Compensation for Nonlinear Distortion of the Frequency Modulation Based Parametric Array Loudspeaker

Yuta Hatano, *Student Member, IEEE*, Chuang Shi, *Member, IEEE*, and Yoshinobu Kajikawa, *Senior Member, IEEE*

***Abstract*—The parametric array loudspeaker (PAL) modulates the audio signal on an ultrasonic carrier. When the modulated signal is transmitted in air, an audio beam is created based on the nonlinear acoustic principle. Each modulation method has advantages and disadvantages. The frequency modulation (FM) is favorable for its low cost and high volume, but the trade-off is its complicated nonlinear distortion, which is difficult to be reduced. In this paper, the Volterra filter is adopted to model the nonlinearity of the FM based PAL. A novel complex inverse system is devised to effectively reduce the nonlinear distortion. Three practical aspects are addressed. Firstly, the computational complexity of the Volterra filter is reduced by the parallel cascade structure with almost no compromise to the model accuracy. Secondly, Volterra filters are identified at discrete input levels to treat the nonlinearity that keeps changing with the time-varying audio input. Thirdly, when the input level is high, a separation approach is proposed to refine the identified Volterra filters, which eventually improves the performance of the proposed inverse system.**

***Index Terms*—Parametric array loudspeaker, frequency modulation, nonlinear distortion, Volterra filter, inverse system.**

I. Introduction

THE linear acoustic models are applicable to daily sound waves, whose amplitudes are smaller than their wavelengths. When the amplitude of a sound wave is large, its waveform is cumulatively distorted during the propagation in a nonlinear medium. Such a sound wave is often called the finite amplitude wave. When two finite amplitude waves propagate together, intermodulation distortion components including the sum and difference frequencies are generated in the intersection zone [1]–[3].

The PAL is a directional sound device making use of the aforementioned nonlinear acoustic principle [4]. The block diagram of the PAL is illustrated in Fig. 1. The audio input is modulated on an ultrasonic carrier. The sideband of the modulated signal leads to one finite amplitude wave and the ultrasonic carrier leads to another. The generation of the intermodulation distortion components takes place in air until the finite amplitude waves are sufficiently absorbed. Therefore, an audio beam, resultant from the difference frequency component, can be indirectly transmitted from the ultrasonic emitter. The audio beam is much narrower than the sound field created by any conventional loudspeaker, whose effective size is equivalent to that of the ultrasonic emitter [5].

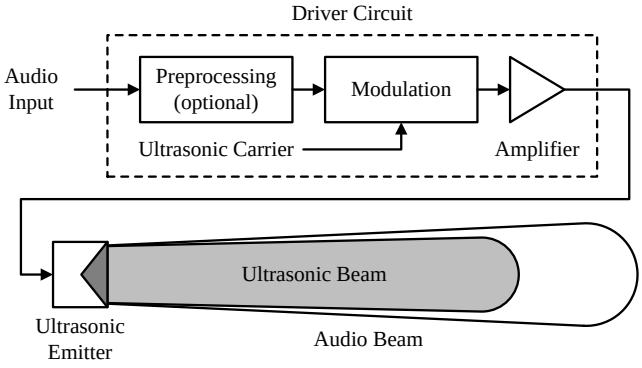

Fig. 1. Block diagram of the parametric array loudspeaker integrated with an optional preprocessing method.

However, the PAL has drawbacks in nonlinear distortion, building cost, and efficiency. Several modulation methods have been developed in the past. Each modulation method has advantages and disadvantages [6]. The amplitude modulation (AM) results in the second harmonic ratio proportional to the modulation index. The square root method, which is a preprocessing method of the AM based PAL, requires the ultrasonic emitter to have a flat frequency response [7]. Equalization of the ultrasonic emitter can greatly reduce the efficiency [8]. The single sideband (SSB) modulation, which adds a quadrature term to the AM, results in a very low distortion level when the output power is from low to moderate [9]. However, the building cost of the SSB based PAL is relatively high, due to the dedicated driver circuit with the matched high-quality ultrasonic emitter. The frequency modulation (FM) is favorable for its low cost and high efficiency. These advantages make the FM based PAL popular in the consumer market. However, the FM based PAL has more complicated nonlinear distortion than the AM and SSB based PALs.

In order to reduce the nonlinear distortion, the Volterra filter is considered to model the nonlinearity of the FM based PAL. Similar works have been carried out with the AM and SSB based PALs [10]–[12]. Inverse systems are available based on the identified Volterra filters [13]–[17]. These approaches are adapted from an established technique for linearizing the

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

Yuta Hatano and Yoshinobu Kajikawa are with the Department of Electrical and Electronic Engineering, Kansai University, Osaka, Japan 564-8680 (email: k040610@kansai-u.ac.jp; kaji@kansai-u.ac.jp).

Chuang Shi is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, China 611731 (email: shichuang@uestc.edu.cn).

conventional loudspeakers [18]–[20]. However, an overlooked fact behind this technique is that the AM does not introduce nonlinearity to the AM based PAL [21]. In contrast, the FM contains high order nonlinearity. When the audio input is time-varying, the nonlinearity modeled between the audio input and audio output of the FM based PAL is also time-varying [22]. The identified Volterra kernels always change with the input level. Especially when the input level is high, low order harmonic components resulting from high order nonlinearity are considerably higher. They may degrade the model accuracy of the identified Volterra filter.

Thus, we would like to devise a complex inverse system that consists of multiple sub-systems [23]. Each sub-system is developed to efficiently reduce the nonlinear distortion of the FM based PAL for a small range of input levels. The activation of a sub-system upon an audio input can be determined on a segment-by-segment basis. The Volterra filters in the sub-systems can be implemented using the parallel cascade structure, in order for the computational complexity to be reduced [24]. Moreover, a separation approach is necessarily proposed to improve the identification process of the Volterra filter for high input levels.

This paper is organized as follows. Section II revisits the fundamental theory of the Volterra filter and its inverse system. The parallel cascade structure is specifically shown in Section II(C). The computational complexity is thereafter compared for different implementation methods. Section III(A) presents the experimental setup and the distortion performance of a FM based PAL under test. Section III(B) examines the feasibility of the complex inverse system with five input levels. In Section III(C), the separation approach is demonstrated to suppress the interference of the third and fourth order nonlinearity in the process of identifying the Volterra filters truncated at the second order. Lastly, Section IV concludes this paper.

II. Theory and Method

A. Volterra filter

The Volterra filter is a versatile system model to represent a nonlinear system with memory [25]. When the Volterra filter is used to model the PAL, the truncated order and finite tap length are considered, *i.e.*

$$y(n) = \sum_{k_1=0}^{N-1} h_1(k_1)\, x(n-k_1) + \sum_{k_1=0}^{N-1}\sum_{k_2=0}^{N-1} h_2(k_1,k_2)\, x(n-k_1)\, x(n-k_2), \quad (1)$$

where N is the tap length; $x(n)$ and $y(n)$ are the discrete input and output signals; h_1 and h_2 are the coefficients of the first and second order Volterra kernels, respectively.

