SYNTHESIS OF VOLTERRA FILTERS FOR THE PARAMETRIC ARRAY LOUDSPEAKER

Chuang Shi and Yoshinobu Kajikawa

Department of Electrical and Electronic Engineering, Kansai University

ABSTRACT

The ultrasound-to-ultrasound Volterra filter (U2VF), which was previously proposed to represent the nonlinear response of the parametric array loudspeaker (PAL), identifies the PAL as a nonlinear system that uses ultrasonic signals as its input. It has been proven that the U2VF is a more generic model as compared to the audio-to-audio Volterra filter (A2VF), when the modulation method is adaptive or the input is time varying. However, there is no explicit solution to a linearization system based on the U2VF. Therefore, this paper proposes a synthesis method to extract A2VFs from the U2VF based on the parallel cascade structure. The mature technique of building a linearization system based on the A2VF can hence be adapted to consort with the U2VF.

***Index Terms*—** Parametric array loudspeaker, Volterra filter, parallel cascade structure, linearization system

1. INTRODUCTION

The finite amplitude wave is a sound wave whose local velocity and amplitude are not negligible as compared to its bulk sound speed and wavelength. The waveform of a finite amplitude wave is distorted cumulatively along its propagation [1]. When a bi-frequency finite amplitude wave is transmitted in a nonlinear medium, the difference frequency is generated and confined in a narrow beam. This phenomena was firstly discovered by Westervelt, and named as the parametric acoustic array [2]. A directional loudspeaker making use of the parametric acoustic array is known as the PAL, in which air is the nonlinear medium [3].

The PAL is advantageous in transmitting a narrow sound beam from a relatively small aperture as compared to the conventional loudspeaker. A typical block diagram of the PAL is provided in Fig. 1. The audio input has to be modulated on an ultrasonic carrier. The modulated signal falls in the ultrasonic frequency band. It is thus referred to as the ultrasound input. The ultrasound input is transmitted by an ultrasonic emitter. The nonlinear acoustic effect will function as a demodulator and restore the audio input in air with some distortion. So far, there have been many modulation methods proposing to suppress the nonlinear distortion of the PAL [4–7]. But the unsatisfactory sound quality still hinders the popularity of the PAL. This is because previous modulation methods assume an ideal ultrasonic emitter and their performance is highly dependent on the accuracy of the Berktay's far-field solution [8].

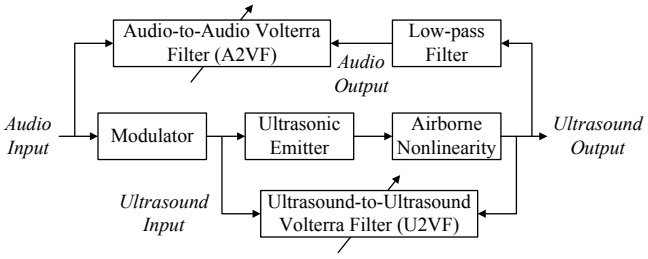


Fig. 1. Block diagram of the PAL and Volterra filter identification structures.

The second-order nonlinear acoustic equation, which is more generic and accurate than the Berktay's far-field solution, has no analytical solution. Thus, Volterra filters are convenient to model the nonlinear response of the PAL [9–11]. Volterra filters are conventionally identified from the audio input to the audio output of the PAL, as shown by the A2VF in Fig. 1. Based on the identified A2VFs, linearization systems are available to compensate for the nonlinear distortion effectively [12]. Moreover, the computational cost of the second-order Volterra kernel can be reduced by implementing few selected diagonals [13] or adopting the parallel cascade structure [14]. A third-order linearization system for the PAL has been examined, but only the main diagonal of the third-order Volterra kernel is implemented to prevent the computational cost from becoming overwhelming [15].

