IDENTIFICATION OF THE PARAMETRIC ARRAY LOUDSPEAKER WITH A VOLTERRA FILTER USING THE SPARSE NLMS ALGORITHM

Chuang Shi and Yoshinobu Kajikawa

Department of Electrical and Electronic Engineering, Kansai University

ABSTRACT

Volterra filters can be applied to a wide range of nonlinear systems, keeping only the low order kernels to yield a good approximation. The parametric array loudspeaker (PAL), as a weak nonlinear acoustic system, is an attractive directional sound reproduction device. Volterra filters have been adopted in the linearization system of the PAL that efficiently reduces the nonlinear distortion with no need of solving the nonlinear acoustic equation. In this paper, the ultrasound-to-ultrasound Volterra filter is proposed, being inspired by the nonlinear acoustic principle, to provide a better systematic representation of the PAL. Experiment results are presented to prove the effectiveness of the proposed approach, where the sparse NLMS algorithm is carried out in the identification.

Index Terms— Parametric array loudspeaker, Volterra filter, sparse NLMS algorithm

1. INTRODUCTION

When two finite amplitude waves are transmitted at close frequencies in a collimated beam, the difference of the two frequencies form a similarly narrow beam. This nonlinear acoustic phenomena was derived by Westervelt and named as the parametric acoustic array [1]. The parametric acoustic array was mainly applied in underwater applications. The first directional sound reproduction device making use of the parametric acoustic array in air was invented in 1983, which is now known as the parametric array loudspeaker (PAL) [2]. The principle of the PAL is commonly explained by the Berktay's far-field solution, although the assumptions made by Berktay are not well validated in air [3].

The sound quality of the PAL is not satisfactory owing to its nonlinear distortion, which is an adverse byproduct of the parametric acoustic array in air. For this reason, there have been many preprocessing methods to suppress the nonlinear distortion of the PAL [4–7]. The audio bandwidth extension has been attempted to improve the perceptual sound quality of the PAL [8, 9]. But all these preprocessing methods have been derived on the basis of the Berktay's far-field solution.

This work is supported by MEXT-Supported Program for the Strategic Research Foundation at Private University, 2013-2017.

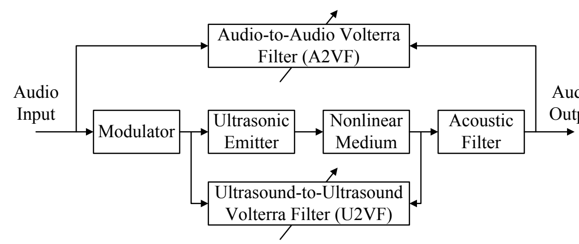

Fig. 1. Block diagram of the parametric array loudspeaker and the Volterra filter identification.

Their performance is usually limited in practice, and thus, an adaptive preprocessing method is much preferred.

In spite of the nonlinear distortion, the PAL is readily deployed in sound field control applications, such as active noise control [10–12], creation of personal listening zone [13], and spatial audio reproduction [14–16]. The precious tailoring of the sound field requires the accurate systematic representation of the PAL. However, the second order nonlinear acoustic equation has no analytical solution, and the numerical solutions add huge computational burdens.

The application of the Volterra filter in the systematic representation of the PAL has been studied since 2002 [17]. After the Volterra filter is identified, the inverse filter can be designed to preprocess the audio input in order to reduce the nonlinear distortion of the PAL efficiently [18–22]. This technique was adapted from its original application in the linearization of the conventional loudspeaker system. In the past studies, both the input and output of the Volterra filter identification were selected by default as the audio input and output of the PAL. In this paper, the present approach is illustrated in Fig. 1 as the audio-to-audio Volterra filter (A2VF) identification. The identified A2VF of the PAL remains accurate if the modulation method and the power of the input audio are identical to those used in the identification.

In this paper, another approach, namely the ultrasound-to-ultrasound Volterra filter (U2VF) identification is proposed to account for the nonlinear acoustic principle of the PAL. The U2VF identification is verified through experiments. The difficulty of the U2VF implementation is a result of its relatively high sampling frequency [23]. Hence, the sparse LMS is adopted to identify the first order U2VF coefficients with a long memory length, and the acoustic delay between the PAL and the microphone can be extracted from the identified first order U2VF. Afterwards, the memory length of the second order U2VF is shortened by realigning the input and output. The identified second order U2VF is used to predict the total harmonic distortion of the PAL in comparison with the measurement results as well as the model results of the A2VF.