The least mean squares (LMS) algorithms are often adopted to obtain the coefficients of the Volterra kernels in the time domain. The output of an unknown nonlinear system is measured when the input is a Gaussian white noise. Alternatively, the frequency response method identifies the Volterra kernels in the frequency domain [26]. The linear and nonlinear frequency responses of an unknown nonlinear system are measured. The coefficients of the Volterra kernels are subsequently calculated based on the inverse Fourier transform of the measured frequency responses.

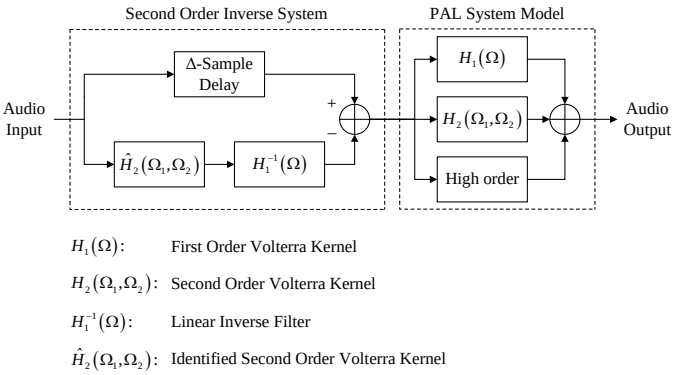

Fig. 2. Block diagram of the inverse system for the parametric array loudspeaker based on the first and second order Volterra kernels.

The frequency response of the first order Volterra kernel, *a.k.a.* the linear frequency response, is calculated from the ratio of the output spectrum $Y(\Omega)$ and the input spectrum $X(\Omega)$ as

$$\hat{H}_1(\Omega) = \frac{Y(\Omega)}{X(\Omega)}, \quad (2)$$

where the input is a sine sweep signal and the output is obtained by the measurement.

Two sine sweep signals are mixed and input into the unknown nonlinear system to measure the second order nonlinear frequency response. The output spectrum $Y(\Omega_1+\Omega_2)$ is divided by the product of input spectra $X(\Omega_1)$ and $X(\Omega_2)$, which is expressed by

$$\hat{H}_2(\Omega_1,\Omega_2) = \frac{Y(\Omega_1+\Omega_2)}{X(\Omega_1)X(\Omega_2)}\frac{N}{\alpha_2}, \quad (3)$$

where α_2 denotes the number of symmetries. $\alpha_2 = 1$ when $\Omega_1 = \Omega_2$; otherwise, $\alpha_2 = 2$.

B. Inverse system

Figure 2 shows the block diagram of the inverse system for the PAL based on the first and second order Volterra kernels. $H_1(\Omega)$ and $H_2(\Omega_1,\Omega_2)$ are the true first and second order models of the PAL, respectively. $H_1^{-1}(\Omega)$ is assumed to be a perfectly designed linear inverse filter yielding $H_1^{-1}(\Omega)H_1(\Omega) = \mathrm{e}^{-\mathrm{j}\Omega\Delta}$, where $\mathrm{j} = \sqrt{-1}$ is the imaginary unit and Δ is the number of samples in a time delay. The tap length of the linear inverse filter must be longer than the tap length of the Volterra filter. In this paper, we choose a tap length of $4N$ for the linear inverse filter.

Truncating the nonlinearity of the PAL at the second order, the combined nonlinear response of the inverse system and the PAL is written as

$$\begin{aligned}&[\mathrm{e}^{-\mathrm{j}\Omega\Delta} - \hat{H}_2(\Omega_1,\Omega_2)H_1^{-1}(\Omega)][H_1(\Omega)+H_2(\Omega_1,\Omega_2)]\\ =&\mathrm{e}^{-\mathrm{j}\Omega\Delta}H_1(\Omega) - \hat{H}_2(\Omega_1,\Omega_2)H_1^{-1}(\Omega)H_1(\Omega_1,\Omega_2)\\ &+\mathrm{e}^{-\mathrm{j}\Omega\Delta}H_2(\Omega_1,\Omega_2) - \hat{H}_2(\Omega_1,\Omega_2)H_1^{-1}(\Omega)H_2(\Omega). \quad (4)\end{aligned}$$

The second order component in (4) can be further manipulated as

$$\begin{aligned} &\mathrm{e}^{-\mathrm{j}\Omega\Delta} H_2(\Omega_1,\Omega_2) - \hat{H}_2(\Omega_1,\Omega_2) H_1^{-1}(\Omega) H_1(\Omega) \\ =&\mathrm{e}^{-\mathrm{j}\Omega\Delta}[H_2(\Omega_1,\Omega_2) - \hat{H}_2(\Omega_1,\Omega_2)]. \end{aligned} \tag{5}$$

Therefore, if there is no identification error, *i.e.* $H_2(\Omega_1,\Omega_2) = \hat{H}_2(\Omega_1,\Omega_2)$, the second order nonlinear distortion of the PAL can be fully eliminated by the inverse system. However, when the Volterra filter is inaccurately identified, the performance of the inverse system is degraded.

C. Parallel Cascade Structure

Moreover, the second order Volterra component is rewritten in a matrix form as

$$\begin{aligned} y_2(n) &= \sum_{k_1=0}^{N-1}\sum_{k_2=0}^{N-1} h_2(k_1,k_2)\, x(n-k_1)\, x(n-k_2) \\ &= \mathbf{X}^T(n)\,\mathbf{H}_2\mathbf{X}(n), \end{aligned} \tag{6}$$

where $\mathbf{X}(n) = [x(n),\ x(n-1),\ \ldots,\ x(n-N+1)]$ is the input signal vector; and $\mathbf{H}_2$ is the matrix form of the coefficients of the second order Volterra kernel.

Without loss of generality, $\mathbf{H}_2$ is assumed to be a symmetric Volterra kernel. It can be decomposed as

$$\mathbf{H}_2 = \sum_{i=0}^{N-1} \lambda_i \mathbf{L}_i \mathbf{L}_i^T, \tag{7}$$

where λ_i is i-th eigenvalue and $\mathbf{L}_i$ is i-th eigenvector given by

$$\mathbf{L}_i = [\ l_{i,0} \quad l_{i,1} \quad \cdots \quad l_{i,N-1}\ ]^T. \tag{8}$$

Substituting (7) to (6) yields the parallel cascade expression of the second order Volterra component as

$$\begin{aligned} y_2(n) &= \sum_{i=0}^{N-1} \lambda_i [\mathbf{X}^T(n)\,\mathbf{L}_i][\mathbf{L}_i^T\mathbf{X}(n)] \\ &= \sum_{i=0}^{N-1} \lambda_i y_{L,i}^2(n), \end{aligned} \tag{9}$$

where $y_{L,i}(n) = \mathbf{X}^T(n)\,\mathbf{L}_i = \mathbf{L}_i^T\mathbf{X}(n)$ denotes the i-th parallel path [27]. Based on (9), the parallel cascade structure is drawn in Fig. 3.

The computational complexity of the parallel cascade structure is adjusted by sorting out the less significant eigenvalues. It is always a trade-off that keeping more eigenvalues result in higher computational complexity but also higher model accuracy. For instance, an inverse system implements the second order Volterra kernel with the tap length of N and the linear inverse filter with the tap length of $4N$. The computational complexity of this inverse system is reduced to $M(N+2)+4N$ multiplications and $MN+4N-1$ additions, when only the most significant M eigenvalues are kept in the parallel cascade structure. When M is relatively small, the computational complexity is $O(N)$ instead of $O(N^2)$. In comparison, the computational complexity of the inverse system implemented with the symmetrical Volterra filter is given by N^2+5N multiplications and $N(N+1)/2+4N-1$ additions. The one dimensional Volterra filter results in the lowest computational complexity of the inverse system, which is $6N$ multiplications and $5N-1$ additions. But the inverse system implemented with the one dimensional Volterra filter is likely to be less efficient to reduce the intermodulation distortion.