As the linearization system is a well established technique for the conventional loudspeaker [16], it has been overlooked that the nonlinearity of the PAL changes with the input level. The linearization system based on the A2VF identified at one input level is no longer efficient at another input level [17, 18]. Hence, the U2VF has been proposed to be a more generic model than the A2VF, when the modulation method and input level are changing [19, 20]. Moreover, since the second-order nonlinear acoustic effect plays the determining role in the PAL [21], the high-order nonlinearity of the PAL is resultant from the modulation method rather than the nonlinear acoustic effect. The U2VF is thus more reasonable to be truncated at the second-order, as compared to the A2VF. However,

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

there is no available linearization system based on the U2VF. Therefore, this paper proposes a synthesis method to extract the A2VFs from the parallel cascade structure of the U2VF, in order that the linearization systems based on the synthesized A2VFs are able to be built accordingly.

2. BACKGROUND THEORY

2.1. Model Equation

Among the nonlinear acoustic model equations, the Berktay's far-field solution is the most widely applied for developing the modulation methods. The Berktay's far-field solution is as simple as an audio-to-audio model. The audio output of the PAL is expressed as

$$p_d = K\frac{\partial^2}{\partial t^2}e^2(t), \tag{1}$$

where K is a joint parameter related to the ultrasonic emitter, observation position, and acoustic properties of air; and $e(t)$ is the envelope of the modulated signal, which is assumed to have an unit amplitude [8].

The Berktay's far-field solution does not consider the frequency response of the ultrasonic emitter. In other words, the ultrasonic emitter is assumed to have an ideally flat frequency response. However, the ultrasonic emitter of the PAL is made up of a number of piezoelectric transducers (PZTs) that have very different frequency responses in practice [22]. Denoting the impulse responses of the PZTs as $s_i(t)$, the audio output of the PAL is rewritten as

$$p_d = s_a(t) * \frac{K}{M}\frac{\partial^2}{\partial t^2}\sum_{i=1}^{M}\{s_i(t) * [e(t)\cos(\omega_c t)]\}^2, \tag{2}$$

where M is the total number of PZTs; ω_c is the angular frequency of the ultrasonic carrier; and $s_a(t)$ is the impulse response of a low-pass filter to exclude the ultrasonic frequencies.

2.2. Modulation Method

The modulated signal of the double sideband (DSB) modulation method is given by

$$e_{DSB}(t)\cos(\omega_c t) = \left[\frac{1+mg(t)}{1+m}\right]\cos(\omega_c t), \tag{3}$$

where m is the modulation index; and $g(t)$ is the audio input. Substituting (3) into (1) yields

$$p_d = \frac{K}{(1+m)^2}\frac{\partial^2}{\partial t^2}\{[2mg(t) + m^2 g^2(t)]\}. \tag{4}$$

Eq. (4) shows that the second harmonic distortion level of the DSB modulation method is proportional to the modulation index. This observation is also derived by substituting (3) into (2).

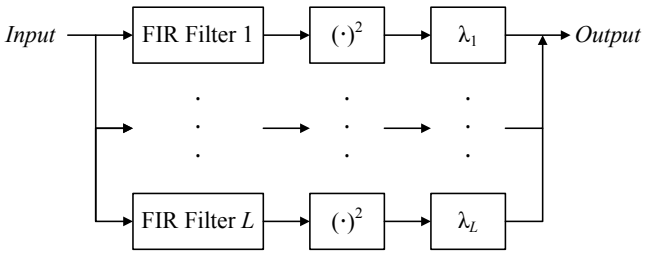


Fig. 2. Parallel cascade structure of the second-order Volterra kernel.

2.3. Volterra Filter Identification

The nonlinear response of the PAL can be expressed by the Volterra filter as

$$y(n) = H_1[x(n)] + H_2[x(n)] + \varepsilon(n), \tag{5}$$

where $x(n)$ and $y(n)$ are the discrete input and output; $\varepsilon(n)$ is the model error, containing the high-order nonlinearity; H_1 and H_2 are the first- and second-order Volterra operators, *i.e.*

$$H_1[x(n)] = \sum_{i=0}^{N-1} h_1(i)\, x(n-i) \tag{6}$$

and

$$H_2[x(n)] = \sum_{i=0}^{N-1}\sum_{l=0}^{N-1} h_2(i,l)\, x(n-i)\, x(n-l), \tag{7}$$

respectively. In (6) and (7), N is the memory length; h_1 and h_2 are the first- and second-order Volterra kernels [23].

