

Low Frequency Audio Sound Field Generated by a Focusing Parametric Array Loudspeaker

Jiaxin Zhong^{ID}, Member, IEEE, Tao Zhuang^{ID}, Ray Kirby^{ID}, Mahmoud Karimi^{ID}, Xiaojun Qiu^{ID}, Haishan Zou^{ID}, and Jing Lu^{ID}, Member, IEEE

Abstract—In this paper, the audio sound field generated by a focusing parametric array loudspeaker (PAL) is investigated both numerically and experimentally. A computationally efficient method is developed to calculate the quasilinear solution of both the Westervelt and Kuznetsov equations. The accuracy of generated sound field using Westervelt equation is initially examined, and the results are then compared with those obtained using the Kuznetsov equation. It is found that the focusing on the true focal location deteriorates the prediction accuracy using the Westervelt equation due to the strong and complicated local effects. To validate the numerical results against experimental data, a 37-channel focusing PAL prototype is constructed by using a field programmable gate array, metal-oxide semiconductor field-effect transistor drivers and the pulse train signal technique. Both the simulation and experimental results show that the audio sound pressure around the true focal location is increased, and the increment becomes larger at low audio frequencies and when the focal point moves close to the radiation surface, which is attributed to the strong local effects in the near field. It has been shown that a focusing PAL can improve the poor low frequency response for a conventional PAL without focusing, and the generated audio sound decays rapidly with the distance in the far field when compared to that without focusing.

Index Terms—Parametric array loudspeaker, focusing audio sound, kuznetsov equation, pulse width modulation.

I. INTRODUCTION

PARAMETRIC array loudspeakers (PALs) have been widely used in many audio applications due to their capability of generating highly directional sound beams [1]. PALs can adopt the phased array technique to steer and focus the audio beams [2]. The steerable PAL has been studied for applications such as active noise control (ANC) and sound reproduction systems [3], [4]. The focusing PAL aims to amplify the audio sound at a small target point by focusing the ultrasonic waves [5]. It has been explored for many potential applications. For example,

Manuscript received 11 April 2022; revised 9 August 2022; accepted 31 August 2022. Date of publication 27 September 2022; date of current version 5 October 2022. The work of Tao Zhuang, Haishan Zou, and Jing Lu was supported by the National Natural Science Foundation of China under Grant 11874219. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefan Bilbao. (*Corresponding author: Haishan Zou*)

Jiaxin Zhong, Ray Kirby, and Mahmoud Karimi are with the Centre for Audio, Acoustics and Vibration, University of Technology Sydney, New South Wales 2007, Australia (e-mail: jiaxin.zhong@student.uts.edu.au; ray.kirby@uts.edu.au; mahmoud.karimi@uts.edu.au).

Tao Zhuang, Xiaojun Qiu, Haishan Zou, and Jing Lu are with the Key Laboratory of Modern Acoustics, Nanjing University, Nanjing 210008, China (e-mail: taozhuang@smail.nju.edu.cn; xjqi@nju.edu.cn; hszou@nju.edu.cn; lujing@nju.edu.cn).

Digital Object Identifier 10.1109/TASLP.2022.3209944

it was used in an ANC system to effectively reduce the global noise radiation in 2011 [6]. It was demonstrated that a virtual sound source can be remotely constructed by focusing the sound beams generated by multiple PALs [7]. A so-called “holographic whisper” was designed to render the spatial audio based on the focusing PAL [8]. The significant distortion and the poor low frequency response limit the engineering applications of the conventional PAL [9]. Although there are many preprocessing techniques proposed to eliminate the distortion [10], [11], [12], only a few studies aims to improve the low frequency response [13]. Apart from the psychoacoustic methods [13], focusing the sound beams on the target point might be an effective alternative to overcome this problem. This work aims to examine the generation of the audio sound field by a focusing PAL to improve the poor low frequency response for a conventional PAL without focusing.

The challenge to be addressed with PALs involves the accurate modelling of the demodulated sound field with a low computational cost. When a PAL generates two intensive ultrasonic beams at difference frequencies, the audio sound is demodulated due to the nonlinear interactions of ultrasound in air [1]. The Westervelt equation is usually used to describe such a nonlinear generation, and the Gaussian beam expansion (GBE) has been proposed to obtain its quasilinear solution under the paraxial approximation [14]. The well-known Khokhlov-Zabolotskaya-Kuznetsov (KZK) equation can be seen as a paraxial approximation of the Westervelt equation [15], and has also been widely used to model the PAL [16]. However, it has been demonstrated that both the GBE and the KZK equation are inaccurate at low audio frequencies, so non-paraxial models were developed to improve the prediction accuracy which apply paraxial approximations for only ultrasound [17], [18]. Recently, a spherical wave expansion (SWE) based on the Westervelt equation was proposed with a relatively low computational load and without loss of any accuracy [19]. The two-dimensional version of the SWE, i.e., a cylindrical wave expansion, was also developed to predict the audio sound generated by a phased array PAL, which improves the prediction agreement against measurements [20], [21]. It is well known the Westervelt equation cannot capture the local nonlinear effects in the near field due to the omission of Lagrangian density. Recently, an algebraic correction for the Westervelt equation was proposed to address this issue [22]. Alternatively, one may use the Kuznetsov equation which addresses exactly the second-order nonlinearity [23]. Although the finite element method has been developed to solve the equation,

the computational load and the required memory scale with the size [24], [25]. To cope with this issue, the SWE was extended to calculate the quasilinear solution based on the Kuznetsov equation [15]. Owing to its advantage, it has also been extended to accommodate the scattering effects by a rigid sphere [26]. In this paper, the SWE method will be further extended to include the focusing behavior and investigate the sound field generated by a focusing PAL.

There are a few studies related to the sound field generated by a focusing PAL. In 1983, Lucas et al. investigated the underwater sound generated by a focusing parametric acoustic array [5]. The difference-frequency wave was found to be more effectively focused than a conventional source. However, the paraxial approximation of the Westervelt equation was used in the mathematical modelling leading to the limited accuracy. Jing et al. examined the accuracy of the Westervelt equation for a focusing transducer generating the second-harmonic field [27]. It was found the error using the Westervelt equation becomes progressively larger in the near field when ka becomes smaller, where k is the acoustic wavenumber and a is the aperture size. Although ka is usually large enough in medical and underwater applications (e.g., $ka = 28$ for the difference-frequency wave in [5]), it is not the case for a PAL as the audio wavelength can be comparable or much larger than the aperture size (e.g., $ka = 0.9$ at 500 Hz with the size of 0.1 m). Besides, it is concluded that the Lagrangian density is of order $1/\sqrt{kF}$ compared to the nonlinear term, where F is the focal length [28], so the error using Westervelt equation increases as the focal point moves close to the radiation surface. Since the Westervelt equation is widely used in the modelling of a PAL, it is necessary to verify its accuracy for a focusing PAL.

There are various kinds of ways to realize a focusing PAL, such as convex and concave lens [5], phased arrays [8], and gradient index phononic crystals [29]. The phased array technique is the most flexible and versatile approach at present because the location of the focal point can be arbitrarily and electronically controlled by applying different time delays to the ultrasound signals. The PAL adopting the phased array technique can be implemented in a digital signal processor board [2]. However, the peripheral electronic components including the power amplifiers, analog-to-digital converters (ADC), and digital-to-analog converters (DAC) scale with the number of independent channels. In addition, the relatively high sample rate is required to enable a high-resolution phase control. To overcome these drawbacks, the pulse train signal technique was proposed recently [30], [31]. A pulse train signal with the fundamental frequency of the resonant frequency of the ultrasonic emitter (typically in the range from 40 kHz to 80 kHz) is generated as the carrier signal, which can be realized by inexpensive microcontrollers. By powering the metal-oxide semiconductor field-effect transistor (MOSFET) drivers with a real time audio signal, an amplitude modulation on the carrier signal is realized, and the audio sound is then generated in air. The high-resolution phase manipulation of different channels can be easily implemented by temporal shifting the pulse train without requirements of a high sample rate system. However, existing phased array PALs utilizing pulse train signals generate

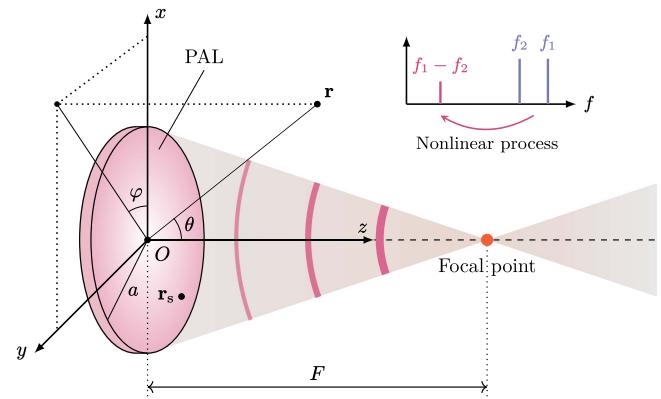


Fig. 1. Sketch of a circular PAL focusing on a focal point located on the radiation axis.

ultrasound with multiple frequencies when a pure-tone audio signal is desired. This is caused by the amplitude modulation process and is inconvenient to validate the prediction model where the ultrasound with only two frequencies is assumed to generate the single-frequency audio sound.