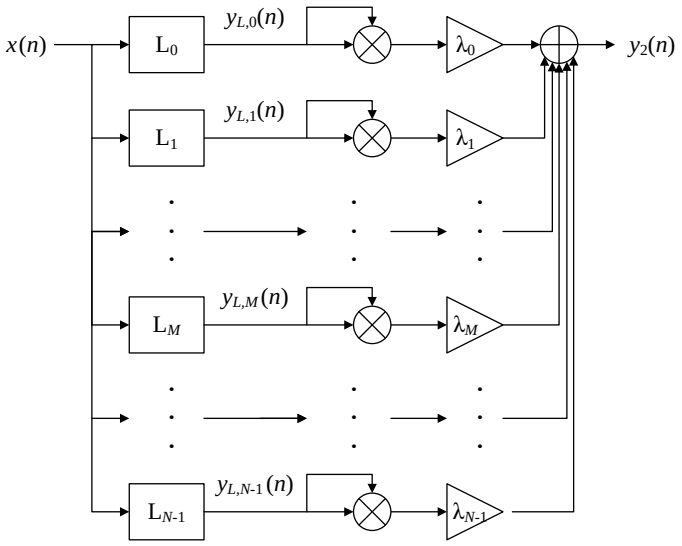

Fig. 3. Block diagram of the parallel cascade structure for the second order Volterra component.

III. Experimental Validation

A. Experimental Setup

The experimental setup is depicted in Fig. 4. The walls of the recording room are treated with sound absorbing textile material to suppress reflection. The absorption coefficients at 500 Hz and 4000 Hz are specified by the manufacturer as 99%. Due to the room size ($2.9 \times 3.1 \times 2.1\ m^3$), the PAL and the microphone are placed diagonally to maintain a distance of 3.0 meters. This distance has been validated to result in almost identical observations with or without an acoustical filter. In [3], [28], it has also been reported that the difference frequency pressures are at close levels in the filtered and unfiltered measurement results when the distance between the PAL and the microphone is more than a quarter of the absorption distance.

The sampling frequency of the measurement system is set to 16 kHz. The tap length of the Volterra filters is set to $N = 512$. The temperature and the relative humidity are recorded at 26°C and 35%, respectively. Therefore, the acoustical delay between the PAL and the microphone is estimated to be 103 samples. The tap length of the Volterra filter is confirmed to be long enough to cope with such an acoustical delay [29].

Two sine sweep signals are prepared with the frequency range from 500 Hz to 2000 Hz. The upper frequency bound of the sine sweep signals should not exceed one eighth of the sampling frequency, in order that the fourth harmonic frequency is not greater than half of the sampling frequency. The input levels are measured at the audio jack of the PAL.

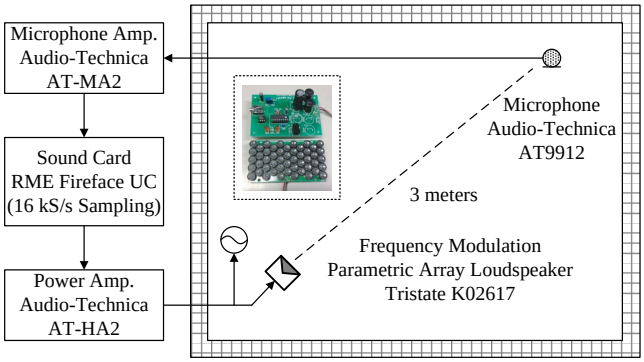

Fig. 4. Experimental setup, with an inserted photograph of the frequency modulation based parametric array loudspeaker.

Five discrete input levels are tentatively selected at 0.05, 0.10, 0.15, 0.20, and 0.25 Volts.

The distortion reduction is the main performance measure. It is defined as the relative amplitude of the undesired frequency component before and after applying the inverse system. The total harmonic distortion (THD) is also used and defined as

$$\mathrm{THD} = \sqrt{\frac{\mathrm{T}_2^2 + \mathrm{T}_3^2 + \mathrm{T}_4^2}{\mathrm{T}_1^2}} \times 100\%, \tag{10}$$

where $\mathrm{T_i}$ represents the amplitude of the i-th order harmonic component $(i \leq 4)$. Furthermore, the second harmonic ratio (SHR), third harmonic ratio (THR), and fourth harmonic ratio (FHR) are defined as

$$\mathrm{SHR} = \frac{\mathrm{T}_2}{\mathrm{T}_1} \times 100\%, \tag{11}$$

$$\mathrm{THR} = \frac{\mathrm{T}_3}{\mathrm{T}_1} \times 100\%, \tag{12}$$

and

$$\mathrm{FHR} = \frac{\mathrm{T}_4}{\mathrm{T}_1} \times 100\%, \tag{13}$$

respectively. In the latter part of this paper, measurement results are presented by the average amount across the frequency range of the sine sweep signals.

The PAL used in the experiment adopts the FM. The driving circuit of the FM based PAL has a very compact size, since it is adapted from the switching amplifier. The modulated signal of the FM is known as

$$u\left(t\right) = g_u \cos\left[\omega_c t + k\int_0^t x(\tau)d\tau\right], \tag{14}$$

where g_u and ω_c are the amplitude and frequency of the ultrasonic carrier, respectively; k is a positive constant, namely the frequency sensitivity; and $x(t)$ is the audio input.

When the audio input is provided by a sine tone, *i.e.* $x(t) = g_a \cos\omega_0 t$, the modulated signal can be rewritten as

$$u\left(t\right) = g_u \cos\left[\omega_c t + \frac{k g_a}{\omega_0}\sin\omega_0 t\right], \tag{15}$$

where g_a and ω_0 denotes the amplitude and frequency of the sine tone. The modulation index of the FM is further defined as

$$m = \frac{k g_a}{\omega_0}. \tag{16}$$

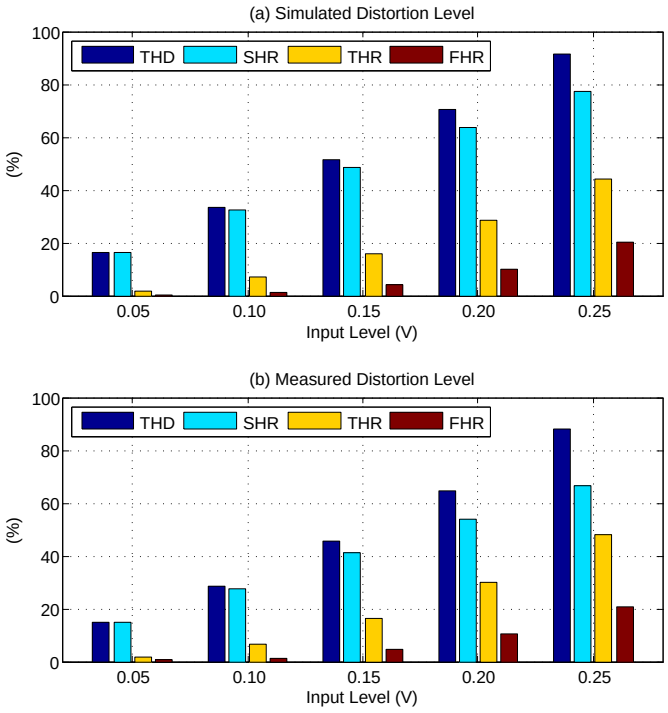

Fig. 5. Harmonic distortion performance of the frequency modulation based parametric array loudspeaker used in the experiment.