The second-order Volterra operator is often rewritten in a matrix equation as

$$H_2[X] = XH_2X^T, \tag{8}$$

where $X := [x(n), \ldots, x(n-N+1)]$ is the row-wise vector of the discrete input; H_2 is the coefficient matrix of the second-order Volterra kernel, given by $H_2(i,l) := h_2(i,l)$.

When H_2 is real-symmetric, it can be decomposed into $V\Lambda V^T$, where $\Lambda := diag(\lambda_1, \ldots, \lambda_N)$ is a diagonal matrix formed from the eigenvalues of H_2 and the columns of V are the corresponding eigenvectors. Without loss of generality, it is assumed that $\lambda_1 > \lambda_2 > \ldots > \lambda_N$.

Eq. (8) is further manipulated as

$$H_2[X] = X\left(V\Lambda V^T\right)X^T = (XV)\,\Lambda\,(XV)^T. \tag{9}$$

This provides the fundamental of the parallel cascade structure. The ith column of V is treated as a finite impulse response (FIR) filter, of which the coefficients are denoted as $v_i(n)$ for convenience. Therefore, each row of XV presents the filtered output of $x(n)$. Eq. (9) is interpreted as a linear combination of the squared outputs of an array of FIR filters, as shown in Fig. 2. The parallel cascade structure helps to reduce the computational cost by implementing the L largest eigenvalues of H_2 only, where L is much smaller than M.

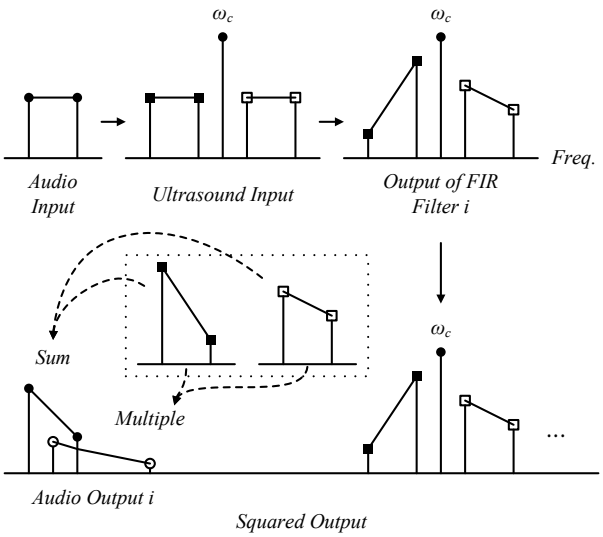
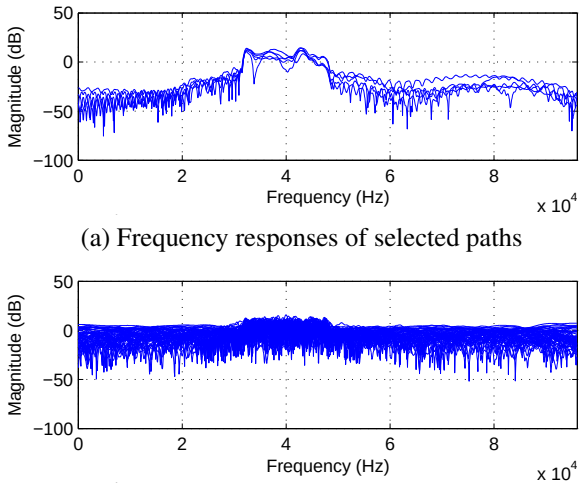


Fig. 3. Illustration of the nonlinear process in the PAL based on the DSB modulation method.



