

## Adaptive personal sound zones systems with online plant modelling

Sipei ZHAO<sup>1</sup> and Ian S. BURNETT<sup>2</sup>

<sup>1,2</sup> Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, University of Technology Sydney,  
Australia

### ABSTRACT

Personal sound zones systems have attracted significant research interest recently due to its broad potential applications in car cabins, mobile devices and public spaces *etc.* Most existing studies focus on optimal performance in stationary environments assuming that accurate plant from loudspeakers to microphones are known *a priori*. However, in practical applications, the acoustic environments may change over time and the accurate plant may not be available. This paper presents an adaptive personal sound system with online plant modeling for non-stationary environments. The control filters are updated based on the filtered- $x$  least mean squares algorithm and the plant models are adapted with the recursive least squares algorithm. Simulations with measured impulse responses are performed to evaluate the performance of a ten-loudspeaker array for the creation of two sound zones. Simulation results demonstrate that an adaptive system with online plant modeling converges to the solution with an ideal plant model for both stationary and non-stationary environments.

Keywords: Personal audio systems, Personal sound zones, Multizone sound, Online plant modelling

### 1. INTRODUCTION

Personal sound zones systems (1), aiming to generate independent listening zones for multiple users in a common physical space, have been studied intensively in the past two decades due to their wide ranging applications. Various methods, such as acoustic contrast control (2), pressure matching (3), mode matching (4) and variable span linear filtering (5), have been proposed to optimize the driving signals of loudspeakers. Optimization of the placement of those loudspeakers has also been investigated (6,7). These various approaches focus on performance in stationary environments based on the assumption that accurate plant is known *a priori*. In practice, however, acoustic environments may be time-varying, hence adaptive systems are required to track changes in the plant.

Recently, the Filtered- $x$  Least Mean Squares (FxLMS) algorithm has been explored to adapt personal sound systems (8). This paper extends that work to include Online Plant Modeling (OPM) for non-stationary environments (9). OPM has been investigated in multichannel Active Noise Control (ANC) systems and two main approaches are found in the literature. The first approach involves the injection of additional random noise (10) while the second uses control filter output signals directly (11). This paper will employ the second approach to avoid interference of the injected noise to personal sound systems. It is noted that personal sound systems differ from ANC systems in that the microphones measure the reproduced sound only, rather than the error signals. To the best of the authors' knowledge, OPM has not been studied in the context of personal sound systems. This paper presents a theoretical formulation of the adaptive personal sound systems with OPM first, and then validates the algorithm by simulations with measured impulse responses.

<sup>1</sup> Sipei.Zhao@uts.edu.au

<sup>2</sup> Ian.Burnett@uts.edu.au

## 2. THEORY

### 2.1 Problem formulation

Figure 1 illustrates a personal sound system with an arc-shaped array of  $L$  loudspeakers and two control zones, i.e., the bright and dark zones. The input signal  $x[n]$  at time instant  $n$  is filtered by a  $K$ -tap Finite Impulse Response (FIR) filter  $\mathbf{w}_l$  ( $l = 1, 2, \dots, L$ ) before being sent to the  $l$ -th loudspeaker for reproduction. The filters are designed to generate the bright and dark zones with high and low sound pressure levels, respectively. The plant from the  $l$ -th loudspeaker to the  $m$ -th microphone in the bright zone is modelled as a FIR filter  $\mathbf{h}_{B,ml}$  with a length of  $J$ . Similarly, the definition  $\mathbf{h}_{D,ml}$  is used for the plant model to the dark zone.

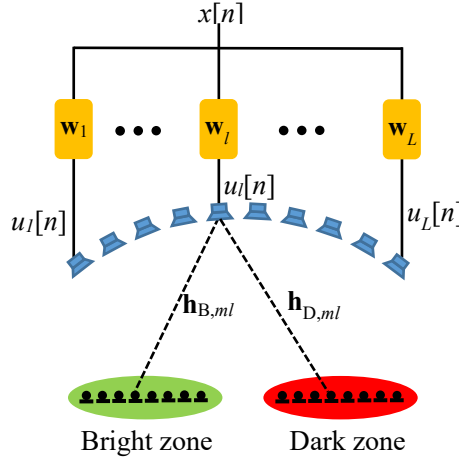


Figure 1 – Diagram of the personal audio systems

The output signal of the  $l$ -th loudspeaker is

$$u_l[n] = \sum_{k=0}^{K-1} x[n-k] w_l(k), \quad (1)$$

which can be written in vector form as

$$u_l[n] = \mathbf{x}[n]^T \mathbf{w}_l, \quad (2)$$

where  $\mathbf{x}[n] = [x[n], x[n-1], \dots, x[n-K+1]]^T$  and  $\mathbf{w}_l = [w_l(1), w_l(2), \dots, w_l(K)]^T$ .