In this paper, the computationally efficient SWE method is extended to accommodate the focusing behavior of PALs by giving the quasilinear solution of both Westervelt and Kuznetsov equations. The accuracy of using Westervelt equation is examined first and then compared with the results obtained using the Kuznetsov equation. The simulation results of audio sound generated by a focusing PAL are presented and analyzed. In the research, a 37-channel focusing PAL prototype was constructed and realized using a field programmable gate array (FPGA) and MOSFET drivers, a modified pulse train signal was developed so that the ultrasonic transducers emitting ultrasound with only two frequencies. The experimental results with this prototype are reported to support the simulation ones.

II. THEORY

Fig. 1 shows a circular PAL focusing on a focal point with a focal length of F . The circular PAL with a radius of a is assumed to be placed in free field. For simplicity, the focal point is set on the radiation axis of the PAL to achieve the axisymmetric focusing effect. A Cartesian coordinate system (x , y , z) is established with its origin, O , at the center of the PAL, and the positive z axis pointing to the focal point. The spherical coordinate system (r , θ , φ) is established with respect to the Cartesian coordinates for further calculations, where r , θ , and φ are the radial distance, zenith angle, and azimuth angle, respectively.

When the ultrasound at the frequencies of f_1 and f_2 ($f_1 > f_2$) is radiated by the PAL, the audio sound with the frequency of $f_a = f_1 - f_2$ is demodulated in air due to the second-order nonlinearity caused by air. The excitation velocity profile on the transducer surface is assumed to be axisymmetric as

$$u(r_s, t) = u_1(r_s)e^{-i\omega_1 t} + u_2(r_s)e^{-i\omega_2 t}, \quad r_s \leq a, \quad (1)$$

where i is the imaginary unit, $r_s = \sqrt{x_s^2 + y_s^2}$ is the distance between a source point $\mathbf{r}_s = (x_s, y_s, 0)$ on the radiation surface

and the center of the PAL, the angular frequency $\omega_i = 2\pi f_i$, t is the time, $u_i(r_s)$ is the excitation velocity profile at f_i , and $i = 1$ and 2. In this paper, the focusing PAL is realized by applying a continuous profile for the ultrasound, i.e., [27]

$$u_i(r_s) = u_0 e^{-i \operatorname{Re}(k_i) \sqrt{r_s^2 + F^2}}, \quad (2)$$

where u_0 is a constant, $\operatorname{Re}(\cdot)$ denotes the real part of the argument, the wavenumber $k_i = \omega_i/c_0 + i\alpha_i$, c_0 is the sound speed, α_i is the sound attenuation coefficient at f_i , and $i = 1$ and 2. (2) ensures that the wave fronts impinging at the focal point are in-phase for ultrasound, so that the maximum constructive interference is achieved. It can simulate the radiation from a concave source [5]. The challenge of modelling a focusing PAL is to obtain the accurate numerical results with a low computational load. Existing methods adopted the KZK equation [5], which was found to be inaccurate for a PAL due to the paraxial approximations [17]. Instead, the more accurate Westervelt and Kuznetsov equations are used and their quasilinear solutions without additional approximations are obtained in the following text.

A. Governing Equations and the Transition Distance

The radiation of the PAL is governed by the Kuznetsov equation in terms of the velocity potential Φ such that $\mathbf{v} = \nabla \Phi$ [15],

$$\begin{aligned} \nabla^2 \Phi - \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2} &= -\frac{\delta}{c_0^2} \nabla^2 \frac{\partial \Phi}{\partial t} + \frac{1}{c_0^2} \frac{\partial}{\partial t} \\ &\left[(\nabla \Phi)^2 + \frac{\beta - 1}{c_0^2} \left(\frac{\partial \Phi}{\partial t} \right)^2 \right], \end{aligned} \quad (3)$$

where \mathbf{v} is the particle velocity vector, δ is the sound diffusivity parameter related to the sound attenuation coefficient, and $\beta = 1.2$ is the nonlinearity coefficient in air. The Kuznetsov equation given by (3) is used instead of the general second-order nonlinear wave equation expressed in terms of the sound pressure because the Kuznetsov equation has the advantage of avoiding the evaluation of the second second-order spatial derivatives of the ultrasonic Lagrangian density [15].

When the local effects are neglected, (3) is reduced to the Westervelt equation [23], [25],

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = -\frac{\delta}{c_0^2} \nabla^2 \frac{\partial p}{\partial t} - \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2 p^2}{\partial t^2}, \quad (4)$$

where p is the sound pressure. The Westervelt equation is easier to calculate than the Kuznetsov equation due to the omission of the Lagrangian density. However, the Westervelt equation cannot capture the complex interactions of ultrasound in the near field and the prediction accuracy is therefore deteriorated.

It was defined in [15] that the Westervelt far field is the region where the Westervelt equation is accurate enough, and the transition distance from the near field to the Westervelt far field can be determined by

$$r_0 = \frac{a^2}{\lambda_u} - \frac{\lambda_u}{4}, \quad (5)$$

where λ_u is the ultrasonic wavelength at the center frequency $f_u = (f_1 + f_2)/2$. Although the Rayleigh distance, i.e., $D = \pi a^2/\lambda_u$, is about π times larger, it is also an alternative to the transition distance given by (5) [25]. It is demonstrated the prediction error of audio sound beyond the transition distance given by (5) is less than 0.3 dB when the PAL radius is larger than 0.05 m, so (5) is more suitable for engineering applications [15]. The validity of (5) has been verified for a conventional PAL with a uniform excitation profile, and it can also be used as the transition distance for a focusing PAL as shown in Section III.

B. Spherical Wave Expansion of the Quasilinear Solution

The ultrasound field is solved first by the SWE method as [15],

$$p(\mathbf{r}, k_i) = \rho_0 c_0 \sum_{n=0}^{\infty} C_n R_n(r, k_i) P_{2n}(\cos \theta), \quad i = 1, 2, \quad (6)$$

where the coefficient and the radial component for ultrasound are [15]

$$C_n = (-1)^n (4n+1) \frac{\Gamma(n + \frac{1}{2})}{\sqrt{\pi} \Gamma(n+1)}, \quad (7)$$

$$\text{and } R_n(r, k_i) = \int_0^a u_i(r_s) j_{2n}(k_i r_s, <) h_{2n}(k_i r_s, >) k_i^2 r_s dr_s, \quad (8)$$

respectively. In addition, $\Gamma(\cdot)$ is the Gamma function, $j_{2n}(\cdot)$ is the spherical Bessel function, $h_{2n}(\cdot)$ is the spherical Hankel function of the first kind, $P_{2n}(\cdot)$ is the Legendre polynomial, $r_{s,>} = \max(r, r_s)$, and $r_{s,<} = \min(r, r_s)$. Contrary to (14) in [15], the velocity profile given by (2) is included in the radial component of ultrasound in (8), which determines the focusing behavior of the PAL. The interval of integration ranges from 0 to a in (8) which includes the aperture effect of the source. The components of the velocity vector in spherical coordinates are correspondingly

$$\begin{cases} v_r(\mathbf{r}, k_i) = -i \sum_{n=0}^{\infty} C_n \frac{d R_n(r, k_i)}{d(k_i r)} P_{2n}(\cos \theta) \\ v_\theta(\mathbf{r}, k_i) = -i \sum_{n=0}^{\infty} C_n \frac{R_n(r, k_i)}{k_i r} \frac{d P_{2n}(\cos \theta)}{d \theta} \\ v_\varphi(\mathbf{r}, k_i) = 0 \end{cases}. \quad (9)$$