(a) Frequency responses of selected paths

(b) Frequency responses of discarded paths

Fig. 4. Identified second-order kernel of the U2VF ($N = 300$ and $L = 5$).

3. PROPOSED SYNTHESIS METHOD

As the first-order kernel of the U2VF is a linear filter operating on the ultrasound input, it does not contribute to the audio output. Both the first-order and second-order kernels of the A2VF are embedded in the second-order kernel of the U2VF. Hence, the U2VF is analyzed by considering the parallel cascade structure.

When $x_u(n) = e(n)\cos(\Omega_c n)$ gives the ultrasound input, the output of the second-order kernel of the U2VF is written as

$$p_s = \sum_{i=1}^{L} \lambda_i \{v_i(n) * [e(n)\cos(\Omega_c n)]\}^2, \tag{10}$$

and equivalently in the frequency domain as

$$p_s = \sum_{i=1}^{L} \lambda_i \{\mathrm{IFFT}[V_i(\Omega) X(\Omega)]\}^2, \tag{11}$$

where Ω_c is the carrier frequency; $V_i(\Omega)$ is the frequency response of the ith FIR filter; and $X(\Omega)$ is the spectrum of the ultrasound input. The audio output is simply obtained as the low-passed ultrasound output.

Eq. (10) agrees with (2), since the second-order derivative is a linear high-pass filter and the joint parameter K is readily absorbed into λ_j. Moreover, (10) can be elaborated by grouping the PZTs by the similarity in their frequency responses. This action can result in L much smaller than M, if there are many PZTs having very similar frequency responses and thus can be grouped together. Moreover, when $L = 1$, (10) is simplified to the one-dimension case, where only the main diagonals are implemented for the Volterra kernels [15]. In a relevant study, the authors have also demonstrated that with the measured frequency response of the whole ultrasonic emitter, Berktay's far-field solution is almost adequate to evaluate different preprocessing methods by simulations [24]. This implies that $L = 1$ is a possible setting to get the lowest computation cost but a compromised model accuracy.

Figure 3 illustrates the nonlinear process of the DSB modulation method. The spectrum of the filtered ultrasound input is divided into two subbands that are cut off at the carrier frequency. They are also referred to as the upper and lower sidebands. To extract the A2VF, the upper sideband is shifted by the amount of the carrier frequency in the frequency domain, while the lower sideband is necessary to be flipped before the frequency shift.

Due to the square operation, the audio output is resultant from the multiplication between the sidebands and the carrier frequency. The synthesized first-order kernel of the A2VF is written in a summation as

$$h_1^S = \sum_{i=1}^{L} \frac{m\lambda_i V_i(\Omega_c)}{4(1+m)^2} \left[v_i(n)\, e^{-j\Omega_c n} + v_i(-n)\, e^{j\Omega_c n}\right], \tag{12}$$

where $j = \sqrt{-1}$ is the imaginary unit.

The second-order audio output is resultant from the multiplication between the upper and lower sidebands. By defining a new diagonal matrix as

$$D = \frac{m^2}{4(1+m)^2} \times \Lambda \times diag[V_1(\Omega_c), \ldots, V_N(\Omega_c)] \tag{13}$$

and two matrices U and Q consisting of $v_i(n)\exp(-j\Omega_c n)$ and $v_i(-n)\exp(j\Omega_c n)$ in row-wise vectors respectively, the coefficient matrix of the synthesized second-order kernel is written in the matrix form as

$$H_2^S = UDQ^T. \tag{14}$$

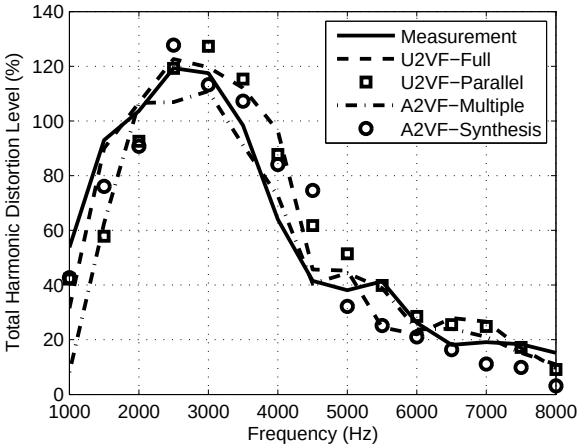


Fig. 5. THD curves when $m = 1.0$.