2. PARAMETRIC ARRAY LOUDSPEAKER

Among the nonlinear acoustic models, the Berktays far-field solution is the most widely used for developing the preprocessing method of the PAL, although the assumptions made by Berktay are not consistently validated in air. The Berktay's far-field solution is an audio-to-audio model. The audio output of the PAL is expressed as

$$p_d = K\frac{\partial^2}{\partial t^2}E^2\left(t\right). \tag{1}$$

where K is a joint parameter related to the ultrasonic emitter, observation position, and acoustic properties of air; and $E\left(t\right)$ is the preprocessed audio input, *a.k.a.* the envelope function.

The modulated audio input of the first PAL is given by

$$E_{DSB}\left(t\right)\cos\left(\omega_c t\right) = \left[1 + mA\left(t\right)\right]\cos\left(\omega_c t\right), \tag{2}$$

where m is the modulation index; $A\left(t\right)$ is the audio input; and ω_c is the angular frequency of the carrier [2]. This in effect is the double sideband (DSB) modulation method. Based on (1), the second harmonic distortion of the PAL using the DSB method is found to be proportional to the modulation index.

Therefore, the square root (SRT) method was introduced to offset the square operation in (1) [4]. The modulated audio input of the SRT method is given by

$$E_{SRT}\left(t\right)\cos\left(\omega_c t\right) = \sqrt{1 + mA\left(t\right)}\cos\left(\omega_c t\right). \tag{3}$$

The drawback of the SRT method was mostly encountered in the implementation, since the ultrasonic emitter was required to have an infinite bandwidth. To overcome this drawback, the single sideband (SSB) modulation method was suggested. The modulated audio input of the SSB method is given by

$$\begin{aligned} E_{SSB}\left(t\right)\cos\left(\omega_c t + \phi\right) = &\left[1 + mA\left(t\right)\right]\cos\left(\omega_c t\right) \\ &\mp mA_H\left(t\right)\sin\left(\omega_c t\right), \end{aligned} \tag{4}$$

where $A_H\left(t\right)$ is the Hilbert transformed audio input;

$$E_{SSB}\left(t\right) = \sqrt{2 + 2mA\left(t\right)}; \tag{5}$$

and

$$\phi = \tan^{-1}\left[\frac{1 + mA\left(t\right)}{\mp mA_H\left(t\right)}\right]. \tag{6}$$

There are two types of the SSB method. They are the upper and lower SSB methods, which depends on the sign of the orthogonal carrier. The SSB method contains an angular modulation term. Thus, it is not a pure amplitude modulation. The amplitude modulation term of the SSB method results in the same envelope function as the SRT method. Based on (1), both the SRT and SSB methods are expected to eliminate the harmonic distortion of the PAL when the Berktay's far-field solution is valid and the ultrasonic emitter is perfect [7]. The modulated audio inputs of the DSB, SRT, and SSB methods are shown in Fig. 2, when the audio input is a sine tone at ω_0.

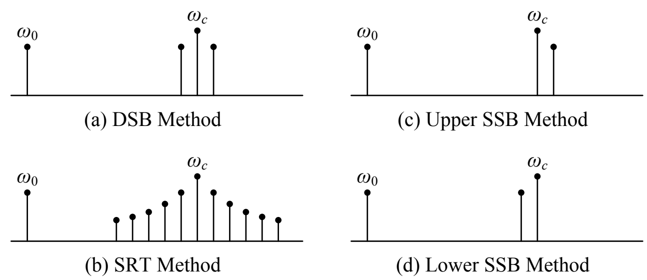

Fig. 2. Spectrum examples of the modulated audio inputs of different preprocessing methods.

3. VOLTERRA FILTER AND NLMS ALGORITHMS

The second order nonlinear system can be expressed by the Volterra filter as

$$y\left(n\right) = H_1\left[x\left(n\right)\right] + H_2\left[x\left(n\right)\right] + e\left(n\right), \tag{7}$$

where $x\left(n\right)$ and $y\left(n\right)$ are the input and output of the system; $e\left(n\right)$ is the model error; H_1 and H_2 are the first and second order Volterra operators, *i.e.*

$$H_1\left[x\left(n\right)\right] = \sum_{i=0}^{N-1} h_1\left(i\right)x\left(n-i\right) \tag{8}$$

and

$$H_2\left[x\left(n\right)\right] = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} h_2\left(i,j\right)x\left(n-i\right)x\left(n-j\right) \tag{9}$$

respectively. In (8) and (9), N is the memory length; h_1 and h_2 are the first and second order Volterra filter coefficients.

