and the receiver results in the system becoming unstable under certain conditions and causes artifacts commonly referred to as whistling or howling. The least mean square (LMS) class of algorithms is commonly used to mitigate this by providing adaptive feedback cancellation (AFC). The speech quality after AFC and the amount of added stable gain (ASG) with AFC are used to assess these algorithms. In this paper, we introduce a variant of the LMS that promotes sparsity in estimating the acoustic feedback path. By using the l_p norm as a diversity measure, the approach does not enforce, but takes advantage of sparsity when it exists. The performance in terms of speech quality, misalignment, and ASG of the proposed algorithm is compared with other proportionate-type LMS algorithms which also leverage sparsity in the feedback path. We demonstrate faster convergence compared with those algorithms, quality improvement of about 0.25 (on a 0–1 objective scale of the hearing-aid speech quality index (HASQI)), and about 5 dB ASG improvement compared with the normalized LMS (NLMS).

Index Terms—Adaptive feedback cancellation, LMS, proportionate adaptation, sparsity, hearing aids

I. INTRODUCTION

In hearing aids (HAs), acoustic feedback is a well-known problem that causes howling and whistling effects that are annoying to the users. Under certain conditions, the receiver signal feeding back to the microphone will make the system become unstable. This not only degrades the audio quality but also limits the amount of amplification that can be provided by the HA. To overcome this many adaptive feedback cancellation (AFC) techniques have been proposed for modern HAs [1].

In AFC, an adaptive filter is continuously adjusting to approximate the impulse response (IR) of the acoustic feedback path. In the adaptation stage, least mean square (LMS) algorithms [2] are the most widely used techniques due to computational simplicity and their effectiveness. However, the estimate given by the LMS is inevitably biased due to the correlation between the incoming signal and the feedback signal [3]. Several methods have been proposed to address this bias issue such as the filtered-X LMS (FXLMS) [4], [5], the prediction-error-method (PEM) based AFC [6], insertion of probe noise [7], phase modulation [8], PEM with frequency shifting [9], the dual-microphone approach [10], etc.

Assuming that the bias problem can be well handled by the above techniques, a question of interest is: Can we further improve the convergence behavior of the AFC from other

aspects? Observing that typical feedback path IRs are (quasi) sparse as shown in Fig. 2, one might think of taking advantage of this sparseness for improvements. This can actually be carried out by the concept of proportionate adaptation that originated from the proportionate normalized LMS (PNLMS) algorithm [11]. The main idea behind proportionate adaptation is to update each filter coefficient independently of the others by assigning to the corresponding step size a weight in proportion to the magnitude of the estimated coefficient. In other words, it redistributes the adaptation gains among all coefficients and emphasizes the large ones in order to speed up their convergence.

However, the original PNLMS has the problem that it is more beneficial for systems with very sparse structures. For AFC application where the feedback path IRs are usually quasi-sparse, other proportionate-type LMS algorithms can be more suitable. For example, the improved PNLMS (IPNLMS) [12] and the IPNLMS- l_0 [13] have the flexibility for identifying systems of different levels of sparsity. Attempts have been made to incorporate these proportionate algorithms into AFC [14]–[16] and improvements have been reported. However, these proportionate algorithms were not formally derived by minimizing any underlying objective functions so that their usage can be further optimized. Moreover, the parameters within these algorithms do not have direct connections to the sparsity degree of the underlying system they aim to identify.

In this paper, we propose a new AFC approach that incorporates sparsity in the feedback path estimate based on the pLMS algorithm of [17], [18]. This method, called Sparsity promoting LMS (SLMS), takes advantage of sparseness in the feedback path when such sparsity is present while not enforcing it. By adding the l_p norm as a diversity measure to the objective function of the ordinary LMS, the algorithm is derived using the affine scaling method [19] in the minimization procedure. The benefit of using the l_p norm is brought by its direct connection to the system sparseness which provides a practical way of parameter selection. Our algorithm has the advantages of enjoying theoretical support, simpler parameter optimization, and more straightforward leverage of (quasi) sparsity in acoustic feedback paths. Simulation results using feedback paths measured with real HAs show that the SLMS outperforms other proportionate-type LMS algorithms in terms of audio quality, misalignment, and added stable gain (ASG).