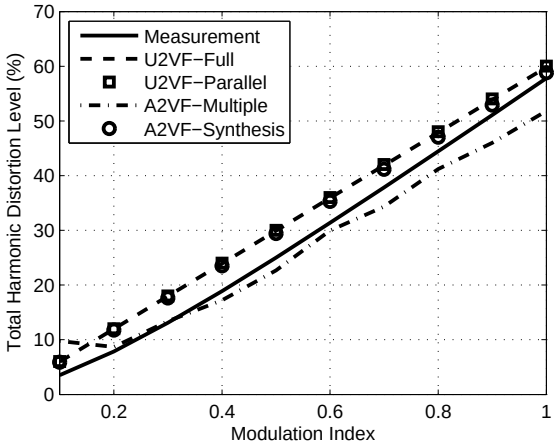


Fig. 6. Averaged THD levels at different modulation indexes.

This synthesized second-order kernel can be implemented in a modified parallel cascade structure, where every path consists of two FIR filters and the outputs of the two FIR filters are multiplied instead of taking the square operation.

4. EXPERIMENT RESULT

The experiment was carried out in a sound proof room, of which the size was $2.9 \times 3.1 \times 2.1$ m^3. The microphone (B&K Type 4191L) was placed 3 meters away from the ultrasonic emitter (Mitsubishi MSP-30E). Both the DAC and ADC had the sampling frequency of 192 kHz and the resolution of 32 bit. The ultrasound output level was kept constant throughout the experiment.

The normalized least mean squares (NLMS) algorithm was adopted. The acoustic delay was offset by 1500 samples. The memory size was set at $N = 300$. A band-passed white noise from 32 kHz and 48 kHz was transmitted directly from the ultrasonic emitter for the U2VF identification. Thereafter, the DSB modulation method was implemented with the modulation index from 0.1 to 1.0 with the interval of 0.1. The carrier frequency was chosen at 40 kHz. The A2VF was identified for every modulation index by a low-passed white noise cut off at 8 kHz.

The identified second-order kernel of the U2VF was decomposed into the parallel cascade structure. Only the largest five eigenvalues were selected, *i.e.* $L = 5$. The frequency responses of the selected and discarded paths were plotted respectively in Figs. 4(a) and 4(b). The selected paths showed similar spectral characteristics to the ultrasonic emitter, but the discarded paths had rather flat frequency responses. Subsequently, the synthesized A2VFs were calculated using the proposed method.

The total harmonic distortion (THD) test was carried out [25]. A sine sweep was created as the testing audio input. The frequency of the sine sweep ranged from 1 kHz to 8 kHz. The input of the A2VFs was provided by the testing audio input, while the input of the U2VFs was provided by the modulated testing audio input. The THD levels were extracted from the outputs of the Volterra filters. Meanwhile, the THD levels were also measured by the experiment, to be used as the reference standard. The THD curves when $m = 1.0$ and the averaged THD levels at different modulation indexes were plotted in Figs. 5 and 6, respectively.

The parallel cased structure of the U2VF (U2VF-Parallel) achieves the closest performance to the full implementation of the U2VF (U2VF-Full), even though there are only five selected paths. The A2VFs have to be identified at every modulation index (A2VF-Multiple) and lead to reasonably high accuracy at every frequency. However, when the modulation index is low, the A2VF is not able to be identified accurately. This is because a low modulation index leads to weak audio output of the PAL and thus a low signal-to-noise ratio with respect to the floor noise in the experiment. The synthesized A2VFs (A2VF-Synthesis) show the similar model accuracy to other Volterra filter representations. Inheriting the advantage of the U2VF, the synthesized A2VF at $m = 0.1$ does not suffer from the low identification accuracy of the A2VF.