The input and output of the system can be continuously acquired and stored in the vector form, of which the size is N by 1. Furthermore, the function *"reshape"* can arrange all the elements of an N by N matrix to an N^2 by 1 vector. The second order nonlinear system in (7) can be rewritten into a linear model as

$$y\left(n\right) = X^T H + e\left(n\right), \tag{10}$$

where

$$X = \left[x, reshape\left(xx^T, N^2, 1\right)\right]^T, \tag{11}$$

and

$$H = \left[h_1, reshape\left(h_2, N^2, 1\right)\right]^T, \quad (12)$$

Both X and H have the size of $N^2 + N$ by 1. Using (10), the NLMS algorithm can be adopted to identify H. The update equation of the NLMS algorithm is given by

$$H_{k+1} = H_k + \alpha \frac{X_k e(k)}{X_k^T X_k}, \quad (13)$$

where α is the step size.

Modified from the reweighted zero-attracting LMS algorithm [24], the update equation of the sparse NLMS algorithm is written as

$$H_{k+1} = H_k + \frac{1}{X_k^T X_k}\left[\alpha X_k e(k) - \beta \frac{\operatorname{sgn}(H_k)}{1 + |H_k|/\beta}\right], \quad (14)$$

where β is the small amplitude that coefficients lower than β will shrink to 0 to create the sparsity.

4. EXPERIMENT AND RESULTS

The experiment is carried out in a meeting room, where the floor noise and reflections are similar to the daily usage of the PAL. The experiment setup is shown in Fig. 3. A white noise signal is generated in the computer and processed by a digital band-pass filter, of which the cut-off frequencies are chosen at 20 kHz and 60 kHz. This band-passed white noise is transmitted by the ultrasonic emitter. A microphone (B&K Type 4191L) is placed 3.2 meters away from the ultrasonic emitter to record the sound reproduced by the PAL setup. Both the DAC and ADC have the sampling frequency of 192 kHz and the resolution of 32 bit.

4.1. First order Volterra filter identification

The sound speed in dry air at 20°C is estimated to be 342 m/s. Thus, there is an acoustic delay between the ultrasonic emitter and the microphone, which is about 9.375 ms and 1800 samples. The causality of the Volterra filter is not satisfied if the memory length is set shorter than the actual acoustic delay. Moreover, the computational complexity of the Volterra filter increases exponentially with the memory length. To use a memory length longer than the estimated acoustic delay of 1800 samples is impractical. It is of significant importance to extract the acoustic delay, in order that the input and output of the identification can be realigned accordingly and the memory length of the Volterra filter can be shortened drastically.

For this purpose, the first order U2VF, which is a linear filter, is first identified using the NLMS and sparse NLMS algorithms. Due to the nonlinear acoustic principle of the PAL, the ultrasound pressure captured by the microphone can be 60 dB higher than the audio sound pressure. The identification of the PAL as the first order U2VF can yield a very good accuracy. However, it is useless for the design of the linearization

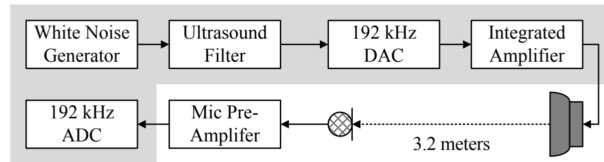

Fig. 3. Experiment setup.

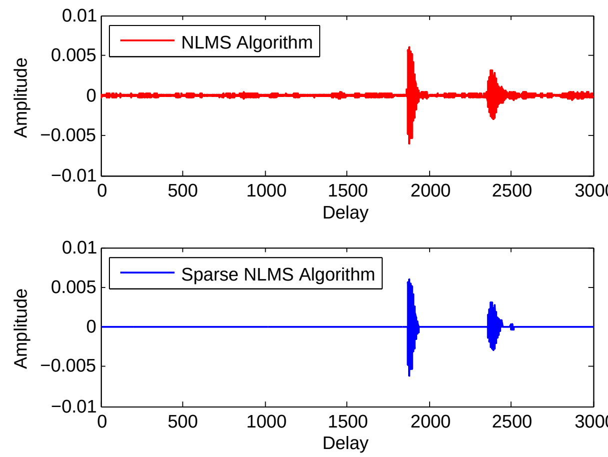

Fig. 4. The first order U2VF coefficients identified by the NLMS and sparse NLMS algorithms.

system, as the first order U2VF does not generate any audible sound from the ultrasound. In this paper, the significance of the linear U2VF is to find out the acoustic delay only.