The SWE of the quasilinear solution of audio sound pressure based on the Westervelt equation is obtained as [15], [19]

$$p_a(\mathbf{r}) = -i \beta \rho_0 \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_l C_m W_{lmn} \mathcal{R}_{lmn}(r) P_{2n}(\cos \theta). \quad (10)$$

where the coefficient involving the Wigner-3j symbol is

$$W_{lmn} = (4n+1) \begin{pmatrix} 2l & 2m & 2n \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (11)$$

and the radial component for audio sound is

$$\begin{aligned} \mathcal{R}_{lmn}(r) &= \int_0^\infty R_l(r_v, k_1) R_m^*(r_v, k_2) j_{2n}(k_a r_v, <) \\ &\times h_{2n}(k_a r_v, >) \times k_a^3 r_v^2 dr_v, \end{aligned} \quad (12)$$

with the superscript “*” denoting the complex conjugate, $r_{v,>} = \max(r, r_v)$, and $r_{v,<} = \min(r, r_v)$.

The quasilinear solution of audio sound pressure based on the Kuznetsov equation can be seen as the superposition of the sound field generated by the cumulative, $p_{\text{cum}}(\mathbf{r})$, and local (non-cumulative) effects, $p_{\text{local}}(\mathbf{r})$,

$$p_a(\mathbf{r}) = p_{\text{cum}}(\mathbf{r}) + p_{\text{local}}(\mathbf{r}), \quad (13)$$

where $A_{lmn} = l(2l+1) + m(2m+1) - n(2n+1)$, [15]

$$\begin{aligned} p_{\text{cum}}(\mathbf{r}) = i\rho_0 \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_l C_m W_{lmn} P_{2n}(\cos \theta) \\ \times [(1-\beta) \mathcal{R}_{lmn}(r) - \mathcal{R}_{r,lmn}(r) \\ - A_{lmn} \mathcal{R}_{\theta,lmn}(r)], \end{aligned} \quad (14)$$

$$\text{and } p_{\text{local}}(\mathbf{r}) = \frac{p(\mathbf{r}, k_1)p^*(\mathbf{r}, k_2)}{2\rho_0 c_0^2} - \frac{\rho_0}{2} \mathbf{v}(\mathbf{r}, k_1) \cdot \mathbf{v}^*(\mathbf{r}, k_2), \quad (15)$$

with the radial components in the radial and the zenithal directions respectively expressed by

$$\begin{aligned} \mathcal{R}_{r,lmn}(r) = \int_0^{\infty} \frac{dR_l(r_v, k_1)}{d(k_1 r_v)} \frac{dR_{2,m}^*(r_v, k_2)}{d(k_2^* r_v)} j_{2n}(k_a r_{v,<}) \\ \times h_{2n}(k_a r_{v,>}) k_a^3 r_v^2 dr_v, \end{aligned} \quad (16)$$

$$\begin{aligned} \text{and } \mathcal{R}_{\theta,lmn}(r) = \int_0^{\infty} R_l(r_v, k_1) R_m^*(r_v, k_2) j_{2n}(k_a r_{v,<}) \\ \times h_{2n}(k_a r_{v,>}) \frac{k_a^3}{k_1 k_2^*} dr_v. \end{aligned} \quad (17)$$

The sound field generated by the local effects given by (15) is determined only by the ultrasound, which can be seen as an algebraic correction to the sound pressure obtained by the cumulative effects given by (14). When the ultrasound behaves like plane waves, it can be observed that (15) is zero so that the local effects are negligible [15], [25]. For a focusing PAL, however, the sound field is intensive and complicated around the focal point indicating that the plane wave condition is not satisfied.

III. SIMULATIONS

Numerical simulations are conducted on MATLAB R2020b. The center frequency of the ultrasound is set as $f_u = 40$ kHz, which is the resonant frequency of the ultrasonic emitter used in experiments. The sound attenuation coefficient, α_u , at 40 kHz is approximately 0.15 Np/m calculated according to ISO 9613-1 with the relative humidity 70% and temperature 20 °C. The radius of the PAL is set as 0.1 m, which is the same as that in experiments. The transition distance from the near field to the Westervelt far field for the audio sound is 1.16 m calculated by (5). The Rayleigh distance is 3.66 m. The vibration velocity profile constant u_0 is set as 0.12 m/s for all cases. The on-surface ultrasound pressure is then $\rho_0 c_0 u_0 = 50$ Pa (125 dB), which is a common value in applications. The reference quantity for the sound pressure level (SPL) is 20 μPa. The accuracy of

Westervelt equation is examined in Section III.A and the sound field generated by a focusing PAL is analyzed in Section III.B.

A. Accuracy of the Westervelt Equation

Fig. 2 compares the SPL for audio sound generated by a PAL using the Westervelt and Kuznetsov equations at 250 Hz, 500 Hz, and 1 kHz, with a focal length of 0.2 m and 2 m and without focusing. The dashed line in the figures denotes the transition distance (1.16 m) from the near field to the Westervelt far field. When the wave propagation is modelled by the Westervelt equation, the audio sound can be seen as the generation of an infinitely large virtual volume source with the source density proportional to the product of two ultrasound pressure $p(\mathbf{r}, k_1)p^*(\mathbf{r}, k_2)$. The integration over such a virtual source represents the cumulative effects, which smooth the calculated sound pressure as shown in Fig. 2 [15], [25]. It is observed that the difference obtained using two equations is large in the near field but negligible in the Westervelt far field for all cases. This can be explained by analyzing the axial sound pressure and the Lagrangian density of the ultrasound at 40 kHz, as shown in Fig. 3. When the Kuznetsov equation is used, the Lagrangian density of ultrasound, representing the local (non-cumulative) effects, is considered, and can be seen as an algebraic correction to the results obtained using the Westervelt equation [15], [25]. Because the ultrasound significantly fluctuates in the near field as shown in Fig. 3, highly oscillatory audio sound pressure can be observed in the near field as shown in Fig. 2. The transition distance is the location of the first maxima of the ultrasound SPL [15], where the Lagrangian density also reaches around its maximum. Therefore, the difference obtained using two equations is negligible beyond the transition distance, which is termed the Westervelt far field in [15].

In the near field, the ultrasound pressure is large and fluctuates significantly, so the local effects are strong which cannot be captured by the Westervelt equation. In the Westervelt far field, the Westervelt equation is accurate enough even for the focusing PAL. For example, the SPL difference obtained using the two equations is less than 0.1 dB and 0.4 dB at 250 Hz when the focal length is 0.2 m and 2 m, respectively. For a focusing PAL, the source density of audio sound is increased around the focal point as it is proportional to the product of the local ultrasound pressure [15]. The cumulative effects captured by the Westervelt equation are therefore increased. For example, the SPL obtained using the Westervelt equation at 500 Hz increases from 51.8 dB to 60.8 dB after the focusing on the focal point of 0.2 m. However, it is interesting to note that the focusing behavior influences more on the local effects, resulting in a large increment in the SPL difference obtained using two equations near the focal point. For example, the SPL difference at 0.2 m is increased from 0.2 dB to 17.3 dB at 500 Hz after the focusing. At lower audio frequencies, this increment is more prominent. As the frequency decreases from 1 kHz to 500 Hz and 250 Hz, the SPL difference increases from 3.4 dB to 17.3 dB and 29.3 dB at 0.2 m, respectively.