II. AFC SYSTEM

Fig 1. shows the typical AFC framework. The AFC filter $W(z)$ is an FIR filter placed in parallel with the HA processing $G(z)$ that continuously adjusts its coefficients to emulate the IR of the feedback path $F(z)$. $x(n)$ is the desired input signal and $d(n)$ is the actual input to the microphone, which contains $x(n)$ and the feedback signal $y(n)$ generated by the HA output $s(n)$ passing through $F(z)$. $\hat{y}(n)$ is the estimate of $y(n)$ given by the output of $W(z)$. $e(n) = d(n) - \hat{y}(n)$ is the feedback-compensated signal which, ideally, should be identical to $x(n)$. In practice, however, the AFC is not perfect and therefore $\hat{y}(n) \neq y(n)$, resulting in distortion between $e(n)$ and $x(n)$.

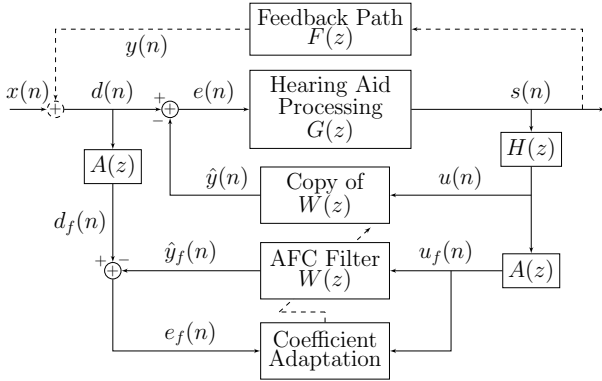


Fig. 1. Block diagram of the AFC framework.

In our system we adopt the PEM based AFC framework [6] where a time-varying pre-filter $A(z)$ is present and adapted using linear prediction of $e(n)$ [20]. We also employ a band-limited filter $H(z)$ from the FXLMS approach to concentrate on the frequency regions where oscillation is likely to occur [5]. This filter can also be viewed as a very rough approximation of the feedback path in the frequency domain [4].

In this framework, the LMS for coefficient adaptation is carried out in the pre-filtered signal domain, where the pre-filtered signal $u_f(n)$ and $e_f(n)$ are used to update the L -tap AFC filter $\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T$ as:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)\mathbf{u}_f(n)e_f(n), \quad (1)$$

where $\mathbf{u}_f(n) = [u_f(n), u_f(n-1), \dots, u_f(n-L+1)]^T$ and a time-varying step size $\mu(n)$ is usually employed to improve the convergence rate:

$$\mu(n) = \frac{\mu}{L\hat{\sigma}^2(n) + \epsilon}, \quad (2)$$

with $\mu > 0$ the step size parameter, ϵ a small positive constant to prevent division by zero, and $\hat{\sigma}^2(n) = \rho\hat{\sigma}^2(n-1) + (1-\rho)(u_f^2(n) + e_f^2(n))$ the power estimate term with a forgetting factor $0 < \rho \leq 1$. This is actually the (modified) normalized LMS (NLMS) [21] and has been widely used in speech processing especially for AFC in HAs [5], [6], [22].

III. PROPORTIONATE ALGORITHMS FOR AFC

To incorporate the proportionate adaptation idea into the AFC, the update rule (1) becomes:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)\mathbf{P}(n)\mathbf{u}_f(n)e_f(n), \quad (3)$$

where

$$\mathbf{P}(n) = \text{diag}\{p_0(n), p_1(n), \dots, p_{L-1}(n)\} \quad (4)$$

is an L -by- L diagonal matrix assigning different weights to the step sizes for different filter taps. We refer to it as the "proportionate matrix".

Applying the PNLMS [11], at the n -th iteration the diagonal entries of $\mathbf{P}(n)$ are computed as:

$$p_l(n) = r_l(n) / \left(\frac{1}{L} \sum_{i=0}^{L-1} r_i(n) \right), \quad (5)$$

for $l = 0, 1, \dots, L-1$ where

$$r_l(n) = \max\{\kappa \max\{\zeta, |w_0(n)|, \dots, |w_{L-1}(n)|\}, |w_l(n)|\}, \quad (6)$$

with positive constants κ and ζ .