Therefore, the modulated signal is expanded into a series as

$$u\left(t\right) = g_u \sum_{i=-\infty}^{+\infty} J_i(m)\cos\left[\left(\omega_c + i\omega_0\right)t\right], \tag{17}$$

where $J_i\left(m\right)$ is the ith order Bessel function of the first kind [30].

Although an AM and FM integrated Berktay's equation is available in [31], it is difficult to cooperate with practical ultrasonic emitters with narrow bandwidths. Alternatively, the harmonic distortion performance of the FM based PAL is readily simulated using the classic Berktay's equation. The experiment and simulation results are plotted in Fig. 5. Since the FM based PAL in the experiment is an off-the-shelf product, the modulation index is empirically set as $m = 6.5 g_a$. The limited bandwidth of the ultrasonic emitter is simulated by restricting $|i| \leq 4$ in (17). Figure 5 confirms that the nonlinearity of the FM based PAL is associated with the input level. The second order harmonic distortion is dominant and quasi-linearly related to the input level. The small discrepancies between the measurement and simulation results are due to the lack of information about the frequency response of the ultrasonic emitter [32].

B. Time-varying input

Thereafter, we devise a complex inverse system that consists of multiple sub-systems, as shown in Fig. 6. Each sub-system is developed to be efficient for a limited range of input levels. The changing nonlinearity caused by the time-varying input can be solved by selecting the suitable sub-system on a segment-by-segment basis. The computational complexity

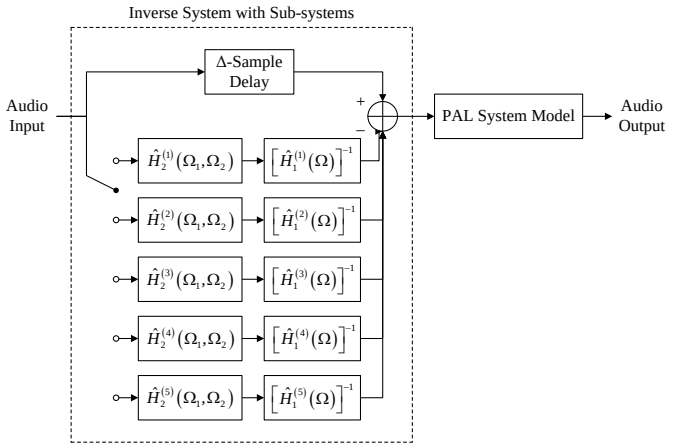

Fig. 6. Block diagram of the complex inverse system consisting of five sub-systems.

of the complex inverse system is not significantly increased, because only one sub-system is activated at any one time.

To validate the feasibility, five Volterra filters are identified based on the experimental setup shown in Fig. 4, when the input levels are set to 0.05, 0.10, 0.15, 0.20, and 0.25 Volts. Afterwards, five sub-systems are correspondingly implemented. The distortion reduction is measured for each sub-system. The measurement results of the second harmonic, sum frequency, and difference frequency components are plotted in Figs. 7, 8, and 9, respectively. The voltage in the legend indicates the input level in the identification process. It is observed that when the input levels for identification and validation are unmatched, the performance of each sub-system significantly worsens for all types of the second order nonlinear distortion.

The performance on the second harmonic component worsens when the input level increases, because the model accuracy is degraded by the interference from high order nonlinearity in the identification process. The performance on the sum frequency component is less affected by the input level. This is due to the fact that high order nonlinearity can result in more low order harmonic components rather than intermodulation distortion components. Moreover, the performance on the difference frequency component is improved when the input level increases. The difference frequency component is relatively weak even before applying the inverse system. The model accuracy of the Volterra filter may be deteriorated by the noise incurred in the measurement system when the input level is low.

Furthermore, the sub-systems in the complex inverse system are implemented with the parallel cascade structure, whereby the top $M = 224$ eigenvalues are selected by trial and error, in order that the performance is not noticeably compromised. This setting allows the numbers of the multiplications and additions to be reduced by 55.7% and 12.5%, as compared to the symmetrical Volterra filter. It should be noted that there is no standard in selecting the number of eigenvalues, as it is always a trade-off that a larger number costs more computational power and results in higher model accuracy. The distortion reductions in the second harmonic, sum frequency,

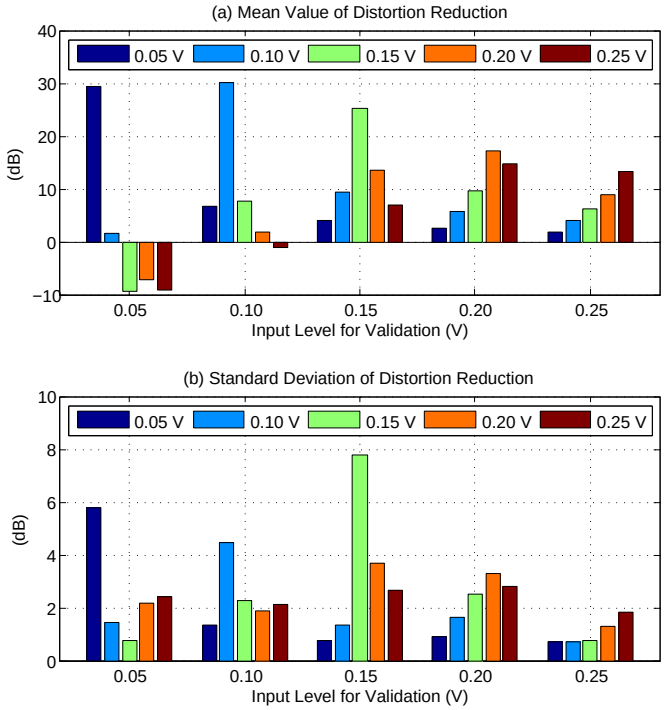

Fig. 7. Distortion reduction in the second harmonic component, validated with five sub-systems built for different input levels of 0.05, 0.10, 0.15, 0.20, and 0.25 Volts.