When the focal length is 0.2 m, it can be found in Fig. 2(b), (e), and (h) that there is a peak at 0.2 m in SPL for both equations which corresponds to the focal point. However, when the PAL

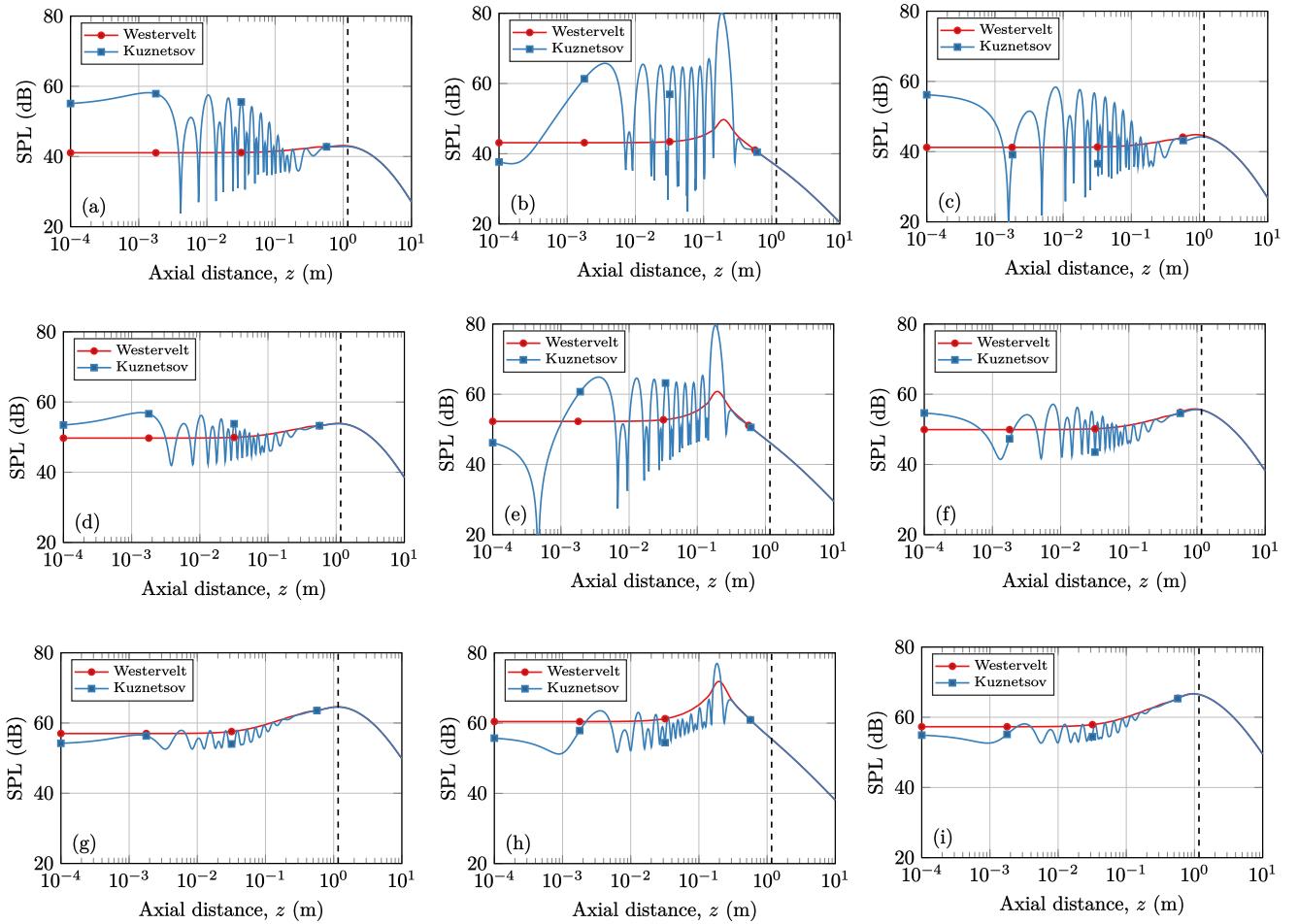


Fig. 2. Axial audio SPL obtained using the Westervelt and Kuznetsov equations without focusing (left column) and with a focal length of 0.2 m (middle column) and 2 m (right column) at 250 Hz (top row), 500 Hz (middle row), and 1 kHz (button row). Red circle, Westervelt equation; blue square, Kuznetsov equation; dashed line, the transition distance.

is focusing on 2 m which is beyond the transition distance, Fig. 2(c), (f), and (i) show that the SPL trend is like that of without focusing. It is observed in Fig. 3 that the ultrasound pressure reaches the last maxima around 0.8 m instead of 2 m. This is known as the focal shift phenomenon due to the diffraction limit, and the shift can be estimated by (15) in [32]. The focusing PAL is unable to focusing the ultrasound beyond the transition distance. In the limiting case when the focal length is infinity, the ultrasound pressure reaches the last maxima at the transition distance given by (5) [15]. Therefore, the audio SPL difference obtained using two equations would be insignificant beyond the transition distance even if the focal point is located there. For example, as shown in Fig. 2 when the focal length is 2 m, the SPL difference beyond the transition distance is below 0.4 dB, 0.2 dB, and 0.1 dB at 250 Hz, 500 Hz, and 1 kHz, respectively. It is concluded that although the transition distance given by (5) was obtained based on the analysis for a conventional PAL in [15], it is still valid for a focusing PAL. It means the Kuznetsov equation must be used when the field point is located before the transition distance, and the results presented in [5] cannot be used for investigating a focusing PAL where the paraxial approximation of the Westervelt equation was adopted.

Therefore, all the subsequent simulations are obtained using the Kuznetsov equation.

B. Audio Sound Generated by a Focusing PAL

The audio sound generated by a conventional PAL without focusing and a focusing PAL with a focal length of 0.2 m and 2 m at 250 Hz, 500 Hz, and 1 kHz are presented in Fig. 4. It is found that the audio beam generated by a conventional PAL is narrow and decays slowly along the radiation axis in all frequencies. For a focusing PAL with a focal length of 0.2 m, the energy is focused on the focal point. Beyond this point, the audio beam decays more rapidly than that without focusing. It shows an energy transfer from the far field to the near field. This is a merit for some applications of PALs. For a conventional PAL, the reflection of the audio sound is strong due to slow decay rate and its interference with the incident sound is annoying when it is used in a room [33]. For a focusing PAL with a focal length of 2 m, the audio sound is slightly amplified when compared to that generated by a conventional PAL. The reason is that the ultrasound wave is ineffectively focused beyond the transition

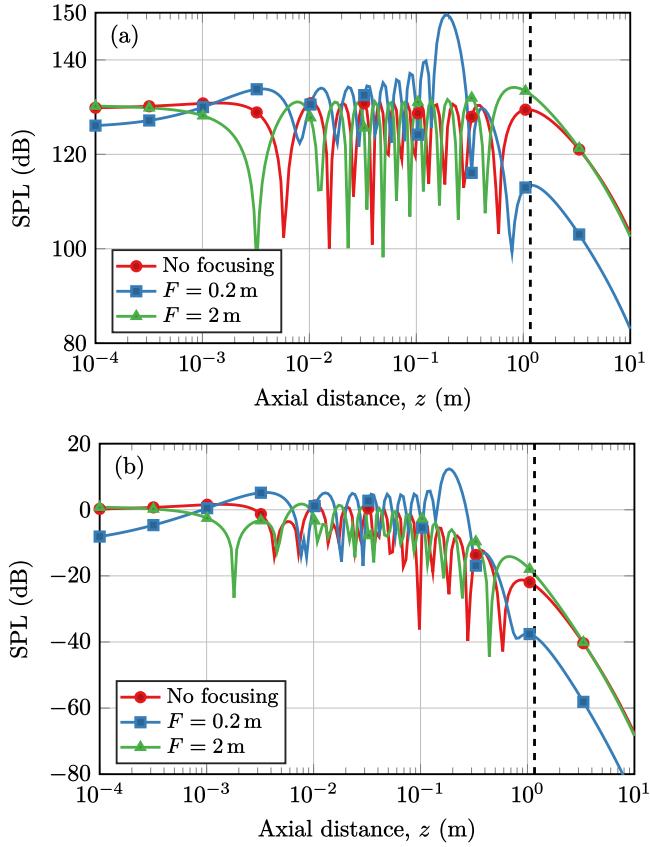


Fig. 3. Axial (a) SPL and (b) the level of normalized Lagrangian density at 40 kHz, with a focal length of 0.2 m and 2 m, and without focusing. The dashed line denotes the transition distance.

distance, which is 1.16 m calculated by (5). Instead, the sound pressure at the true focal location (0.8 m) is increased.