When using the IPNLMS [12], we have:

$$r_l(n) = (1-\alpha) \frac{\|\mathbf{w}(n)\|_1}{L} + (1+\alpha)|w_l(n)|. \quad (7)$$

Substituting (7) into (5) gives:

$$p_l(n) = \frac{(1-\alpha)}{2} + (1+\alpha) \frac{|w_l(n)|}{\frac{2}{L} \sum_{i=0}^{L-1} |w_i(n)| + \delta}, \quad (8)$$

where a small constant $\delta > 0$ is added to prevent division by zero and $-1 \leq \alpha \leq 1$ is a parameter that can be chosen for different sparsity levels: When $\alpha = 1$ it behaves like the PNLMS while for $\alpha = -1$ it reduces to the NLMS.

When using the IPNLMS- l_0 [13], the l_0 norm approximation

$$\|\mathbf{w}(n)\|_0 = \lim_{\beta \rightarrow \infty} \sum_{i=0}^{L-1} [1 - e^{-\beta|w_i(n)|}] \approx \sum_{i=0}^{L-1} [1 - e^{-\beta|w_i(n)|}] \quad (9)$$

is used to replace the l_1 norm in the IPNLMS. This gives:

$$r_l(n) = (1-\alpha) \frac{\|\mathbf{w}(n)\|_0}{L} + (1+\alpha) [1 - e^{-\beta|w_l(n)|}], \quad (10)$$

which results in:

$$p_l(n) = \frac{(1-\alpha)}{2} + (1+\alpha) \frac{1 - e^{-\beta|w_l(n)|}}{\frac{2}{L} \sum_{i=0}^{L-1} [1 - e^{-\beta|w_i(n)|}] + \delta}, \quad (11)$$

with another parameter β that provides control for identifying systems with different degrees of sparsity.

The PNLMS is more suitable for very sparse systems. The IPNLMS and IPNLMS- l_0 have parameters (α and β) for fitting different sparsity degrees but without direct connections to the system sparsity levels. Moreover, the formulations of these algorithms do not have theoretical foundation.

IV. PROPOSED SPARSITY PROMOTING LMS ALGORITHM

In this section we present the derivation of the proposed AFC algorithm that promotes sparsity. Let $E[\cdot]$ denote the expectation operator. In the pre-filtered signal domain of AFC, the ordinary LMS considers the following mean squared error (MSE) minimization problem:

$$\begin{aligned} \min_{\mathbf{w}} J(\mathbf{w}) &= E[|e_f(n)|^2] = E[|d_f(n) - \hat{y}_f(n)|^2] \\ &= E[|d_f(n) - \mathbf{w}^T \mathbf{u}_f(n)|^2] \\ &= E[|d_f(n)|^2] - 2\mathbf{w}^T \mathbf{b} + \mathbf{w}^T \mathbf{R} \mathbf{w}, \end{aligned} \quad (12)$$

where $\mathbf{R} \triangleq E[\mathbf{u}_f(n)\mathbf{u}_f^T(n)]$ and $\mathbf{b} \triangleq E[\mathbf{u}_f(n)d_f(n)]$.

To enforce sparsity on the solution \mathbf{w} , we modify (12) as:

$$J(\mathbf{w}) = E[|d_f(n)|^2] - 2\mathbf{w}^T \mathbf{b} + \mathbf{w}^T \mathbf{R} \mathbf{w} + \gamma \|\mathbf{w}\|_p^p, \quad (13)$$

where the l_p norm diversity measure $\|\mathbf{w}\|_p^p = \sum_{l=0}^{L-1} |w_l|^p$, $p \in (0, 2]$ is added with a regularization parameter $\gamma > 0$ to promote sparsity in the solution \mathbf{w} [17], [18]. From [19], the gradient of the l_p norm term w.r.t. \mathbf{w} is given by:

$$\nabla_{\mathbf{w}} \|\mathbf{w}\|_p^p = p \Pi(\mathbf{w}) \mathbf{w}, \quad (14)$$

where $\Pi(\mathbf{w}) = \text{diag}(|w_l|^{p-2})$. Applying this to the gradient of (13) and setting to zero, we obtain the following condition for the optimal solution \mathbf{w}_o :