The sound pressure at the  $m$ -th microphone in the bright zone is

$$p_{B,m}[n] = \sum_{l=1}^L \sum_{j=0}^{J-1} u_l[n-j] h_{B,ml}(j) = \sum_{l=1}^L \mathbf{h}_{B,ml}^T \mathbf{u}_l[n], \quad (3)$$

where  $h_{ml}(j)$  is the  $j$ -th coefficient of the acoustic path from the  $l$ -th loudspeaker to the  $m$ -th microphone  $\mathbf{h}_{B,ml} = [h_{B,ml}(0), h_{B,ml}(1), \dots, h_{B,ml}(J-1)]^T$  and  $\mathbf{u}_l[n] = [u_l[n], u_l[n-1], \dots, u_l[n-J+1]]^T$ .

Substituting Eq. (2) into Eq. (3) yields

$$p_{B,m}[n] = \sum_{l=1}^L \mathbf{h}_{B,ml}^T \mathbf{X}[n] \mathbf{w}_l, \quad (4)$$

where  $\mathbf{X}(n) = [\mathbf{x}[n], \mathbf{x}[n-1], \dots, \mathbf{x}[n-J+1]]^T$ . By vectorizing the sound pressure at all the microphones, the sound pressure in the bright zone is

$$\mathbf{p}_B[n] = \begin{bmatrix} \mathbf{r}_{B,11} & \cdots & \mathbf{r}_{B,1L} \\ \vdots & \ddots & \vdots \\ \mathbf{r}_{B,M1} & \cdots & \mathbf{r}_{B,ML} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_L \end{bmatrix} = \mathbf{R}_B \mathbf{w}, \quad (5)$$

where  $\mathbf{r}_{B,ml} = \mathbf{h}_{B,ml}^T \mathbf{X}$  are the filtered input signals based on the slow adaptation assumption.

Similarly, the sound pressure in the dark zone can be written as

$$\mathbf{p}_D[n] = \mathbf{R}_D \mathbf{w} \quad (6)$$

with  $\mathbf{r}_{D,ml} = \mathbf{h}_{D,ml}^T \mathbf{X}$  being the  $(m, l)$ -th element of the matrix  $\mathbf{R}_D$ .

The cost function of the weighted pressure matching method is

$$J(\mathbf{w}) = \mathbb{E} \left\{ \kappa (\mathbf{p}_B - \mathbf{p}_T)^T (\mathbf{p}_B - \mathbf{p}_T) + (1 - \kappa) \mathbf{p}_D^T \mathbf{p}_D \right\}, \quad (7)$$

where  $\kappa$  is the weighting factor and  $\mathbf{p}_T$  the target sound field in the bright zone, which is defined as

$$\mathbf{p}_T = \mathbf{H}_T^T \mathbf{x}, \quad (8)$$

where  $\mathbf{H}_T = [\mathbf{h}_{T,1}, \mathbf{h}_{T,2}, \dots, \mathbf{h}_{T,M_B}]^T$  with  $\mathbf{h}_{T,m}$  being the target impulse responses from a virtual source to the  $m$ -th microphone in the bright zone.

Substituting Eqs. (5), (6) and (8) into Eq. (7) yields

$$J(\mathbf{w}) = \mathbf{w}^T [\kappa \mathbf{Z}_B + (1 - \kappa) \mathbf{Z}_D] \mathbf{w} - \kappa (\mathbf{w}^T \mathbf{y}_B + \mathbf{y}_B^T \mathbf{w}) + \kappa y_T, \quad (9)$$

where  $\mathbf{Z}_B = \mathbb{E} \{ \mathbf{R}_B^T \mathbf{R}_B \}$ ,  $\mathbf{Z}_D = \mathbb{E} \{ \mathbf{R}_D^T \mathbf{R}_D \}$ ,  $\mathbf{y}_B = \mathbb{E} \{ \mathbf{R}_B^T \mathbf{p}_T \}$  and  $y_T = \mathbb{E} \{ \mathbf{p}_T^T \mathbf{p}_T \}$ .