Fig. 5 shows the audio SPL on the radiation axis ($\theta = 0$) generated by a PAL with and without focusing at 250 Hz, 500 Hz, 1 kHz, and 2 kHz. In the near field, the audio sound field fluctuates with the axial distance for all cases. The audio sound around the focal point is generally amplified when compared to that without focusing. For example, when the PAL is focused at 0.3 m, the SPL is increased from 42.1 dB to 64.8 dB at the focal point when the audio frequency is 250 Hz. It is observed that as the focal point becomes closer to the radiation surface, the SPL increment around the focal point becomes larger. For example, the SPL increment increases from 1.9 dB to 41.0 dB when the focal point changes from 1 m to 0.2 m at 250 Hz. This is because the focusing gain, defined by the ratio of the Rayleigh distance to the focal length, $G = D/F$, becomes larger resulting in a larger ultrasound field at the focal point [32]. Therefore, both the cumulative and local effects given by (14) and (15) are more significant around the focal point. It is also noted that the amplification is more prominent at lower audio frequencies. The reason is that the audio sound pressure determined by the cumulative effects given by (14) drops about 12 dB as the audio frequency is halved [15], while that determined by the local effects given by (15) changes a little. Although the contribution

from the cumulative effects is neglected, that from the local effects is retained.

In the Westervelt far field, there is no fluctuations for the audio SPL, and it decreases as the axial distance increases. It can be found the audio sound generated by a focusing PAL is smaller than that by a conventional PAL without focusing. As the focal point becomes closer to the radiation surface, the audio SPL is smaller in the Westervelt far field. For example, the audio SPL generated by a conventional PAL without focusing is 61.1 dB at 10 m and 2 kHz, while it respectively decreases to 46.0 dB, 50.4 dB, 55.0 dB, and 59.0 dB, when the focal length is 0.2 m, 0.3 m, 0.5 m, and 1 m. The ultrasound pressure on the radiation axis ($\rho = 0$) and the focal plane ($z = F$) can be respectively approximated by [34]

$$p(\rho = 0, z, k_i) = \frac{\rho_0 c_0 u_0}{1 - z/F} \left[1 - e^{iD(1/z - 1/F)} \right] e^{ik_i z}, \quad (18)$$

$$\text{and } p(\rho, z = F, k_i) = -i\rho_0 c_0 u_0 G \text{jinc}(k a \rho / F) e^{ik_i F}, \quad (19)$$

where $\text{jinc}(x) = 2J_1(x)/x$ is the jinc function, and $J_1(x)$ is the first order Bessel function. As the focal point becomes closer to the radiation surface, the focal length decreases and the focusing gain $G = D/F$ becomes larger. Therefore, the ultrasound pressure in the far field decreases according to (18), while it increases in the focal plane according to (19). It demonstrates that more energy is transferred from the far field to the near field as the focusing gain becomes larger.

Fig. 6 shows the SPL increment, which is defined as the SPL with focusing minus the SPL without focusing, at the focal point after the focusing as a function of the focal length at different audio frequencies. The SPL increment generally increases with some fluctuations as the focal point moves close to the PAL. The increment is smaller at higher audio frequencies. The reason is that the audio sound pressure caused by the cumulative effects increases by approximately 12 dB as the audio frequency is doubled [9], while the local effects change a little at different audio frequencies. Consequently, the influence of the local effects is weaker at higher audio frequencies. As the focal length increases, the SPL increment decreases rapidly at all frequencies resulting from smaller focusing gains. When the focal point is beyond the transition distance, the ultrasound is inefficiently focused on the focal point, and the SPL increment decreases and approaches to 0 dB. In the limiting case when the focal length is infinity, the radiation can be seen as a conventional PAL without focusing. The SPL increment can even be negative especially at low audio frequencies indicating that the SPL is decreased at the focal point after the focusing. For example, the SPL increment is -6 dB when the focal length is 0.5 m at 250 Hz, and the SPL as a function of the axial distance can be found in Fig. 5(a) for this case. This phenomenon is caused by the destructive interference between the cumulative and local effects as can be seen by (13). At small focal lengths, the local effects are dominant, while at large focal lengths, the cumulative are dominant. When the cumulative and local effects are comparable for a moderate focal

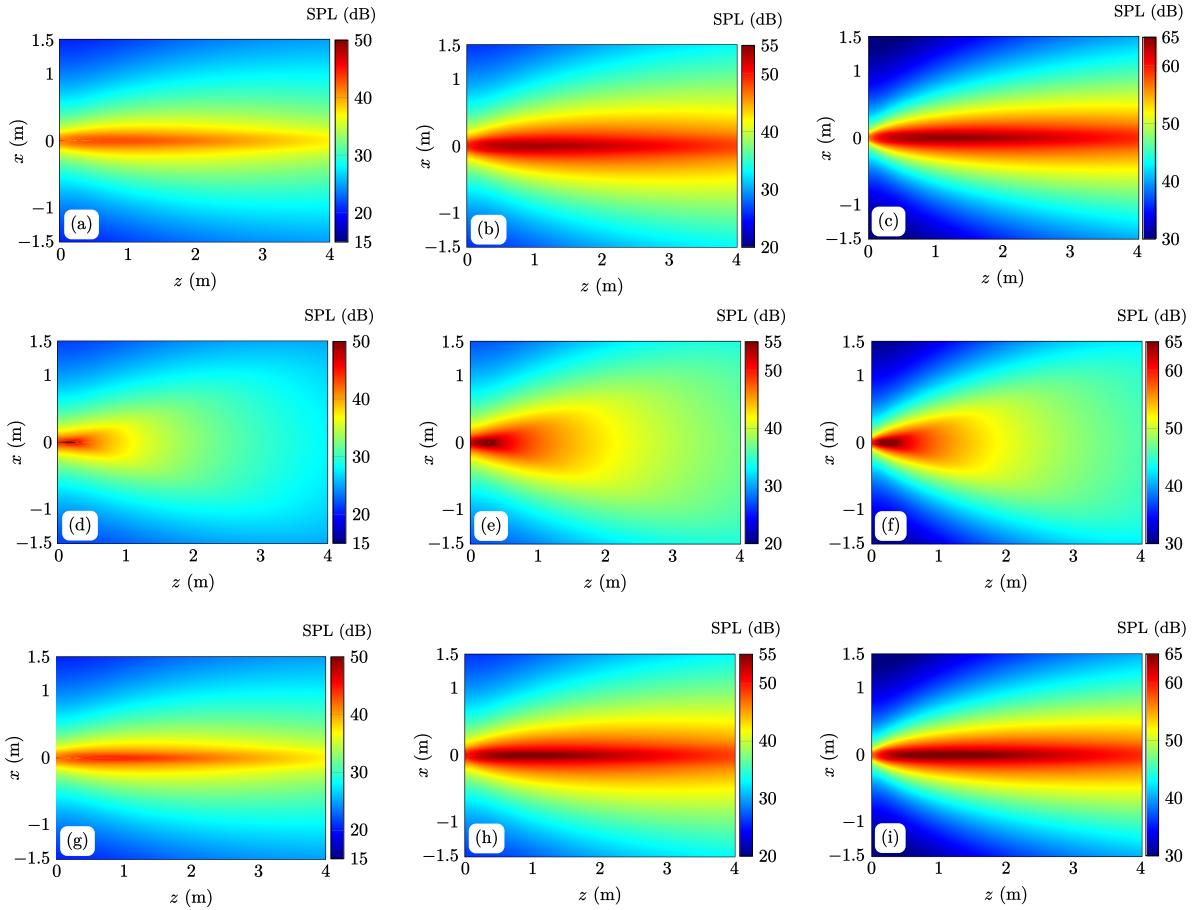


Fig. 4. SPL distribution for audio sound generated by a conventional PAL without focusing (top row) and a focusing PAL with a focal length of 0.2 m (middle row) and 2 m (bottom row) on its radiation axis, at 250 Hz (left column), 500 Hz (middle column), and 1 kHz (right column).