$$\mathbf{w}_o = \Pi^{-\frac{1}{2}}(\mathbf{w}_o) (\Pi^{-\frac{1}{2}}(\mathbf{w}_o) \mathbf{R} \Pi^{-\frac{1}{2}}(\mathbf{w}_o) + \tilde{\gamma} \mathbf{I})^{-1} \Pi^{-\frac{1}{2}}(\mathbf{w}_o) \mathbf{b}, \quad (15)$$

where $\tilde{\gamma} \triangleq \frac{\gamma p}{2}$ and \mathbf{I} denotes the identity matrix. This suggests the following iterative procedure for computing \mathbf{w}_o :

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{M}(n) (\mathbf{M}(n) \mathbf{R} \mathbf{M}(n) + \tilde{\gamma} \mathbf{I})^{-1} \mathbf{M}(n) \mathbf{b} \\ &= \mathbf{M}(n) (\mathbf{R}(n) + \tilde{\gamma} \mathbf{I})^{-1} \mathbf{b}(n), \end{aligned} \quad (16)$$

where we have defined:

$$\mathbf{M}(n) \triangleq \Pi^{-\frac{1}{2}}(\mathbf{w}(n)) = \text{diag}(|w_l(n)|^{\frac{2-p}{2}}) \quad (17)$$

and auxiliary variables: $\mathbf{R}(n) \triangleq \mathbf{M}(n) \mathbf{R} \mathbf{M}(n)$ and $\mathbf{b}(n) \triangleq \mathbf{M}(n) \mathbf{b}$.

At the n -th iteration, for $\mathbf{w} \in \mathbb{R}^L$, we define the corresponding affinely scaled variable as [19]:

$$\mathbf{q}(\mathbf{w}) \triangleq \Pi^{\frac{1}{2}}(\mathbf{w}(n)) \mathbf{w} = \mathbf{M}^{-1}(n) \mathbf{w}. \quad (18)$$

With this transformation into the affine scaling domain, we have at the n -th iteration:

$$\mathbf{q}(n) \equiv \mathbf{q}(\mathbf{w}(n)) = \mathbf{M}^{-1}(n) \mathbf{w}(n). \quad (19)$$

$$\mathbf{q}(n+1) \equiv \mathbf{q}(\mathbf{w}(n+1)) = \mathbf{M}^{-1}(n) \mathbf{w}(n+1). \quad (20)$$

Using (20), the update rule (16) can be equivalent to:

$$\mathbf{q}(n+1) = (\mathbf{R}(n) + \tilde{\gamma} \mathbf{I})^{-1} \mathbf{b}(n). \quad (21)$$

Observing that (21) is actually the minimizer of the following quadratic objective function:

$$J(\mathbf{q}) = E[|d_f(n) - \mathbf{q}^T \mathbf{M}(n) \mathbf{u}_f(n)|^2] + \tilde{\gamma} \|\mathbf{q}\|_2^2, \quad (22)$$

we can derive the steepest descent algorithm as:

$$\begin{aligned} \mathbf{q}(n+1) &= \mathbf{q}(n) - \frac{\mu}{2} \nabla_{\mathbf{q}} J(\mathbf{q}) \\ &= (1 - \mu \tilde{\gamma}) \mathbf{q}(n) + \mu (\mathbf{b}(n) - \mathbf{R}(n) \mathbf{q}(n)), \end{aligned} \quad (23)$$

where $\mu > 0$ is the step size parameter.

Transforming (23) back to the original coefficient domain using (19) and (20) we obtain:

$$\mathbf{w}(n+1) = (1 - \mu \tilde{\gamma}) \mathbf{w}(n) + \mu \mathbf{M}^2(n) (\mathbf{b} - \mathbf{R} \mathbf{w}(n)). \quad (24)$$