Applying the Stochastic Gradient Descent (SGD) algorithm to Eq. (9), the control filter update equation is derived as

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu [\kappa \tilde{\mathbf{R}}_B^T (\mathbf{p}_B - \mathbf{p}_T) + (1 - \kappa) \tilde{\mathbf{R}}_D^T \mathbf{p}_D], \quad (10)$$

where  $\mu$  is the step-size and  $\tilde{\mathbf{R}}_B$  and  $\tilde{\mathbf{R}}_D$  are the estimated filtered reference signals based on the plant model for the bright and dark zones, respectively, since true plants are usually not available. Online modelling of the plants will be discussed in next section. It is clear that Eq. (10) is a modified multichannel FxLMS algorithm with an extra term  $(1 - \kappa) \tilde{\mathbf{R}}_D^T \mathbf{p}_D$  for sound energy suppression in the dark zone. When  $\kappa = 1$ , Eq. (10) degenerates to the standard multichannel FxLMS algorithm for sound pressure matching in the bright zone without any control for the dark zone; when  $\kappa = 0.5$ , Eq. (10) is equivalent to the standard multichannel FxLMS algorithm for sound pressure matching in both bright and dark zones with zero target sound pressure in the dark zone.

## 2.2 Online Plant Modeling

As mentioned above, the plants between the loudspeakers and microphones are changing in non-stationary environments and accurate plants are not available for control filter update in Eq. (10); therefore, online plant modeling based on the Recursive Least-Squares (RLS) algorithm is utilized. The block diagram of the system is illustrated in Figure 2, where the FxLMS algorithm (green blocks) is used to update the control filters while the two RLS algorithms are employed to model the plants from the loudspeakers to the bright (blue blocks) and dark (yellow blocks) zones, respectively.

In the RLS algorithms, the control filter outputs are used as the inputs while the differences between the microphone measurements and the modelled results are used as the error signals. The cost function for the plant model from the loudspeakers to the  $m$ -th microphone in the bright zone is defined as (12)

$$\xi_m[n] = \sum_{i=1}^n \lambda^{n-i} \left| p_{B,m}[n] - \tilde{\mathbf{h}}_{B,m}^T \mathbf{u}[n] \right|^2 + \delta \lambda^n \left\| \tilde{\mathbf{h}}_{B,m} \right\|^2, \quad (11)$$

where  $\lambda$  and  $\delta$  are the forgetting factor and regularization parameter, respectively,  $\tilde{\mathbf{h}}_{B,m} = [\tilde{\mathbf{h}}_{B,m1}^T, \dots, \tilde{\mathbf{h}}_{B,ml}^T, \dots, \tilde{\mathbf{h}}_{B,mL}^T]^T$ , and  $\mathbf{u}[n] = [\mathbf{u}_1^T[n], \dots, \mathbf{u}_l^T[n], \dots, \mathbf{u}_L^T[n]]^T$ . Following the RLS derivation process (12) and collocating the update equations for different microphones, the bright zone plant model is updated via

$$\tilde{\mathbf{H}}_B[n] = \tilde{\mathbf{H}}_B[n-1] + \mathbf{k}[n] \xi_B^H[n-1], \quad (12)$$

where  $\tilde{\mathbf{H}}_B = [\tilde{\mathbf{h}}_{B,1}, \dots, \tilde{\mathbf{h}}_{B,m}, \dots, \tilde{\mathbf{h}}_{B,M_B}]$ ,  $\xi_B[n] = \mathbf{p}_B[n] - \tilde{\mathbf{H}}_B[n-1] \mathbf{u}[n]$ , and

$$\mathbf{k}[n] = \frac{\lambda^{-1} \mathbf{Q}[n-1] \mathbf{u}[n]}{1 + \lambda^{-1} \mathbf{Q}[n-1] \mathbf{u}[n]} \quad (13)$$

with the matrix  $\mathbf{Q}$  being updated as

$$\mathbf{Q}[n] = \lambda^{-1} \mathbf{Q}[n-1] - \lambda^{-1} \mathbf{k}[n] \mathbf{u}^H[n] \mathbf{Q}[n-1]. \quad (14)$$