The identified Volterra filter coefficients of the NLMS and sparse NLMS algorithms are shown in Fig. 4, when the memory length is temporarily set to 3000. The step size and small amplitude are selected as $\alpha = 0.01$ and $\beta = 1e-5$, respectively. It is observed in Fig. 4 that the sparse NLMS algorithm fast converges to the satisfactory situation, where the acoustic delay is easily distinguishable and read as 1865 samples. The microphone output will be realigned by this acoustic delay in the second order U2VF identification.

4.2. Second order Volterra filter identification

The memory length of the first and second order Volterra filters is reset to 600. The step size and small amplitude are kept as $\alpha = 0.01$ and $\beta = 1e-5$, respectively. The A2VF is identified by the NLMS algorithm. The U2VF is identified by the NLMS and sparse NLMS algorithms. The identified coefficients of the first order U2VF after realignment are plotted in Fig. 5. In comparison with Fig. 4, the difference between the coefficients before and after realignment is due to the involvement of the second order U2VF. Moreover, the sparsity in the first order U2VF remains validated in Fig. 5.

The identified A2VF, U2VF, and sparse U2VF can be evaluated by the total harmonic distortion (THD). The THD

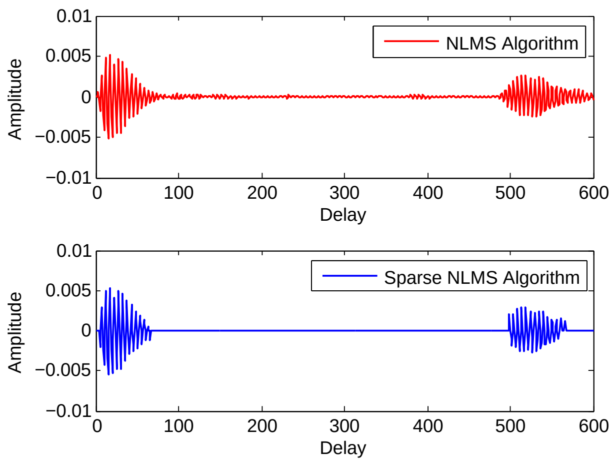

Fig. 5. The first order U2VF coefficients identified by the NLMS and sparse NLMS algorithms after realignment.

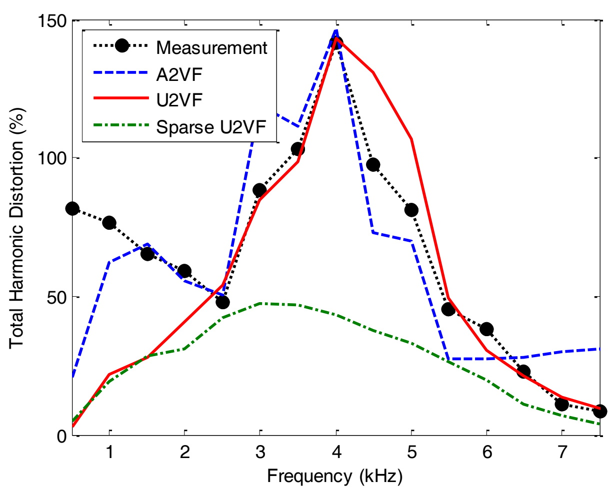

Fig. 6. THD values of the DSB method.