Finally, to derive the adaptive algorithm, we replace \mathbf{R} and \mathbf{b} in (24) by their corresponding instantaneous values $\mathbf{u}_f(n)\mathbf{u}_f^T(n)$ and $\mathbf{u}_f(n)d_f(n)$, respectively [2]. This gives:

$$\mathbf{w}(n+1) = (1 - \mu \tilde{\gamma}) \mathbf{w}(n) + \mu \mathbf{M}^2(n) \mathbf{u}_f(n) e_f(n), \quad (25)$$

where $\mathbf{M}^2(n) = \text{diag}(|w_l(n)|^{2-p})$ from (17). This is actually the pLMS algorithm proposed in [17]. Note that $\mathbf{M}^2(n)$ has a similar role as the proportionate matrix $\mathbf{P}(n)$ in (3). In fact, it indicates that the weight assigned to each step size is in proportion to $|w_l(n)|^{2-p}$. Based on this and the previous discussion on proportionate adaptation for AFC, we propose the following update rule:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n) \mathbf{P}(n) \mathbf{u}_f(n) e_f(n), \quad (26)$$

where $\mu(n)$ is the normalized step size as (2), $\mathbf{P}(n) = \text{diag}\{p_0(n), p_1(n), \dots, p_{L-1}(n)\}$ as (4) with $p_l(n) = r_l(n) / (\frac{1}{L} \sum_{i=0}^{L-1} r_i(n))$, for $l = 0, 1, \dots, L-1$ as (5), and:

$$r_l(n) = |w_l(n)| + c|^{2-p}, \quad (27)$$

where $p \in (0, 2]$ and $c > 0$ is a small positive constant.

Several observations can be made here. First, comparing (26) to (25), we have set γ to zero as it has negligible effect as long as we keep it small [18]. In such case, due to the presence of $\mathbf{P}(n)$, the algorithm will still exploit, though not enforce, the sparsity of the solution if it already exists. We then refer to it as Sparsity promoting LMS (SLMS). However, a practical issue arises when $\gamma = 0$ for (25): Once any of the AFC filter taps becomes 0, it will not get updated anymore. We therefore suggest adding a small constant $c > 0$ to all the taps as in (27) before they are used to compute $\mathbf{P}(n)$.

The parameter p is much more influential: When p is smaller (close to or less than 1), the algorithm becomes more sparsity promoting due to the nature of the l_p norm; while when $p = 2$, it reduces to the NLMS. This indicates that a sparse system would benefit more from a smaller p while for a dispersive system, p close to 2 would be more preferable. From this it becomes clearer the advantage of incorporating the l_p diversity measure. For the quasi-sparse feedback IRs in AFC, we expect that the optimal p value would lie between 1 and 2.

V. SIMULATION RESULTS

The above algorithms are evaluated using computer simulations in MATLAB at a sampling rate of 16 kHz. The experimental set up was as follows: The HA processing $G(z) = gz^{-d}$ with $g = 20$ and d corresponding to a delay of 8 ms. The feedback path IRs were measured using a behind-the-ear HA with open fitting on a dummy head and truncated to a length of 263 samples as shown in Fig. 2. The AFC filter length was $L = 100$ to cover the significant part of the IRs. The forgetting factor $\rho = 0.985$. The step size parameter $\mu = 0.005$. Small positive constants $\epsilon = \delta = c = 10^{-6}$. The band-limited filter $H(z) = 1 - 1.8z^{-1} + 0.81z^{-2}$ as used in [4]. The pre-filter $A(z)$ was an FIR filter of order 20 updated every 10 ms via linear prediction of $e(n)$ [20].

For the purpose of evaluation, the following metrics were used. First, to measure the distortion between $e(n)$ and $x(n)$,

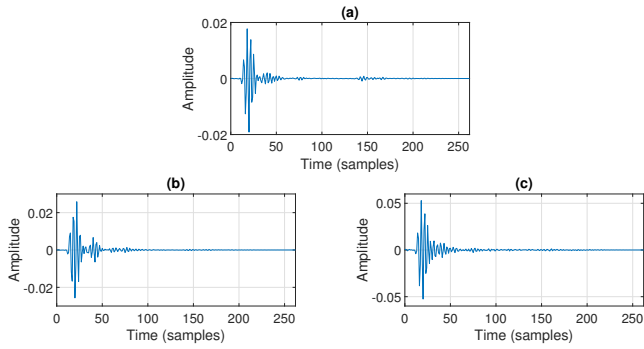


Fig. 2. Measured acoustic feedback path IRs of (a) f_1 : no obstruction, (b) f_2 : with a cellphone close to the ear, and (c) f_3 : with a cellphone right on the ear.