It has been shown in (13) that, as long as the control filter coefficients  $\mathbf{w}$  are not all zeros,  $\tilde{\mathbf{H}}_B$  will converge to the true plant  $\mathbf{H}_B$ . A similar procedure can be followed to update the plant model for the dark zone, i.e.,

$$\tilde{\mathbf{H}}_D[n] = \tilde{\mathbf{H}}_D[n-1] + \mathbf{k}[n] \xi_D^H[n-1], \quad (15)$$

where  $\xi_D[n] = \mathbf{p}_D[n] - \tilde{\mathbf{H}}_D[n-1] \mathbf{u}[n]$ . It is noted that the control filter output  $\mathbf{u}[n]$ , the matrix  $\mathbf{Q}[n]$ , and the intermediate vector  $\mathbf{k}[n]$  are independent of the microphone positions and hence are the same for both the bright and dark zones.

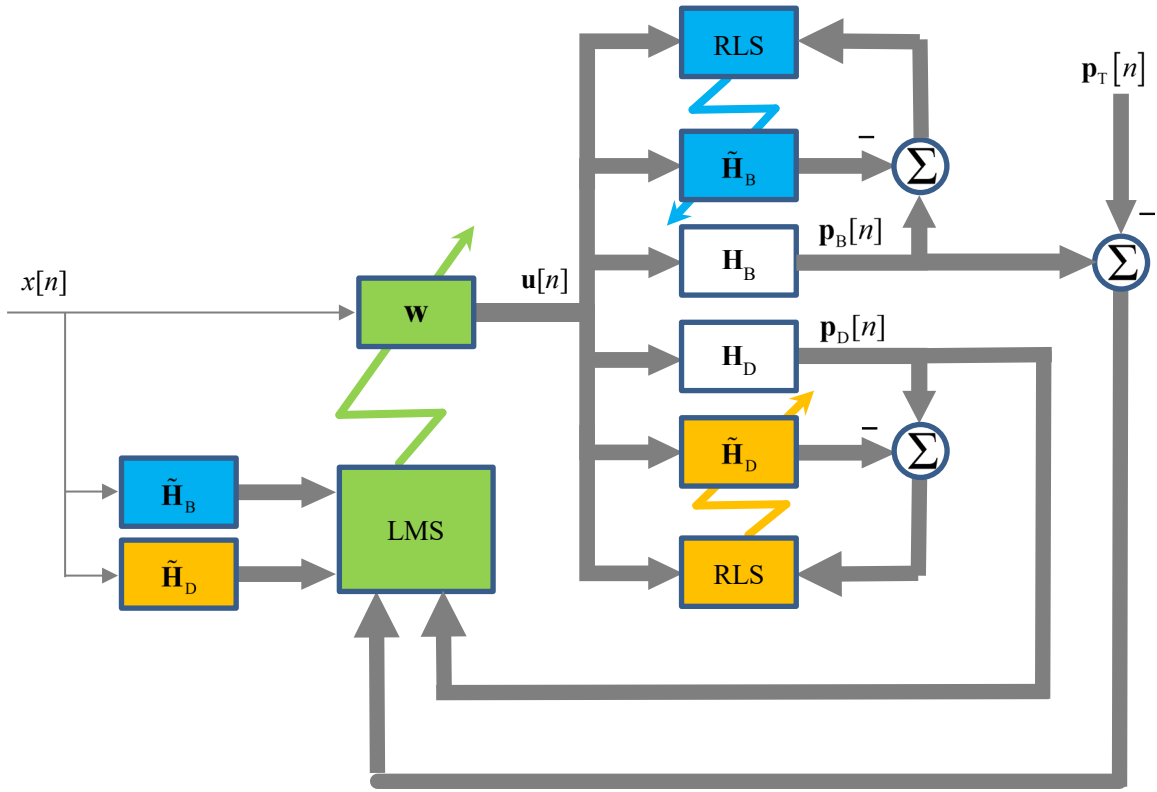


Figure 2 – Block diagram of the adaptive weighted pressure matching method

The complete algorithm for adaptive personal sound zone systems with online plant modeling is illustrated in Figure 2 and summarized in Table 1.