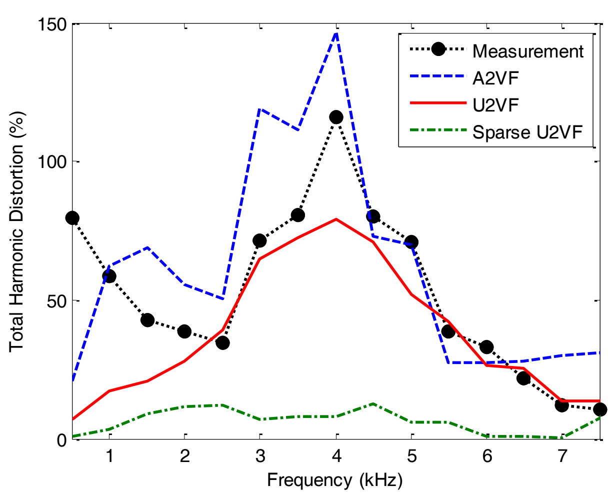

Fig. 7. THD values of the SRT method.

used in this paper is defined as the root mean square ratio of the harmonics to the fundamental frequency. A testing audio input is generated as a sine sweeps ranging from 500 Hz to 7500 Hz. The THD values are extracted from the outputs of the identified Volterra filters when the input is provided by the testing audio input. The THD values are also measured by the same experiment setup in Fig. 3, but the input is replaced by the sine sweep.

The THD values of the DSB and SRT methods are plotted in Figs. 6 and 7, respectively. In these two figures, the A2VF produces the same curve. This is because that the A2VF has to be identified for every different preprocessing method, but in this paper, the A2VF is only identified for the DSB method. In Fig. 6, the A2VF provides accurate predictions to the THD values of the DSB method from 1500 Hz to 6500 Hz. But it cannot match the measured THD of the SRT method as shown in Fig. 7. The limitation of the A2VF is clearly demonstrated.

The U2VF shows the consistent accuracy in Figs. 6 and 7, when it is used to predict the THD values of the DSB and SRT methods. The drawback of the U2VF is observed in the frequency band lower than 2000 Hz. This is probably because of the background noise of the room, where the measurement has been carried out. The measured THD values may include some errors. Another obvious discrepancy occurs in the THD plot of the SRT method, when the input frequency is 4000 Hz. Except for this frequency, the U2VF provides accurate predictions to the THD values of both the DSB and SRT methods from 2500 Hz to 7500 Hz.

The sparse U2VF exhibits the poorest performance in this comparison. Therefore, it is concluded that the second order Volterra filter contains very few sparsity when it is applied in the systematic representation of the PAL. Nevertheless, the symmetry in the second order Volterra filter coefficients is still effective to reduce the computational complexity.

5. CONCLUSIONS

In this paper, the second order Volterra filters are studied for the systematic presentation of the PAL. The acoustic delay between the PAL and the microphone is much longer in the U2VF than in the A2VF, because a higher sampling frequency is used to cover the ultrasonic frequency range. Hence, it is of significant importance to extract the acoustic delay, in order that the realignment of the input and output can be carried out and the memory length of the U2VF can be reduced. The sparse NLMS algorithm has been effectively applied to identify the first order Volterra filter and reveal the acoustic delay. Furthermore, the identified U2VF and A2VF are compared by experiments. The U2VF outperforms the A2VF when they are used to predict the THD values of the DSB and SRT methods. This agrees with the authors' previous conclusions based on the simulation results that the modulated audio input in the ultrasound frequency range is a better identification input for the PAL [23].

6. REFERENCES

[1] P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Amer.*, vol. 35, no. 4, pp. 535-537, 1963.

[2] M. Yoneyama, J. Fujimoto, Y. Kawamo, and S. Sasabe, "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design," *J. Acoust. Soc. Amer.*, vol. 73, no. 5, pp. 1013-1020, 1983.

[3] H. O. Berktay, "Possible exploitation of nonlinear acoustics in underwater transmitting applications," *J. Sound Vib.*, vol. 2, no. 4, pp. 435-461, 1965.

[4] T. Kamakura, M. Yoneyama, and K. Ikegaya, "Developments of parametric loudspeaker for practical use," in *Proc. 10th Int. Symp. Nonlinear Acoust.*, Kobe, Japan, 1984.

[5] K. Aoki, T. Kamakura, and Y. Kumamoto, "Parametric loudspeaker: Characteristics of acoustic field and suitable modulation of carrier ultrasound," *Electron. Commun. Jpn.*, vol. 74, no. 9, pp. 7682, 1991.

[6] F. J. Pompei, "The use of airborne ultrasonics for generating audible sound beams," *J. Audio Eng. Soc.*, vol. 47, no. 9, pp. 726-731, 1999.