the hearing-aid speech quality index (HASQI) was used [23]. The HASQI score ranges from 0 to 1, where the higher the score, the better the quality. For evaluating convergence performance the normalized misalignment was considered:

$$\text{Misalignment} = 10 \log_{10} \frac{\int_0^{\pi} |F(e^{j\omega}) - \hat{F}(e^{j\omega})|^2 d\omega}{\int_0^{\pi} |F(e^{j\omega})|^2 d\omega}, \quad (28)$$

where $F(e^{j\omega})$ and $\hat{F}(e^{j\omega})$ are the frequency responses of the measured and estimated feedback IRs, respectively. Note that the estimated feedback response is given by the band-limited filter and the AFC filter together as $\hat{F}(e^{j\omega}) = H(e^{j\omega})W(e^{j\omega})$. For estimating the ASG we used [24]:

$$\text{ASG} = 20 \log_{10} \left(\min_{\omega} \frac{1}{|F(e^{j\omega}) - \hat{F}(e^{j\omega})|} \right) - 20 \log_{10} \left(\min_{\omega} \frac{1}{|F(e^{j\omega})|} \right). \quad (29)$$

In our first experiment we examine the effect of p in the SLMS on channels with different sparsity levels. In addition to the measured feedback path f_1 we considered two other artificial channels as plotted in Fig. 3. The input was a stationary speech-shaped noise. Fig. 4 shows the convergence behavior in terms of misalignment. We can see that for the feedback path case, $p = 1.5$ outperforms other values. For the sparser system, a smaller p around 1.2 is more preferable; while p around 1.8 gives the best performance for the dispersive one. These results show that the SLMS is exploiting the underlying system structure in the way we expect. Furthermore, in Fig. 4 (a) we can see that the SLMS (with a good choice of p) can improve the convergence rate over the NLMS and still maintains low steady-state error, which is not achievable by only using a larger step size parameter μ for the NLMS.

In the next experiment, we ran the AFC system with the SLMS on 25 male and 25 female speech signals from TIMIT database and measured the corresponding HASQI of the feedback-compensated signal $e(n)$. Average HASQI scores over the 50 speech files for different values of p are shown in Fig. 5. We can see that the optimal p almost lies in the same range even as the feedback IR differs. This means, for a given HA device, if we have some rough knowledge about the

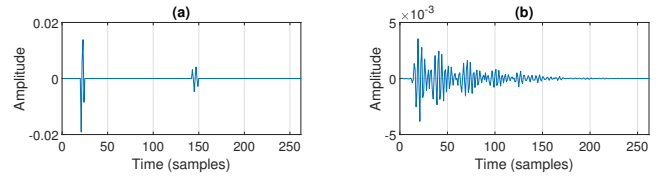


Fig. 3. Artificial IRs of (a) a sparse channel and (b) a dispersive channel.

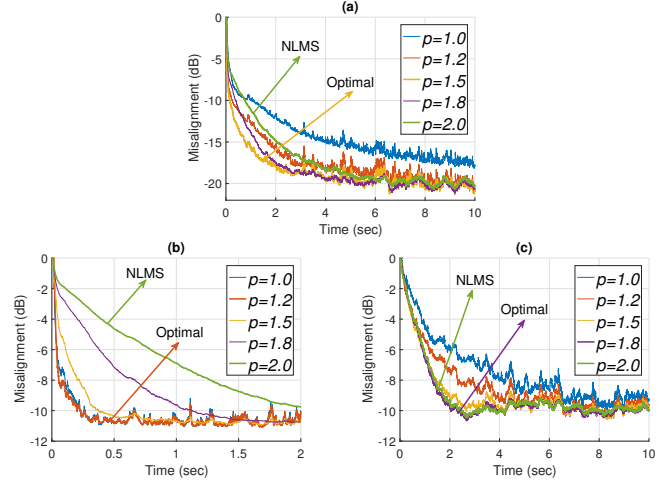


Fig. 4. Misalignment comparison of SLMS with different p on (a) the measured feedback path f_1 , (b) the sparse channel, and (c) the dispersive channel with speech-shaped noise input.

sparsity degree of its feedback channel, the SLMS is robust since p is not very sensitive near the optimal point. From the results we found p around 1.5 to be a good choice, which also corresponds to the result we had in Fig. 4 (a).