Table 1 – Algorithm for adaptive personal sound zone systems with online plant modeling

Initialization		$\mathbf{w}_l[0] = [1, 0, \dots, 0]^T$ for $l = 1, 2, \dots, L$ ;
		$\tilde{\mathbf{H}}_B[0] = \mathbf{0}$ ; $\tilde{\mathbf{H}}_D[0] = \mathbf{0}$ ; $\mathbf{Q}[0] = \delta^{-1}\mathbf{I}$ .
Control filter update	Step 1	Update the control filters based on Eq. (10);
	Step 2	Update $\mathbf{k}[n]$ according to Eq. (13);
Online plant modeling	Step 3	Update bright zone plant model based on Eq. (12);
	Step 4	Update dark zone plant model based on Eq. (15);
	Step 5	Update the matrix $\mathbf{Q}$ according to Eq. (14);
	Step 6	Repeat Steps 1 – 5.

### 3. SIMULATIONS

#### 3.1 System Setup

Simulations with measured room impulse responses were performed to validate the proposed algorithm for both fixed and time-varying plants. The system setup is shown in Figure 3, where ten loudspeakers are placed along an arc with a radius of 1.5 m to create two listening zones that are separated by 0.8 m. The interval distance between loudspeakers is 0.15 m. Eight microphones are used in both the bright and dark zones. The microphones are moved from Position 1 to Position 2 to investigate the tracking performance of the algorithm for non-stationary plants.

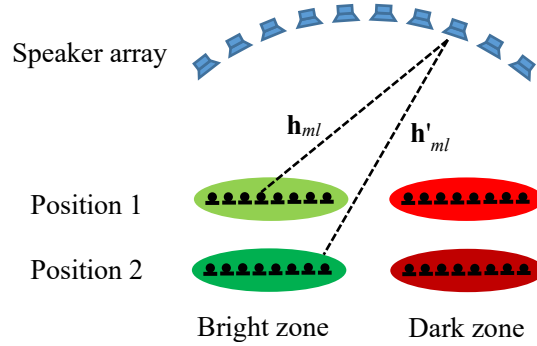


Figure 3 – Simulation system setup.