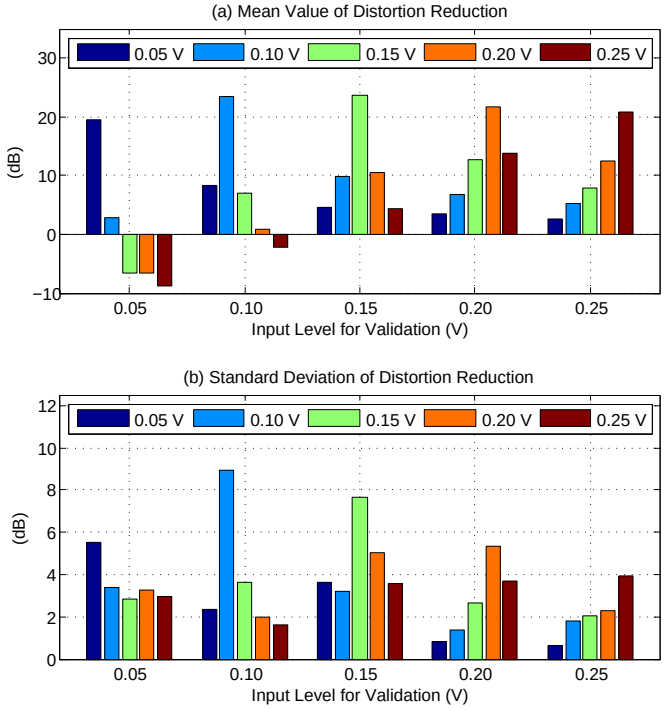

Fig. 8. Distortion reduction in the sum frequency component, validated with five sub-systems built for different input levels of 0.05, 0.10, 0.15, 0.20, and 0.25 Volts.

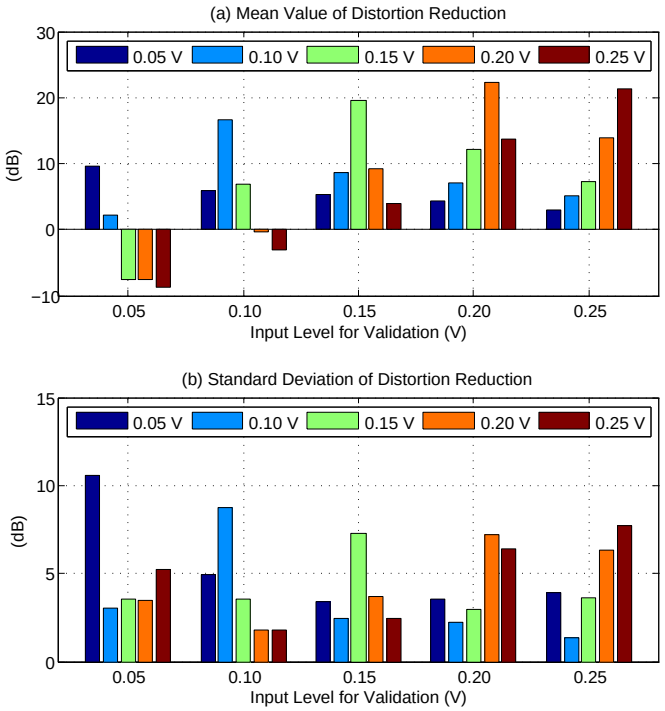

Fig. 9. Distortion reduction in the difference frequency component, validated with five sub-systems built for different input levels of 0.05, 0.10, 0.15, 0.20, and 0.25 Volts.

TABLE I
COMPARISON BETWEEN DIFFERENT IMPLEMENTATION METHODS WHEN THE INPUT LEVELS FOR IDENTIFICATION AND VALIDATION ARE SAME.

Input Level (V)	0.05	0.10	0.15	0.20	0.25
Distortion reduction in the second harmonic component (dB)					
Symmetrical Volterra Filter	29.35	30.22	25.13	17.16	13.21
Parallel Cascade Structure	25.11	28.83	22.35	17.56	11.82
Distortion reduction in the sum frequency component (dB)					
Symmetrical Volterra Filter	19.54	23.37	23.64	21.81	20.80
Parallel Cascade Structure	19.09	22.98	25.39	22.09	20.13
Distortion reduction in the difference frequency component (dB)					
Symmetrical Volterra Filter	9.63	16.72	19.49	22.35	21.23
Parallel Cascade Structure	8.75	18.02	19.96	22.75	19.69

and difference frequency components are summarized in Table I. The sub-systems implemented with the parallel cascade structure is able to yield equivalent distortion reduction performance to those implemented with the symmetric Volterra filter. However, when the input level is large, the inverse system cannot reduce the second harmonic component consistently. This is due to the degraded model accuracy rather than the implementation methods.

C. Identification accuracy upon high input level

According to the above observation, a separation approach is necessary to suppress the interference from high order nonlinearity in the process of identifying the first and second order Volterra kernels. The derivation of the separation approach is presented in this subsection.

Firstly, the third and fourth order nonlinear frequency responses are expressed as

$$H_3(\Omega_1,\Omega_2,\Omega_3) = \frac{Y_3(\Omega_1+\Omega_2+\Omega_3)}{X(\Omega_1)X(\Omega_2)X(\Omega_3)}\frac{N^2}{\alpha_3} \tag{18}$$

and

$$H_4(\Omega_1,\Omega_2,\Omega_3,\Omega_4) = \frac{Y_4(\Omega_1+\Omega_2+\Omega_3+\Omega_4)}{X(\Omega_1)X(\Omega_2)X(\Omega_3)X(\Omega_4)}\frac{N^3}{\alpha_4} \tag{19}$$

where $Y_3(\Omega_1+\Omega_2+\Omega_3)$ and $Y_4(\Omega_1+\Omega_2+\Omega_3+\Omega_4)$ are output spectra; $X(\Omega_1)$, $X(\Omega_2)$, $X(\Omega_3)$, and $X(\Omega_4)$ are input spectra; α_3 and α_4 denote the numbers of symmetries in the third and fourth order nonlinear responses, respectively.

When the input is a sine tone, the high order intermodulation distortion generates the output at the fundamental and second harmonic frequencies as

$$Y_3(\Omega) = 3H_3(\Omega,\Omega,-\Omega)X^2(\Omega)X(-\Omega)/N^2 \tag{20}$$

and

$$Y_4(2\Omega) = 4H_4(\Omega,\Omega,\Omega,-\Omega)X^3(\Omega)X(-\Omega)/N^3, \tag{21}$$

where $\alpha_3 = 3$ and $\alpha_4 = 4$ have been substituted. The above two equations imply that the measured output spectra contain interferences as

$$Y(\Omega) = Y_1(\Omega) + Y_3(\Omega) \tag{22}$$

and

$$Y(2\Omega) = Y_2(2\Omega) + Y_4(2\Omega), \tag{23}$$

where

$$Y_1(\Omega) = H_1(\Omega)X(\Omega) \tag{24}$$

and

$$Y_2(2\Omega) = H_2(\Omega,\Omega)X^2(\Omega)/N \tag{25}$$

denote the desired output spectra of the linear and second order harmonic components, respectively.

Hence, the identified first and second order Volterra kernels by the frequency response method lead to the inaccurate linear and second order nonlinear frequency responses as

$$\hat{H}_1(\Omega) = H_1(\Omega) + 3H_3(\Omega,\Omega,-\Omega)S(\Omega) \tag{26}$$

and

$$\hat{H}_2(\Omega,\Omega) = H_2(\Omega,\Omega) + 4H_4(\Omega,\Omega,\Omega,-\Omega)S(\Omega), \tag{27}$$

where $S(\Omega) = X(\Omega)X(-\Omega)/N^2$ is the power spectrum.