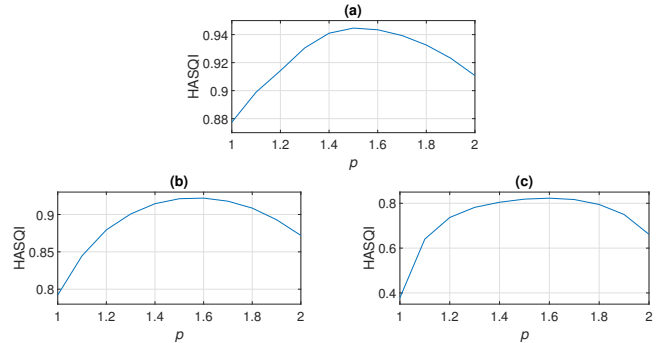


Fig. 5. Effect of p on speech quality of SLMS for (a) f_1 , (b) f_2 , and (c) f_3 .

We also compare the SLMS with other proportionate algorithms for the AFC. The parameter settings were – PNLMS: $\kappa = 0.1$ and $\zeta = 0.01$; IPNLMS: $\alpha = 0$; IPNLMS- l_0 : $\alpha = 0$ and $\beta = 150$. For the SLMS we used $p = 1.5$. The NLMS (which is equivalent to IPNLMS and IPNLMS- l_0 with $\alpha = 0$ and SLMS with $p = 2$) is also compared. Fig. 6 compares the tracking performance in terms of misalignment and ASG with speech-shaped noise input. To model a highly time-varying feedback back environment, the feedback path was initially f_1 , switching to f_2 then f_3 at the 1/3 and 2/3 of the input sequence, respectively. We can see that the proportionate algorithms

basically have faster convergence speed compared to the NLMS. Among all, the SLMS shows the best convergence behavior and can provide up to about 5 dB additional ASG compared to the NLMS.

Finally, for further verification, we ran the algorithms on the speech dataset and measured the average HASQI under 4 different feedback scenarios as shown in Fig. 7. We see that the SLMS outperforms all the other ones, especially obvious under an adverse feedback situation such as the last two cases (about 0.25 HASQI improvement compared to the NLMS in the last case).

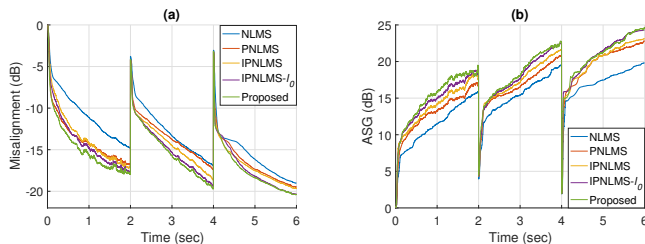


Fig. 6. Misalignment and ASG with speech-shaped noise input.

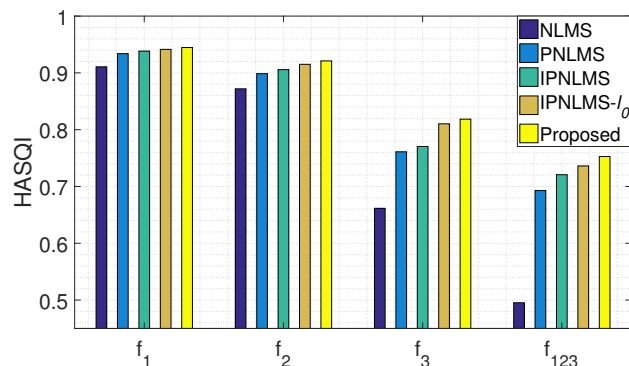


Fig. 7. Comparison of speech quality for different feedback environments. The first three cases were fixed environments with f_1 , f_2 , and f_3 . The last case f_{123} was the feedback path changing from f_1 to f_2 then f_3 at 1/3 and 2/3 of the input sequence, respectively.