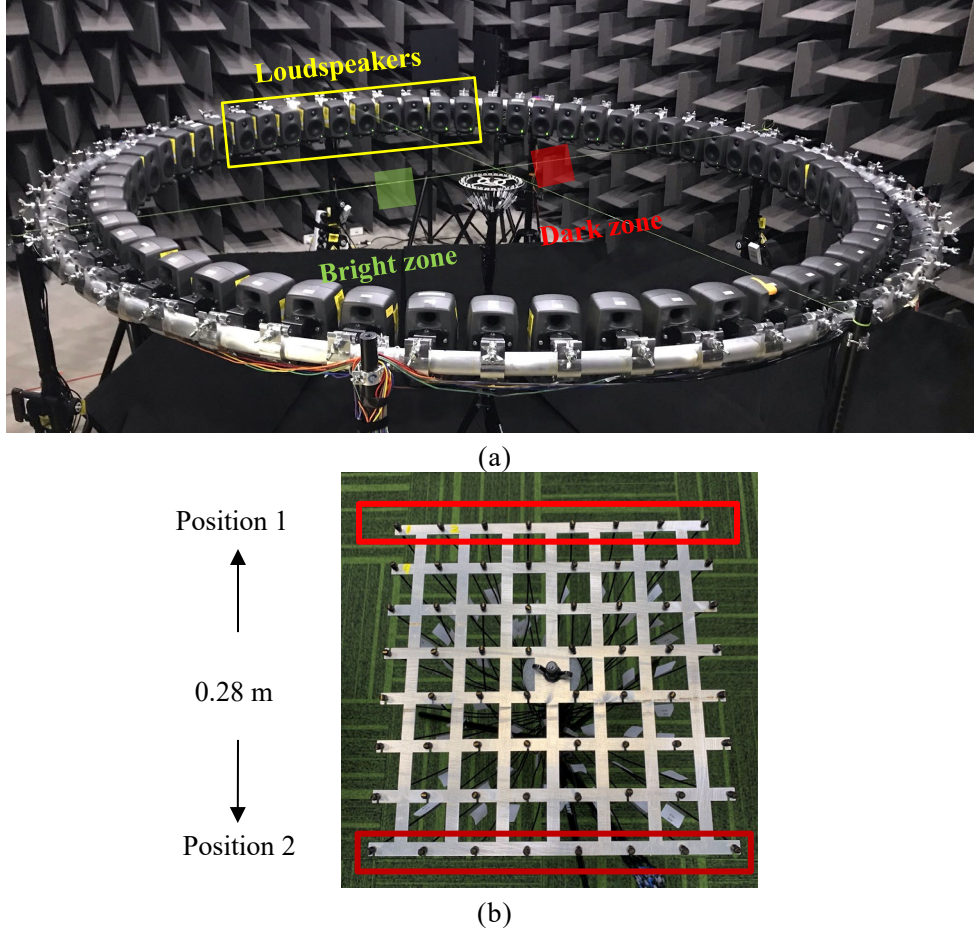


Figure 4 – Photos of (a) the loudspeaker array and (b) the microphone array (not to scale).

The impulse responses were measured in the hemi-anechoic chamber at UTS Tech Lab, with sound absorbent materials on the ground, as depicted in Figure 4. In the experiment, sixty loudspeakers were uniformly placed along a circular truss (Figure 4(a)) and an  $8 \times 8$  array of 64 microphones (Figure 4(b)) were used in both the bright (green square in Figure 4(a)) and dark (red square in Figure 4(a)) zones to measure the impulse responses. However, only the impulse responses from ten loudspeakers that are within the yellow rectangles in Figure 4(a) to eight microphones that are within the red rectangles in Figure 4(b) were used in the simulations to reduce computational burden. The impulse responses were measured at a sampling rate of 48 kHz but were down sampled to 4 kHz for the simulations in this paper.

In the simulations, the target impulse responses were a delayed version of the impulse responses from the center loudspeaker (the 5-th loudspeaker from left in the yellow rectangle in Figure 4(a)) to the bright zone. The filter length is 128 for all the impulse responses, plant models and control filters in the simulations. The step size in the FxLMS algorithm for control filter update is  $\mu = 0.001$ . The forgetting factor and regularization parameter in the RLS algorithm for the online plant modeling are  $\lambda = 0.99995$  and  $\delta = 10^{-5}$ , respectively. The input signals are bandpass (100 Hz – 1 kHz) filtered white Gaussian noise.

### 3.2 Evaluation metrics

The performance of the algorithm is evaluated based on two metrics, i.e., the Mean Square Error (MSE) in the bright zone and the Acoustic Contrast (AC) between the bright and dark zones, which are defined as

$$MSE = 10 \log_{10} \left( \frac{\|\mathbf{p}_B - \mathbf{p}_T\|^2}{\|\mathbf{p}_T\|^2} \right) \quad (16)$$

and

$$AC = 10 \log_{10} \left( \frac{\|\mathbf{p}_B\|^2}{\|\mathbf{p}_D\|^2} \right), \quad (17)$$

respectively.

### 3.3 Simulation Results

Simulations for the stationary case, with the bright and dark zones fixed at Position 1, are performed first, and the results for ideal plant model and OPM are compared in Figure 5. Figure 5(a) and (b) show that, for the ideal plant model (accurate plants are used for control filter update), the MSE decreases and the AC increases with time as the algorithm converges, consistent with expectations. In addition, as the weight factor  $\kappa$  increases, both the MSE and AC decrease, meaning a better sound quality in the bright zone but lower sound energy suppression in the dark zone. This agrees with the designed cost function in Eq. (9). At a time of 100 s, the MSEs are approximately -16 dB ( $\kappa = 0.2$ ), -19 dB ( $\kappa = 0.5$ ) and -20.5 dB ( $\kappa = 0.8$ ), and the ACs are 18 dB ( $\kappa = 0.2$ ), 16.5 dB ( $\kappa = 0.5$ ) and 13 dB ( $\kappa = 0.8$ ), respectively. Similar results can be observed in Figure 5(c) and (d) for the online plant modeling algorithm. It is clear that, without any *a priori* knowledge of the plant, the algorithm converges to the optimal solution although the convergence is slow over the first few seconds.