After assuming that a small perturbation in the input level has no significant effect on the nonlinearity of the PAL, a separation approach can be feasible. The output spectra are measured twice for each input level. The first measurement is carried out as normal. In the second measurement, the input level is multiplied with a perturbation factor σ. For example, when $\sigma = 0.95$, the input levels of 0.20 and 0.19 Volts are both measured in order to get a refined model for the input level of 0.20 Volts. Thereafter, two sets of simultaneous equations are formulated as

$$\begin{bmatrix} 1 & 3S(\Omega) \\ \sigma & 3\sigma^3 S(\Omega) \end{bmatrix} \begin{bmatrix} H_1(\Omega) \\ H_3(\Omega,\Omega,-\Omega) \end{bmatrix} = \begin{bmatrix} H_{11}(\Omega) \\ H_{12}(\Omega) \end{bmatrix} \tag{28}$$

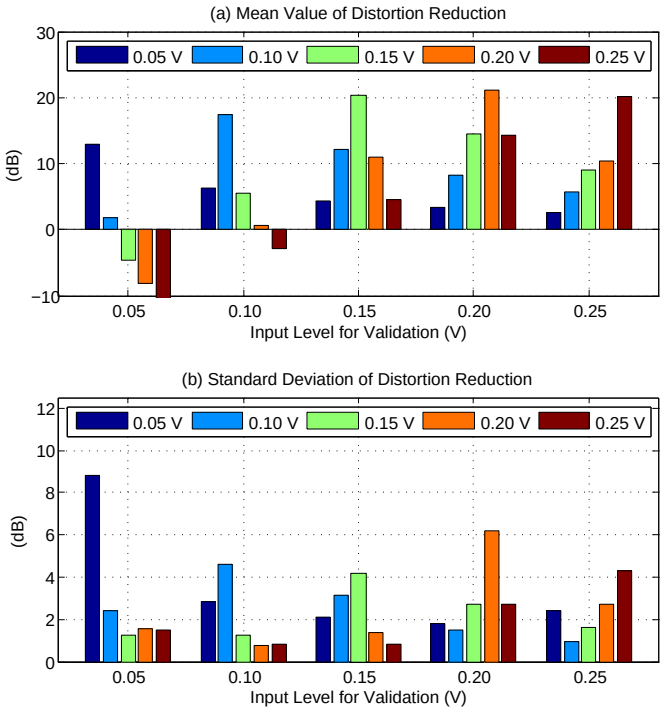

Fig. 10. Distortion reduction in the second harmonic component when the interference of both the third and fourth order nonlinearity is suppressed in the identification process of the Volterra filter.

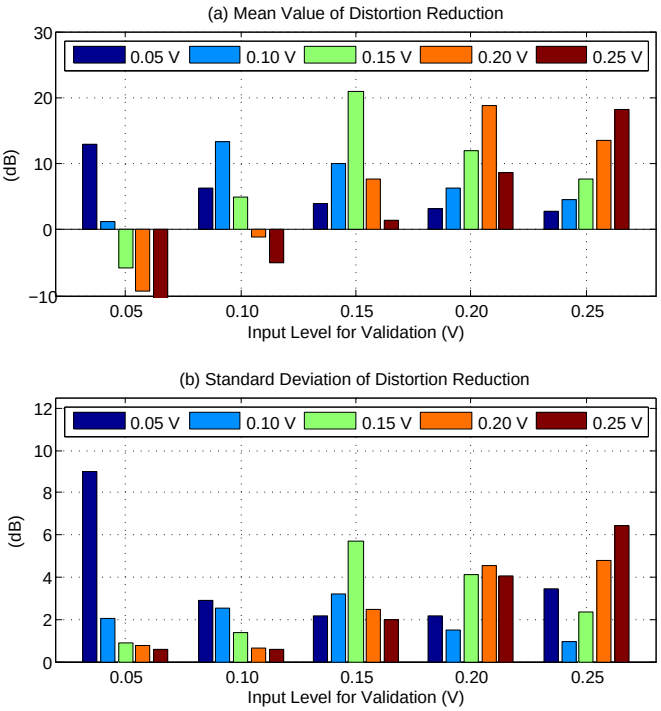

Fig. 11. Distortion reduction in the second harmonic component when the interference of the fourth order nonlinearity is suppressed in the identification process of the Volterra filter.

and

$$\begin{bmatrix} 1 & 4S(\Omega) \\ \sigma^2 & 4\sigma^4 S(\Omega) \end{bmatrix} \begin{bmatrix} H_2(\Omega,\Omega) \\ H_4(\Omega,\Omega,\Omega,-\Omega) \end{bmatrix} = \begin{bmatrix} H_{21}(\Omega,\Omega) \\ H_{22}(\Omega,\Omega) \end{bmatrix}, \tag{29}$$

where $H_{11}(\Omega)$ and $H_{12}(\Omega)$ are the linear frequency responses measured before and after the perturbation; $H_{21}(\Omega,\Omega)$ and $H_{22}(\Omega,\Omega)$ are the second order nonlinear frequency responses measured before and after the perturbation.

The solutions to (28) and (29) yield the improved estimates of the linear and second order nonlinear frequency responses as

$$\hat{H}_1(\Omega) = \frac{H_{12}(\Omega) - \sigma^3 H_{11}(\Omega)}{\sigma - \sigma^3} \tag{30}$$

and

$$\hat{H}_2(\Omega,\Omega) = \frac{H_{22}(\Omega,\Omega) - \sigma^4 H_{21}(\Omega,\Omega)}{\sigma^2 - \sigma^4}, \tag{31}$$

respectively.

The distortion reduction in the second harmonic component is plotted in Fig. 10, when the Volterra filters in the complex inverse system are refined using both (30) and (31) to suppress the interference of both the third and fourth order nonlinearity. The performance of the sub-systems improves when the input levels for identification are larger than 0.15 Volts. These input levels lead to notable high order nonlinearity of the FM based PAL.

Moreover, Fig. 11 shows the distortion reduction in the second harmonic component when only the interference of the fourth order nonlinearity is suppressed using (31). The improvement of the sub-systems is observed when the input levels for identification are larger than 0.15 Volts. However,

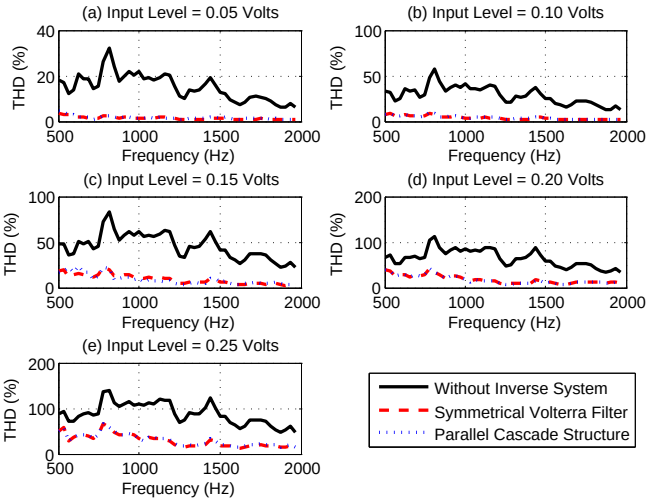

Fig. 12. Total harmonic distortion performance before and after applying the inverse system, whereby the separation approach is applied only if the performance can be improved.

the increment in the distortion reduction is less significant, as compared to Fig. 10. This is due to the fact that both the first and second order Volterra kernels are affected by the changing input level of the FM based PAL.