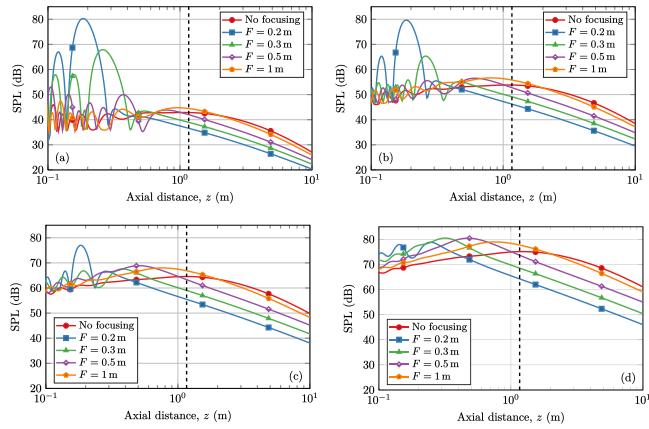


Fig. 5. Axial SPL generated by a PAL with and without focusing at (a) 250 Hz, (b) 500 Hz, (c) 1 kHz, and (d) 2 kHz. The dashed line denotes the transition distance.

length, the destructive interference between them can happen resulting in an audio SPL decrement.

The conventional PAL suffers from the poor low frequency response because the SPL decreases about 12 dB as the frequency is halved [9]. The above analysis shows the low frequency

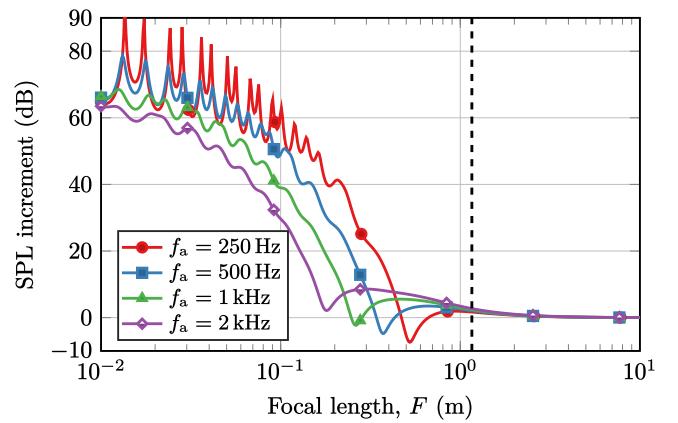


Fig. 6. The SPL increment at the focal point after the focusing as a function of the focal length at 250 Hz, 500 Hz, 1 kHz, and 2 kHz. The dashed line denotes the transition distance.

audio sound can be amplified near the listener by using the focusing PAL with the focal point being around the listener. The mechanism is that focusing the ultrasound on the listener improves both the cumulative and local effects, so the audio sound field is amplified. However, it should be noted that there

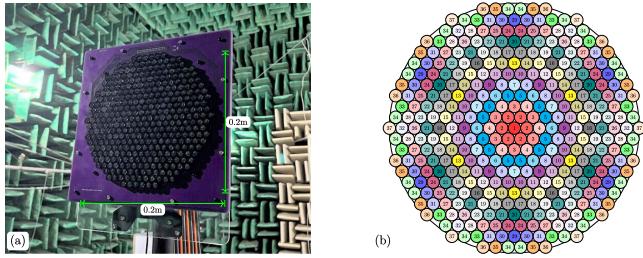


Fig. 7. The prototype of a 37-channel circular focusing PAL with a radius of 0.1 m consisting of 367 circular ultrasonic emitters with a radius of 5 mm: (a) Photo; (b) Sketch, where the number inside the small circles denotes the channel number.

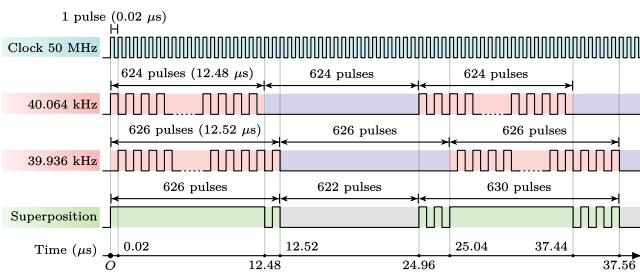


Fig. 8. Timing diagram of signals for generating the audio sound at 128 Hz.

exists a focal shift phenomenon which means the true focal location with the maximal sound pressure shifts relative to the geometric focal point. In addition, the audio SPL at the focal point might be decreased at the focal point after the focusing. Therefore, the focusing PAL should be designed according to a specific audio application.

IV. EXPERIMENTS

A. Prototype of a Focusing PAL

To validate the above findings by experiments, a prototype of a circular focusing PAL with a radius of 0.1 m was fabricated as shown in Fig. 7. The prototype consists of 367 circular ultrasonic emitters (Murata MA40S4S, Kyoto, Japan) which has a resonance frequency of 40 kHz and a radius of 5 mm. The ultrasonic emitters are arranged in a compact hexagonal array as shown in Fig. 7(b) to minimize the interelement separation. To approximate the axisymmetric phase distribution as assumed in the theory, the emitters with the same distance to the centroid of the PAL are grouped together as a channel and assigned the same phase. There are 37 channels in total and the channel number is marked inside each emitter in Fig. 7(b). The distance between the centroid of the PAL and all emitters allocated to a same channel, as well as the number of emitters for each channel, are listed in Table I.

The focusing PAL is constructed with an FPGA (Xilinx XC7A100T, San Jose, CA) and MOSFET drivers (Microchip MIC4127, Chandler, AZ). The logical diagram of the circuits is the same as Fig. 8 in [31]. The FPGA is programmed to generate 37-channel independent rectangular pulse signals which are fed into the MOSFET drivers. The output of them is connected to the positive pins of ultrasonic emitters. The generation of the

TABLE I
THE NUMBER OF EMITTERS IN EACH CHANNEL AND THE DISTANCE TO THE CENTROID OF THE PAL

Channel number	Number of emitters	Distance to the centroid ($\times 5$ mm)
1	1	0
2	6	2
3	6	$2\sqrt{3} \approx 3.464$
4	6	4
5	12	$2\sqrt{7} \approx 5.292$
6	6	6
7	6	$4\sqrt{3} \approx 6.928$
8	12	$2\sqrt{13} \approx 7.211$
9	6	8
10	12	$2\sqrt{19} \approx 8.718$
11	12	$2\sqrt{21} \approx 9.165$
12	6	10
13	6	$6\sqrt{3} \approx 10.392$
14	12	$4\sqrt{7} \approx 10.583$
15	12	$2\sqrt{31} \approx 11.136$
16	6	12
17	12	$2\sqrt{37} \approx 12.166$
18	12	$2\sqrt{39} \approx 12.490$
19	12	$2\sqrt{43} \approx 13.115$
20	6	$8\sqrt{3} \approx 13.856$
21	18	14
22	12	$4\sqrt{13} \approx 14.422$
23	12	$2\sqrt{57} \approx 15.100$
24	12	$2\sqrt{61} \approx 15.620$
25	12	$6\sqrt{7} \approx 15.875$
26	6	16
27	12	$2\sqrt{67} \approx 16.371$
28	12	$2\sqrt{73} \approx 17.088$
29	6	$10\sqrt{3} \approx 17.321$
30	12	$4\sqrt{19} \approx 17.436$
31	12	$2\sqrt{79} \approx 17.776$
32	6	18
33	12	$4\sqrt{21} \approx 18.330$
34	24	$2\sqrt{91} \approx 19.079$
35	12	$2\sqrt{93} \approx 19.287$
36	12	$2\sqrt{97} \approx 19.698$
37	6	20

pulse signal for one channel is illustrated in Fig. 8, where two ultrasonic frequencies at 40.064 kHz and 39.936 kHz are used to generate an audio sound wave at 128 Hz. The FPGA has a clock rate of 50 MHz, which means the clock cycle (the duration of one pulse) is 0.02 μs.