The results in Figure 5 demonstrate the convergence of the proposed algorithm in stationary environments. To investigate the performance in non-stationary environments, the microphones were placed at Position 1 initially, and then moved to Position 2 at a time point of 30 s. This is similar to head movement or seat adjustment in car cabins or home theatres. The results for  $\kappa = 0.5$  without and with OPM are shown for comparison in Figure 6. It is clear that the algorithm without OPM diverges, with the MSE approaching infinity and the AC becoming around 0. In contrast, with OPM, the personal sound system tracks the changes in the plant and adapts to the new environment with satisfactory performance in terms of both MSE and AC. Similar results are obtained for other values of  $\kappa$  but are not shown here for brevity.

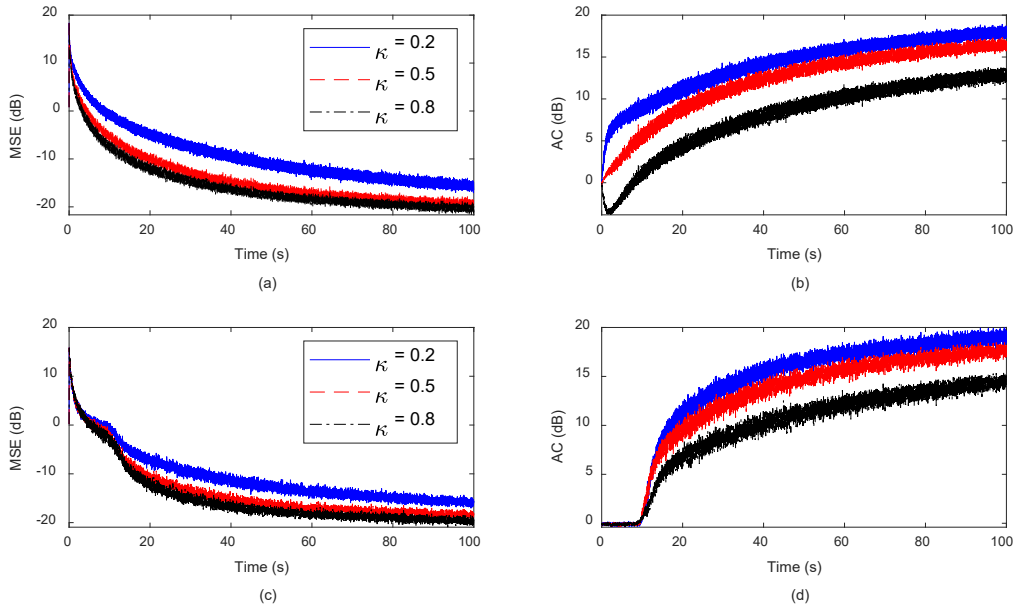


Figure 5 – (a) MSE and (b) AC for ideal plant model; (c) MSE and (d) AC for OPM.

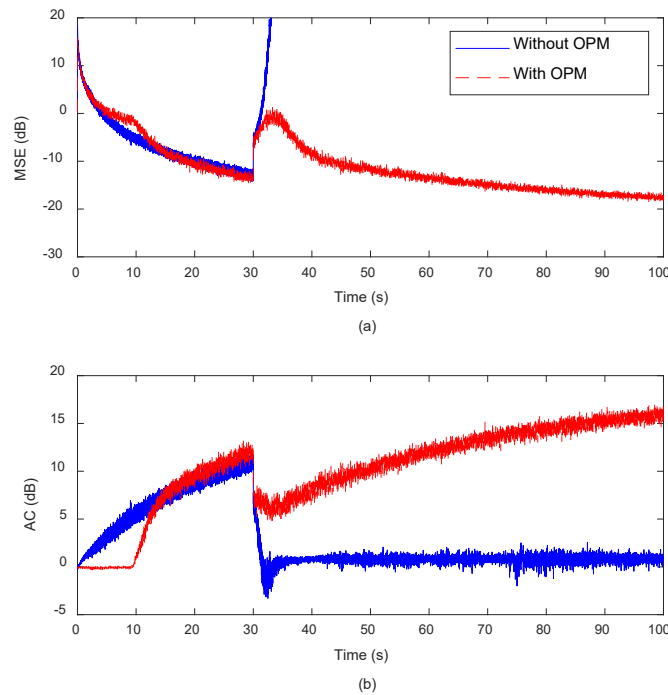


Figure 6 – Comparison of (a) MSE and (b) AC for the adaptive personal sound systems without and with OPM when  $\kappa = 0.5$ .