When the parallel cascade structure is adopted for implementation, the performance of the inverse system is summarized in Table II. The numbers in bold indicate the optimized performance that every sub-systems achieves. Figure 12 further shows the THD performance with respect to the frequency before and after applying the inverse system, whereby the

TABLE II
HARMONIC DISTORTION PERFORMANCE BEFORE AND AFTER APPLYING THE INVERSE SYSTEM.

Input Level (V)	0.05	0.10	0.15	0.20	0.25
THD (%)					
Without Inverse System	15.05	28.72	45.50	64.65	88.12
Symmetrical Volterra Filter	**1.42**	**3.91**	**9.06**	19.05	32.67
(After Separation)	6.77	5.92	9.16	**17.59**	**28.80**
Parallel Cascade Structure	**1.58**	**4.16**	9.39	19.17	35.28
(After Separation)	6.75	5.90	**8.91**	**17.29**	**30.44**
SHR (%)					
Without Inverse System	14.92	27.59	41.18	54.02	66.70
Symmetrical Volterra Filter	**0.59**	**0.99**	**3.19**	8.37	15.96
(After Separation)	6.41	4.72	4.60	**5.49**	**7.70**
Parallel Cascade Structure	**0.99**	**1.27**	**3.49**	8.34	18.04
(After Separation)	6.43	4.56	4.04	**5.17**	**9.82**
THR (%)					
Without Inverse System	1.60	6.76	16.20	29.82	47.94
Symmetrical Volterra Filter	**1.07**	3.60	7.66	14.65	23.48
(After Separation)	1.67	**3.24**	**6.94**	**13.13**	**20.95**
Parallel Cascade Structure	**1.06**	3.76	7.81	14.79	24.66
(After Separation)	1.63	**3.43**	**7.10**	**12.96**	**21.38**
FHR (%)					
Without Inverse System	0.61	1.39	4.45	10.33	20.68
Symmetrical Volterra Filter	0.62	0.94	3.24	**7.98**	**14.64**
(After Separation)	**0.58**	**0.91**	**2.97**	8.69	16.64
Parallel Cascade Structure	0.60	0.96	3.35	**8.04**	**15.61**
(After Separation)	**0.58**	**0.89**	**3.11**	8.72	17.68

separation approach is applied only if the performance can be improved. In general, the parallel cascade structure degrades the overall performance by a trivial amount since a relatively large number of eigenvalues have been selected. The SHR maintains below 10% by both implementations of the inverse system. When the input levels are not greater than 0.15 Volts, the THD manages to be lower than 10%. This is an encouraging performance for the FM based PAL. The separation approach is validated to be effective when the input levels are relatively large, such as 0.20 and 0.25 Volts. Although the separation approach reduces the THD by about 5% when the input level is 0.25 Volts, high order inverse systems are still needed to reduce high order nonlinear distortion caused by the FM.

A preliminary subjective test is carried out with 17 subjects. The purpose of this subjective test is to associate the distortion reduction with the perceptual improvement. Three tonal (500 Hz, 700 Hz, 1200 Hz) and three dual-tonal (500 Hz and 700 Hz, 500 Hz and 1200 Hz, 700 Hz and 1200 Hz) stimuli are examined at the input level of 0.20 Volts. The mean opinion score (MOS) is adapted to evaluate the perceptual quality before and after applying the inverse system. The range of the MOS is set from 1 to 5. The MOS of 1 means that the distortion is extremely obvious, while the MOS of 5 means the distortion is unable to be recognized. Before the inverse system is applied, the lowest MOS is found at 1.63 for the dual-tonal stimulus of 700 Hz and 1200 Hz. This is because the difference frequency component at 500 Hz is significant. Similarly, the dual-tonal stimulus of 500 Hz and 1200 Hz also results in a very low MOS of 1.64, due to the difference frequency component at 700 Hz. The averaged MOS of the six stimuli is 1.80. After applying the inverse system, the MOSs for the dual-tonal stimuli containing 1200 Hz are increased to 3.28 and 2.66, respectively. The averaged MOS of the six stimuli is increased to 2.82. This simple subjective test demonstrates that the improvement in the perceptual clarity of tonal and multi-tonal stimuli is distinguishable when the inverse system achieved about 20 dB of distortion reduction. It is also worth noting that besides sound reproduction, the PAL has been used for in-situ absorption coefficient measurement, active noise control, and non-destructive examination [33]–[35], where the input signals are tonal or multi-tonal. The proposed inverse system will be helpful to increase the performance of these applications.

IV. CONCLUSIONS

In this paper, a complex inverse system based on multiple Volterra filters was devised for the FM based PAL. Three practical concerns were addressed. Firstly, the parallel cascade structure was validated to be able to maintain the performance and reduce the computational complexity of the Volterra filters. Secondly, the FM based PAL was demonstrated to have changing nonlinearity with respect to the input level. Therefore, the complex inverse system had to consist of multiple sub-systems. In the experiments, the sub-systems were prepared for five input levels. The measurement results showed that when the input levels for identification and validation matched, the sub-systems achieved the best performance. There was no single sub-system with the ability to cover a large range of input levels. Lastly, high order nonlinearity incurred in the FM based PAL was treated by the proposed separation approach, which was validated to be an effective and necessary measure when the input level was large.

REFERENCES

[1] P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Amer.*, vol. 35, no. 4, pp. 535–537, 1963.
[2] H. O. Berktay, "Possible exploitation of non-linear acoustics in underwater transmitting applications," *J. Sound Vib.*, vol. 2, no. 4, pp. 435-461, 1965.
[3] M. B. Bennett and D. T. Blackstock, "Parametric array in air," *J. Acoust. Soc. Amer.*, vol. 57, no. 3, pp. 562–568, 1975.
[4] 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. 1532–1536, 1983.
[5] C. Shi and Y. Kajikawa, "A convolution model for computing the far-field directivity of a parametric loudspeaker array," *J. Acoust. Soc. Amer.*, vol. 137, no. 2, pp. 777–784, 2015.
[6] P. Ji, E. L. Tan, W. S. Gan, and J. Yang, "A comparative analysis of preprocessing methods for the parametric loudspeaker based on the Khokhlov-Zabolotskaya-Kuznetsov equation for speech reproduction," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 4, pp. 937–946, 2011.
[7] T. Kamakura, M. Yoneyama, and K. Ikegaya, "Developments of parametric loudspeaker for practical use," in *Proc. 10th Int. Symp. Nonlinear Acoust.*, Kobe, Japan, 1984, pp. 147–150.
[8] W. Kim and V. W. Sparrow, "Audio application of the parametric array: Implementation through a numerical model," in *Proc. 113th Audio Eng. Soc. Conv.*, Los Angeles, California, 2002, pp. 1–16.
[9] K. Aoki, T. Kamakura, and Y. Kumamoto, "Parametric loudspeaker: Characteristics of acoustic field and suitable modulation of carrier ultrasound," *Electron. Comm. Jpn.*, vol. 74, no. 9, pp. 76–82, 1991.
[10] K. 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, pp. 1–5.