Based on the clock signal, two signals are generated first. The first signal aims to radiate the ultrasound at 40.064 kHz. A sequence of rectangle signals is generated with the fundamental frequency of 25 MHz, and the total duration is 12.48 μs. Followed by them is a sequence of logic low lasting for 12.48 μs.

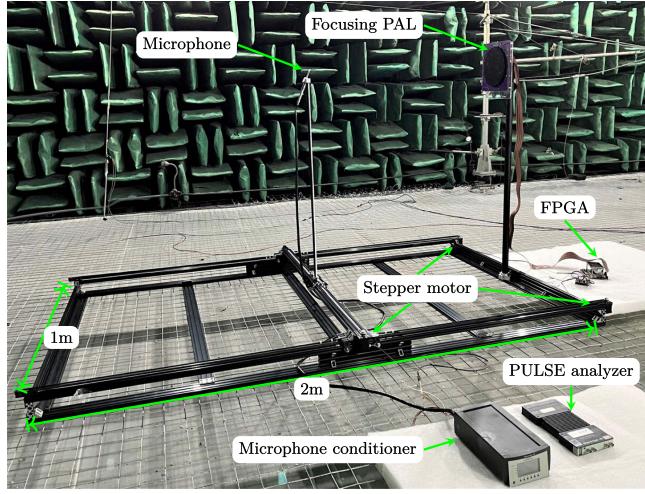


Fig. 9. Photo of the experiment setup.

This signal has the spectral components at 40.064 kHz ($= 50$ MHz / 1248) and 25 MHz with higher harmonic components. Because the ultrasonic emitters have narrow bandwidths centered at the resonant frequency, only the frequency component at 40.064 kHz is radiated by the emitter [31]. The same process is applied for the second signal aiming to radiate the ultrasound at 39.936 kHz. The superposition of the above two signals is fed into the ultrasonic emitters, and the ultrasound waves at two frequencies of 40.064 kHz and 39.936 kHz are therefore generated. The amplitude of the generated sound can be controlled by the input voltage for MOSFET drivers. The phase difference between channels can be realized by the time shifting of signals with a high resolution of only $0.02 \mu\text{s}$ (50 MHz). The advantage of this technique is that no digital-analog converters and the implementation of signal processing algorithms are required. The FPGA can also be replaced by inexpensive microcontrollers [30], [31].

B. Experimental Setup

The experiments were conducted in a full anechoic room at Nanjing University with dimensions of $11.4 \text{ m} \times 7.8 \text{ m} \times 6.7 \text{ m}$ (height). The relative humidity and temperature were 68% and 13°C , respectively. The photo of the experiment setup is presented in Fig. 9. A condenser microphone Brüel & Kjær Type 4135 was used to measure the sound pressure. The signal is conditioned by a Brüel & Kjær Type 2690 conditioner and analyzed by a PULSE analyzer (Brüel & Kjær Type 3160). To avoid spurious sound induced by the intensive ultrasound, the microphone was covered by a piece of small and thin plastic film. The preliminary experimental results show the ultrasound level at the focal point is around 127 dB without using this acoustic filter, when the focal length is 0.2 m. The measured result decreases to around 95 dB with the acoustic filter. Therefore, a large portion of the spurious sound is removed with this acoustic filter [14]. It is noted that this filter also results in an insertion loss of about 3 dB for an audio sound wave at 1 kHz. As shown in Fig. 9, the sound field in a rectangular area was measured with dimensions of $1 \text{ m} \times 2 \text{ m}$. The PAL surface is perpendicular to

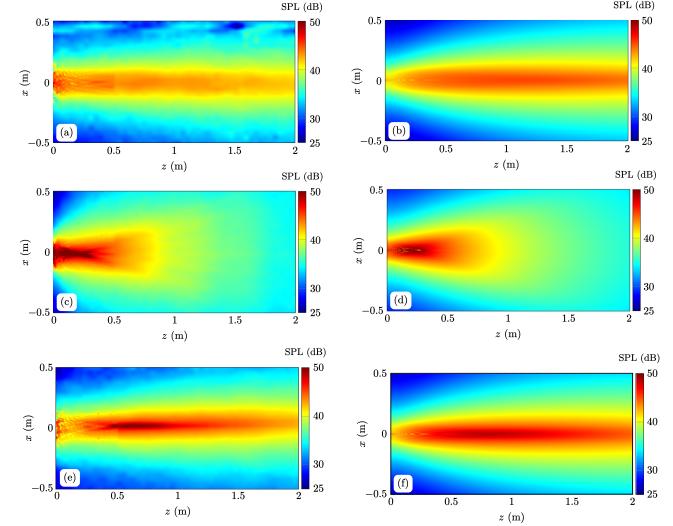


Fig. 10. Measurements (left column) and simulation results (right column) of the audio sound field at 512 Hz, without focusing (top row), and with a focal length of 0.2 m (middle row) and 1 m (bottom row).

the measurement plane. The microphone was mounted at the same height of the PAL center, and it is scanned by three stepper motors which are controlled by personal computers.

C. Experimental Results

The measurement results in the rectangular area and the corresponding simulation results at 512 Hz without focusing and with a focal length of 0.2 m and 1 m are presented in Fig. 10. Since the sound field is more complicated in the near field, the spacing between the measurement points was set to 1 cm in the region of $-0.15 \text{ m} \leq x \leq 0.15 \text{ m}$ and $0 \leq z \leq 0.5 \text{ m}$, while the spacing of 4 cm was chosen in other regions. The predictions are in good agreement with the measured sound field.

The audio sound is highly directional for a conventional PAL without focusing, and it is focused in the near field after the focusing. It is noted the location of the true focal location which has the maximal SPL when the focal length is 1 m is shifted to about 0.66 m. This is due to the focal shift phenomenon and can be estimated from (15) in [32]. However, some discrepancies between measurements and simulation results are observed at the location close to the radiation surface and the true focal location. This might be caused by the spatial aliasing effect as the size of the ultrasonic emitters (1 cm) exceeds one half of the ultrasonic wavelength (0.43 cm). In addition, misalignments and individual deviations of magnitude and phase responses of ultrasonic emitters may also affect the generated sound field [31].

Fig. 11 compares the measurements and simulation results of the audio SPL along the radiation axis ($\theta = 0$) at 512 Hz and 1024 Hz with a focal length of 0.2 m, and at 512 Hz with a focal length of 1 m. In Fig. 11(a) and (b), although the focusing on the focal point is well predicted by the simulation, the measurements are lower and smoother than predictions. This can be attributed to the fact that the sound waves are insufficiently focused due