The above results demonstrate the efficacy of the proposed adaptive personal sound system with online plant modeling for non-stationary environments. However, several challenges need to be tackled before the systems can be deployed for practical applications. Firstly, in practice, it may be infeasible to place the monitoring microphones near the listeners' ears, hence remote sensing (14) or virtual sensing (15) techniques may be required for sound acquisition. Secondly, the computational burden of the multichannel FxLMS and RLS algorithms are too high for real time processing; therefore, computationally efficient adaptive algorithms are desirable for practical applications. Finally, algorithms with faster convergence properties are needed to reduce audible interference when the acoustic environments change. These issues will be further investigated in the future.

#### 4. CONCLUSIONS

This paper presents an adaptive personal sound system with online plant modeling for non-stationary environments, where the FxLMS algorithm is used for control filter update and the RLS algorithm for plant model adaptation. Simulation results with measured impulse responses for a ten-channel system demonstrate that the proposed system works well in both stationary and non-stationary environments. Future work will investigate computationally efficient adaptive algorithms with faster convergence speed for multichannel personal sound systems and employ remote sensing and head-tracking techniques for listener position tracking.

#### ACKNOWLEDGEMENTS

Computational facilities were provided by the UTS eResearch High Performance Compute Facilities.

#### REFERENCES

1. Druyvesteyn WF, Garas J. Personal sound. *J Audio Eng Soc.* 1997;45(9):685–701.
2. Choi J-W, Kim Y-H. Generation of an acoustically bright zone with an illuminated region using multiple sources. *J Acoust Soc Am.* 2002;111(4):1695–700.
3. Kirkeby O, Nelson PA. Reproduction of plane wave sound fields. *J Acoust Soc Am.* 1993;94(5):2992–3000.



4. Wu YJ, Abhayapala TD. Spatial multizone soundfield reproduction: Theory and design. *IEEE Trans Audio, Speech Lang Process.* 2011;19(6):1711–20.
5. Lee T, Nielsen JK, Jensen JR, Christensen MG. A unified approach to generating sound zones using variable span linear filters. In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings.* 2018. p. 491–5.
6. Zhao S, Burnett IS. Evolutionary array optimization for multizone sound field reproduction. *J Acoust Soc Am.* 2022;151(4):2791–801.
7. Radmanesh N, Burnett IS. Generation of isolated wideband sound fields using a combined two-stage lasso-LS algorithm. *IEEE Trans Audio, Speech Lang Process.* 2013;21(2):378–87.
8. Vindrola L, Melon M, Chamard J, Gazengel B, Vindrola L, Melon M, et al. Use of the filtered-x least-mean-squares algorithm to adapt personal sound zones in a car cabin. *J Acoust Soc Am.* 2021;150(3):1779–93.
9. Simon Galvez MF, Elliott SJ, Cheer J. Time Domain Optimization of Filters Used in a Loudspeaker Array for Personal Audio. *IEEE/ACM Trans Audio Speech Lang Process.* 2015;23(11):1869–78.
10. Akhtar MT, Abe M, Kawamata M, Nishihara A. Online secondary path modeling in multichannel active noise control systems using variable step size. *Signal Processing.* 2008;88(8):2019–29.
11. Hu M, Xue J, Lu J. Online multi-channel secondary path modeling in active noise control without auxiliary noise. *J Acoust Soc Am.* 2019;146(4):2590–5.
12. Haykin S. *Adaptive Filter Theory.* Fifth Edit. Pearson; 2014. 913 p.
13. Hu M, Xue J, Lu J. Online multi-channel secondary path modeling in active noise control without auxiliary noise. *J Acoust Soc Am.* 2019;146(4):2590–5.
14. Xiao T, Zhao S, Qiu X, Halkon B. Using a retro-reflective membrane and laser doppler vibrometer for real-time remote acoustic sensing and control. *Sensors.* 2021;21(11).
15. Moreau D, Cazzolato B, Zander A, Petersen C. A review of virtual sensing algorithms for active noise control. *Algorithms.* 2008;1(2):69–99.