[11] W. Ji and W. S. Gan, "Identification of a parametric loudspeaker system using an adaptive Volterra filter," *Applied Acoust.*, vol. 73, no. 12, pp. 1251–1262, 2012.
[12] 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.
[13] K. C. M. Lee and W. S. Gan, "Bandwidth-efficient recursive *p*th–order equalization for correcting baseband distortion in parametric loudspeakers," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 2, pp. 706–710, 2006.
[14] 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.
[15] O. Hoshuyama and T. Okayasu, "Decimation in Volterra filter for distortion reduction of ultrasonic parametric loudspeaker," presented at *Proc. 2014 IEICE Soc. Conf.*, Tokushima, Japan, 2014.
[16] Y. Mu, P. Ji, W. Ji, M. W, and J. Yang, "Multipoint linearization of parametric loudspeakers using Volterra filters," in *Proc. 2014 APSIPA Annu. Summit Conf.*, Siem Reap, Cambodia, 2014, pp. 1–6.
[17] 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.
[18] 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.
[19] T. Helie and M. Hasler, "Volterra series for solving weakly nonlinear partial differential equations: Application to a dissipative Burgers' equation," *Int. J. Control*, vol. 77, no. 12, pp. 1071–1082, 2004.
[20] S. S. Payal and V. J. Mathews, "Equalization of nonlinear systems modeled using the Burgers equation" in *Proc. 2015 European Sig. Process. Conf.*, Nice, France, 2015, pp. 674–678.
[21] C. Shi and Y. Kajikawa, "Volterra model of the parametric array loudspeaker operating with ultrasonic frequencies," *J. Acoust. Soc. Amer.*, vol. 140, no. 5, 3643–3650, 2016.
[22] 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 Acoust. Conf.*, Singapore, 2015, pp. 334–337.
[23] C. Shi and Y. Kajikawa, "Synthesis of Volterra filters for the parametric array loudspeaker," in *Proc. 41st IEEE Int. Conf. Acoust. Speech Sig. Process.*, Shanghai, China, 2016.
[24] 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, pp. 1073–1077.
[25] V. J. Mathews, "Adaptive polynomial filters," *IEEE Sig. Process. Mag.*, vol. 8, no. 3, pp. 10–26, 1991.
[26] M. Tsujikawa, T. Shiozaki, Y. Kajikawa, and Y. Nomura, "Identification and elimination of second order nonlinear distortion of loudspeaker systems using digital Volterra filter," in *Proc. 2000 European Sig. Process. Conf.*, Tampere, Finland, 2000, pp. 1–4.
[27] T. M. Panicker, V. J. Mathews, and G. L. Sicuranza, "Adaptive parallel-cascade truncated Volterra filters," *IEEE Trans. Sig. Process.*, vol. 46, no. 10, pp. 2664–2673, 1998.
[28] I. O. Wygant, M.Kupnik, J. C. Windsor, W. M. Wright, M. S. Wochner, G. G. Yaralioglu, M. F. Hamilton, and B. T. Khuri-Yakub, "50 kHz capacitive micromachined ultrasonic transducers for generation of highly directional sound with parametric arrays," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 56, no. 1, pp. 193–203, 2009.
[29] C. Shi and Y. Kajikawa, "Identification of the parametric array loudspeaker with a Volterra filter using the sparse NLMS algorithm," in *Proc. 40th IEEE Int. Conf. Acoust. Speech Sig. Process.*, Brisbane, Australia, 2015, pp. 3372–3376.
[30] J. M. Chowning, "The synthesis of complex audio spectra by means of frequency modulation," *J. Audio Eng. Soc.*, vol. 21, no. 7, pp. 526–534, 1973.
[31] M. A. Averkiou, Y. S. Lee, and M. F. Hamilton, "Self-demodulation of amplitude- and frequency-modulated pulses in a thermoviscous fluid," *J. Acoust. Soc. Amer.*, vol. 94, no. 5, pp. 2876–2883, 1993.
[32] C. Shi and Y. Kajikawa, "Effect of the ultrasonic emitter on the distortion performance of the parametric array loudspeaker," *Applied Acoust.*, vol. 112, pp. 108–115, 2016.
[33] N. Tanaka and M. Tanaka, "Active noise control using a steerable parametric array loudspeaker," *J. Acoust. Soc. Amer.*, vol. 127, no. 6, pp. 3526–3537, 2010.
[34] Z. Kuang, C. Ye, M. Wu, and J. Yang, "A method for measuring the absorption coefficient of materials in situ using a parametric loudspeaker," in *Proc. 17th Int. Congr. Sound Vib.*, Cairo, Egypt, 2010, pp. 1–7.
[35] P. Calicchia, S. D. Simone, L. D. Marcoberardino, J. Marchal, "Near- to far-field characterization of a parametric loudspeaker and its application in non-destructive detection of detachments in panel paintings," *Applied Acoust.*, vol. 73, pp. 1296–1302, 2012.

Yuta Hatano received his B.Eng. and M.Eng. degrees in electrical and electronic engineering from Kansai University, Osaka, Japan in 2014 and 2016, respectively. He joined Fujitsu Ten Ltd., Kobe, Japan from 2016. He is currently a visiting researcher at Kansai University, Osaka, Japan. His research interests include signal processing and electro-acoustics. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and Acoustical Society of Japan (ASJ). He received the Excellent Student Paper Award in SISA 2014.

Chuang Shi received his bachelor's degree in computer and information technology from Beijing Jiaotong University, Beijing, China in 2005, master's degree in precision instrument from Tsinghua University, Beijing, China in 2008 and Ph.D. degree in electrical and electronic engineering from Nanyang Technological University, Singapore in 2013. He was a postdoctoral researcher at Nanyang Technological University, Singapore, the University of Electro-Communications, Tokyo, Japan and Kansai University, Osaka, Japan. He is currently an Associate Professor at the University of Electronic Science and Technology of China (UESTC), Chengdu, China. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), Acoustical Society of America (ASA), Institute of Electronics, Information, and Communication Engineers (IEICE) and Asia Pacific Signal and Information Processing Association (APSIPA). He has a specific research interest in applying adaptive, array and nonlinear signal processing techniques in acoustics.

Yoshinobu Kajikawa received his B.Eng. and M.Eng. degrees in electrical engineering from Kansai University, Osaka, Japan in 1991 and 1993, respectively. He received his D.Eng. degree in communication engineering from Osaka University, Osaka, Japan in 1997. He joined Fujitsu Ltd., Kawasaki, Japan in 1993 and engaged in research on active noise control. In 1994, he joined Kansai University, Osaka, Japan, where he is now a Professor. His current research interests lie in the area of signal processing for audio and acoustic systems. He is a senior member of the Institute of Electrical and Electronics Engineers (IEEE) and Institute of Electronics, Information, and Communication Engineers (IEICE) , and a member of the European Association for Signal Processing (EURASIP), Asia Pacific Signal and Information Processing Association (APSIPA), Acoustical Society of America (ASA) and Acoustical Society of Japan (ASJ). He is currently serving as an associate editor for the IET Signal Processing, the chair of Chapter Operations Committee of IEEE Kansai Section, and a member of SLA TC in APSIPA. He is also a member-at large in APSIPA. He was a vice president of IEICE Engineering and Science Society. He is the author or co-author of more than 190 articles in journals and conference proceedings, and has been responsible for 8 patents. He received the 2012 Sato Prize Paper Award from the ASJ and the Best Paper Award in APCCAS 2014.