[7] C. Shi and Y. Kajikawa, "A comparative study of preprocessing methods in the parametric loudspeaker," in *Proc. 2014 APSIPA Annu. Summit Conf.*, Siem Reap, Cambodia, 2014.

[8] C. Shi, H. Mu, and W. S. Gan, "A psychoacoustical preprocessing technique for virtual bass enhancement of the parametric loudspeaker," in *Proc. 38th Int. Conf. Acoust. Speech Signal Process.*, Vancouver, Canada, 2013.

[9] C. Shi and W. S. Gan, "A preprocessing method to increase high frequency response of a parametric loudspeaker," in *Proc. 2013 APSIPA Annu. Summit Conf.*, Kaohsiung, Taiwan, 2013.

[10] N. Tanaka and M. Tanaka, "Active noise control using a steerable parametric array loudspeaker," *J. Acoust. Soc. Am.*, vol. 127, no. 6, pp. 3526-3537, 2010.

[11] B. Lam, W. S. Gan, and C. Shi, "Feasibility of a length-limited parametric source for active noise control applications," in *Proc. 21st Int. Congr. Sound Vib.*, Beijing, China, 2014.

[12] K. Tanaka, C. Shi, and Y. Kajikawa, "Study on active noise control system using parametric array loudspeakers," in *Proc. 7th Forum Acusticum*, Krakow, Poland, 2014.

[13] A. Wright, A. Evans, A. Linney, and M. Lincoln, "The listening room: A speech-based interactive art installation," in *Proc. 15th Int. Conf. Multimedia*, Augsburg, Germany, 2007.

[14] T. Masayoshi and A. Shigeaki, "Stereo reproduction using parametric loudspeakers," in *Proc. 20th Int. Congr. Acoust.*, Sydney, Australia, 2010.

[15] S. Aoki, M. Toba, and N. Tsujita, "Sound localization of stereo reproduction with parametric loudspeakers," *Applied Acoust.*, vol. 73, no. 12, pp. 1289-1295, 2012.

[16] C. Shi, H. Nomura, T. Kamakura, and W. S. Gan, "Spatial aliasing effects in a steerable parametric loudspeaker for stereophonic sound reproduction," *IEICE Trans. Fund. Electron. Commun. Computer Sci.*, vol. E97-A, no.9, pp.1859-1866, 2014.

[17] C. M. Lee, J. Yang, W. S. Gan, and M. H. Er, "Modeling nonlinearity of air with Volterra kernels for use in a parametric array loudspeaker," in *Proc. 112th AES Conv.*, Munich, Germany, 2002.

[18] W. Ji and W. S. Gan, "Identification of a parametric loudspeaker system using an adaptive Volterra filter," *Appl. Acoust.*, vol. 73, no. 12, pp. 1251-1262, 2012.

[19] O. Hoshuyama, K. Yoshino, S. Tateoka, K. Hirose, and A. Oyamada, "Distortion reduction using volterra filter for ultrasonic parametric loudspeaker in nearfield usage," *IEICE Tech. Rep.*, vol. 113, no. 28, pp. 97-102, 2013.

[20] Y. Mu, P. Ji, W. Ji, M. Wu, and J. Yang, "On the linearization of a parametric loudspeaker system by using third-order inverse Volterra filters," in *Proc. 1st Int. Conf. Sig. Inf. Process.*, Beijing, China, 2013.

[21] O. Hoshuyama and T. Okayasu, "Decimation in Volterra filter for distortion reduction of ultrasonic parametric loudspeaker," in *Proc. 2014 IEICE Soc. Conf.*, Tokushima, Japan, 2014.

[22] Y. Hatano, C. Shi, and Y. Kajikawa, "A study on linearization of nonlinear distortions in parametric array loudspeakers," in *Proc. 2014 Int. Workshop Smart Inf. Media Syst.*, Ho Chi Minh City, Vietnam, 2014.

[23] C. Shi and Y. Kajiakawa, "Volterra filters for representing the parametric acoustic array in air," in *Proc. 29th IEICE Sig. Process. Symp.*, Kyoto, Japan, 2014.

[24] Y. Chen, Y. Gu, and A. O. Hero III, "Sparse LMS for system identification," in *Proc. 34th Int. Conf. Acoust. Speech Sig. Process.*, Taipei, Taiwan, 2009.