VI. CONCLUSION

In this paper we introduce the SLMS for AFC to exploit sparsity in estimating the feedback path. This approach extends the LMS algorithm by incorporating the l_p diversity measure in the objective function. We derive update rules and discuss guidelines for choosing the system parameters. We present simulation results with speech-shaped stimulus and speech segments with real-world feedback paths, including conditions where the feedback path changes over time. The results show that for the SLMS (i) choice of p is not very sensitive around the optimal point; (ii) HASQI improvement of about 0.25 over the NLMS; (iii) ASG improvement by 5 dB compared to the NLMS; and (iv) better performances compared with other proportionate-type LMS algorithms.

ACKNOWLEDGMENT

This work was supported by NIH/NIDCD grants #1R21DC015046 and #1R01DC015436.

REFERENCES

- [1] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, 2011.
- [2] S. Haykin, *Adaptive filter theory*, 5th edition, Pearson, 2013.
- [3] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 4, pp. 443–453, 2000.
- [4] J. Hellgren, "Analysis of feedback cancellation in hearing aids with Filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, 2002.
- [5] H.-F. Chi, S. X. Gao, S. D. Soli, and A. Alwan, "Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids," *Speech Commun.*, vol. 39, no. 1-2, pp. 147–161, 2003.
- [6] A. Spriet, S. Doclo, M. Moonen, and J. Wouters, "Feedback control in hearing aids," *Springer Handbook of Speech Process.*, pp. 979–1000, 2008.
- [7] M. Guo, S. H. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancement," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, 2012.
- [8] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Proc. Europ. Signal Process. Conf. (EUSIPCO)*, 2012, pp. 2000–2004.
- [9] F. Strasser and H. Puder, "Adaptive feedback cancellation for realistic hearing aid applications," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 12, pp. 2322–2333, 2015.
- [10] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2012, pp. 149–152.
- [11] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, 2000.
- [12] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2002, pp. 1881–1884.
- [13] C. Paleologu, J. Benesty, and S. Ciochină, "An improved proportionate NLMS algorithm based on the l_0 norm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2010, pp. 309–312.
- [14] H. Shen and L. Zhang, "A new variable step-size algorithm on acoustic feedback suppression for digital hearing aids," in *Proc. IEEE Int. Conf. Audio, Lang., Image Process. (ICALIP)*, 2014, pp. 171–175.
- [15] F. Albu, C. R. C. Nakagawa, and S. Nordholm, "Proportionate algorithms for two-microphone active feedback cancellation," in *Proc. Europ. Signal Process. Conf. (EUSIPCO)*, 2015, pp. 290–294.
- [16] Vasundhara, G. Panda, and N. B. Puhana, "A VSS sparseness controlled algorithm for feedback suppression in hearing aids," in *Proc. IEEE Int. Symp. Signal Process. Info. Tech. (ISSPIT)*, 2015, pp. 151–156.
- [17] B. D. Rao and B. Song, "Adaptive filtering algorithms for promoting sparsity," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2003, pp. 361–364.
- [18] Y. Jin, *Algorithm development for sparse signal recovery and performance limits using multiple-user information theory*, Ph.D. dissertation, Univ. California, San Diego, 2011.
- [19] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 187–200, 1999.
- [20] S. Gazor and T. Liu, "Adaptive filtering with decorrelation for coloured AR environments," *IEE Proc. Vision, Image, Signal Process.*, vol. 152, no. 6, pp. 806–818, 2005.
- [21] J. E. Greenberg, "Modified LMS algorithms for speech processing with an adaptive noise canceller," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 4, pp. 338–351, 1998.
- [22] H. Schepker, L. T. T. Tran, S. Nordholm, and S. Doclo, "Improving adaptive feedback cancellation in hearing aids using an affine combination of filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2016, pp. 231–235.
- [23] J. M. Kates and K. H. Arehart, "The hearing-aid speech quality index (HASQI) version 2," *J. Audio Eng. Soc.*, vol. 62, no. 3, pp. 99–117, 2014.
- [24] J. M. Kates, "Room reverberation effects in hearing aid feedback cancellation," *J. Acoust. Soc. Am.*, vol. 109, no. 1, pp. 367–378, 2001.