5. CONCLUSIONS

A modification to the classic Berktay's far-field solution has been worked out to include the frequency responses of the PZTs in the PAL. Existing Volterra filter representations of the PAL are able to be explained via the new model equation. The parallel cascade structure of the U2VF is shown to be equivalent to the new model equation. Hence, the A2VFs are synthesized from the U2VF for different modulation indexes. The synthesized A2VFs demonstrate the similar model accuracy to other Volterra filter representations in the experiment. The synthesis of the A2VFs provides a feasible approach to build the linearization system based on the U2VF.

6. REFERENCES

[1] M. F. Hamilton and D. T. Blackstock, *Nonlinaer Acoustics*. San Diego, CA: Academic Press, 1998.

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

[3] 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.

[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] T. D. Kite, J. T. Post, and M. F. Hamilton, "Parametric array in air distortion reduction by preprocessing," in *Proc. 16th Int. Congr. Acoust.*, Seattle, 1998.

[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] 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. 76-82, 1991.

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] Y. Hatano, C. Shi, S. Kinoshita and Y. Kajikawa, "A study on linearization of nonlinear distortions in parametric array loudspeakers," in *Proc. 2015 European Sig. Process. Conf.*, Nice, France, 2015.

[15] Y. Mu, P. Ji, W. Ji, M. Wu, and J. Yang, "Modeling and compensation for the distortion of parametric loudspeakers using a one-dimension Volterra filter," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 22, no. 12, pp. 2169-2181, 2014.

[16] G. L. Sicuranza and A. Carini, "Filtered-X affine projection algorithm for multichannel active noise control using second-order Volterra filters," *IEEE Sig. Process. Lett.*, vol. 11, no. 11, pp. 853-857, 2004.

[17] Y. Hatano, C. Shi, S. Kinoshita and Y. Kajikawa, "A study on compensating for the distortion of the parametric array loudspeaker with changing nonlinearity," in *Proc. 2015 Western Pacific Acoustic. Conf.*, Singapore, 2015.

[18] Y. Hatano, C. Shi, S. Kinoshita and Y. Kajikawa, "Linearization of the parametric array loudspeaker upon varying input amplitudes," in *Proc. 2015 APSIPA Annu. Summit Conf.*, Hong Kong, 2015.

[19] C. Shi and Y. Kajikawa, "Identification of the parametric array loudspeaker with a Volterra filter using the sparse NLMS algorithm," in *Proc. 40th Int. Conf. Acoust. Speech Signal Process.*, Brisbane, Australia, 2015.

[20] C. Shi and Y. Kajiakawa, "Ultrasound-to-ultrasound Volterra filter identification of the parametric array loudspeaker," in *Proc. 2015 Int. Conf. Digital Sig. Process.*, Singapore, 2015.

[21] T. Helie and M. Hasler, "Volterra series for solving weakly non-linear partial differential equations: Application to a dissipative Burgers' equation," *Int. J. Control*, vol. 77, no. 12, pp. 1071-1082, 2004.

[22] T. Kamakura, H. Nomura, and G. T. Clement, "Linear and nonlinear ultrasound fields formed by planar sources with random pressure distributions," *Acoust. Sci. Technol.*, vol. 36, no. 3, pp. 208-215, 2015.

[23] V. J. Mathews, "Adaptive polynomial filters," *IEEE Sig. Process. Mag.*, vol. 8, no. 3, pp. 10-26, 1991.

[24] C. Shi and Y. Kajikawa, "Fast evaluation of preprocessing methods of the parametric array loudspeaker," in *Proc. 2015 Western Pacific Acoustic. Conf.*, Singapore, 2015.

[25] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique", in *Proc. 108th Audio Eng. Soc. Conv.*, Paris, France, 2000.