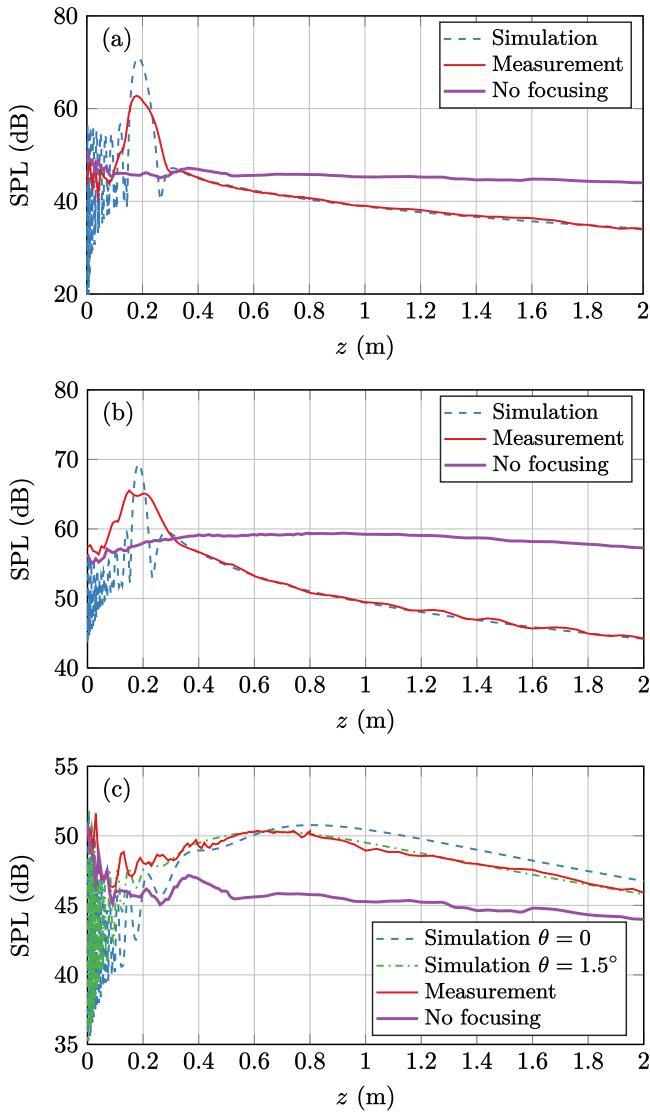


Fig. 11. Comparison of measurements and simulation results of audio SPL along the radiation axis (a) at 512 Hz with a focal length of 0.2 m; (b) at 1024 Hz with a focal length of 0.2 m; (c) at 512 Hz with a focal length of 1 m. The thick solid lines represent the measurements of the audio SPL without focusing.

to the discretized shaded phased array rather than a continuous profile given by (2). The acoustic filter used for removing the spurious sound also reduces audio component resulted from the local effects to some extent due to the reduction of the ultrasound pressure at the measurement point, but this doesn't affect the main discovery of the work. Moreover, the highly oscillating behavior of the sound field close to the radiation surface were not fully captured in the measured data which might because the size of the microphone (0.64 cm) is larger than the spacing of adjacent peaks of the predicted curve. The generated sound field changes rapidly even with a small deviation from the radiation axis. Fig. 11(c) shows that the predictions at $\theta = 1.5^\circ$ agree better with measurements than the predictions at $\theta = 0$ representing the radiation axis. It means the imperfect positioning of the PAL and the microphone could result in large measurement errors. If the measurement plane is imperfectly perpendicular to the PAL

surface, the true focal point deviates more from the measurement plane as the focal point moves farther away, so the measurement error is larger with a larger focal length. Nevertheless, a focusing PAL has been made in the laboratory, and the audio sound field can be well predicted by the methods presented in this paper.

For comparison, the measured audio SPLs generated by a conventional PAL without focusing are also presented in Fig. 11. It can be observed that the audio SPL around the focal point is increased in all cases. For example, Fig. 11(b) shows that the measured increment of the audio SPL at the focal point (0.2 m) after the focusing is 7.1 dB at 1024 Hz. When the audio frequency decreases to 512 Hz, as shown in Fig. 11(a), the measured SPL increment at the focal point increases to 15.7 dB. The increment due to focusing is larger at lower audio frequencies. As discussed in Section III, this is because the audio sound pressure determined by cumulative effects drops about 12 dB as the audio frequency is halved, while that determined by local effects changes only a little at different audio frequencies. Consequently, the influence of the local effects is more significant at lower audio frequencies.

When the focal point moves to 1 m, as shown in Fig. 11(c), measurement shows the SPL increment is smaller and the maximal increment is only 4.7 dB at 0.6 m. The improvement of the audio sound response is larger when the focal point moves closer to the radiation surface. However, the sound pressure at almost all points ranging from 0 to 2 m is amplified, which agrees with the simulation results presented in Fig. 5. It is also observed that the audio SPL decays more rapidly at large distances after focusing. For example, the audio SPL at 2 m decreases from 44.0 dB and 57.3 dB, to 33.9 dB and 44.2 dB, at 512 Hz and 1024 Hz, respectively, after focusing with a focal length of 0.2 m. This can be an advantage in applications because the reflections of the audio beam would be smaller if there exists a reflecting surface behind the listener.

It is noted that there are safety concerns on intensive ultrasonic exposures caused by the PAL. The exposure would be more intensive around the focal point. Existing guidelines for the safe usage of airborne ultrasound cannot be readily applied for a PAL because the ultrasound waves emitted by the PAL has a relatively narrow spectral band [1]. Some studies pointed out that the temporary hearing loss should not occur for levels below 140 dB when exposed to the sufficiently high frequency airborne ultrasound [35], but the specific allowed ultrasound exposure level for the PAL is still unclear at present. The ultrasound level is measured around 127 dB at the focal point for the cases in Fig. 11(a) and (b), where the focal length is 0.2 m. In real applications, one can reduce the focusing gain and/or the ultrasound pressure level to reduce the ultrasound exposure.

V. CONCLUSION

In this paper, the audio sound field generated by a focusing PAL is investigated by both simulations and experiments. A computationally efficient method was developed to calculate the quasilinear solution of both the Westervelt and Kuznetsov equations. The audio SPL calculated using the Westervelt equation is found to be inaccurate in the near field for a focusing PAL.

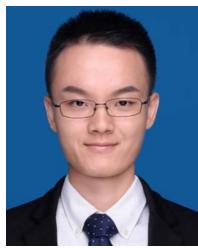
The transition distance from the near field to the Westervelt far field can be determined by (5). It was found that the focal shift phenomenon in ultrasonic beams results in a similar shift in audio beams. Furthermore, the focusing around the true focal location deteriorates the prediction accuracy using the Westervelt equation due to the strong and complex local effects there. Both the simulation and experimental results demonstrate that the audio SPL around the true focal location is increased, and the increment is larger at low audio frequencies and when the focal point moves close to the radiation surface. The audio sound pressure increment can be attributed to the increment of both the cumulative and non-cumulative (local) effects. The source density of the virtual audio source is increased around the focal point, so the audio sound pressure contributed from the cumulative effects is increased. In addition, the ultrasound Lagrangian density characterizing the local effects is increased due to the focusing behavior, so the audio sound around the focal point is furtherly amplified. It is concluded that the focusing PAL has the advantage of improving the poor frequency response for a conventional PAL without focusing. The methods and results presented in this work provide a guidance and improve the understanding of the focusing PAL.

The nonlinear wave equation is solved in the frequency domain in this work, so the pure-tone audio sound field is considered. To verify the theoretical model, a modified pulse train signal was designed to ensure the transducers emit ultrasound at only two frequencies. In real applications when a wide band audio signal is expected to be present, the time-varying audio signal after amplification can be fed into the MOSFET drivers as the input voltage, instead of a constant voltage in this work. A 40 kHz pulse train signal is fed into the input pin of the MOSFET drivers, so that the audio signal is amplitude modulated by an ultrasonic signal at 40 kHz. Consequently, the wide band audio signal can be parametrically demodulated in air. The phase of each channel can also be tuned by shifting the pulse train signal. The analysis in the frequency domain in this work provides a guidance for the applications with wide band audio signals in future.

REFERENCES

- [1] W.-S. Gan, J. Yang, and T. Kamakura, "A review of parametric acoustic array in air," *Appl. Acoust.*, vol. 73, no. 12, pp. 1211–1219, 2012.
- [2] W.-S. Gan, J. Yang, K.-S. Tan, and M.-H. Er, "A digital beamsteerer for difference frequency in a parametric array," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 3, pp. 1018–1025, May 2006.
- [3] N. Tanaka and M. Tanaka, "Active noise control using a steerable parametric array loudspeaker," *J. Acoustical Soc. Amer.*, vol. 127, no. 6, pp. 3526–3537, 2010.
- [4] C. Shi, Y. Kajikawa, and W.-S. Gan, "An overview of directivity control methods of the parametric array loudspeaker," *APSIPA Trans. Signal Inf. Process.*, vol. 3, pp. 1–12, 2014.
- [5] B. G. Lucas, J. N. Tjøtta, and T. G. Muir, "Field of a parametric focusing source," *J. Acoustical Soc. Amer.*, vol. 73, no. 6, pp. 1966–1971, 1983.
- [6] N. Tanaka and M. Tanaka, "Mathematically trivial control of sound using a parametric beam focusing source," *J. Acoustical Soc. Amer.*, vol. 129, no. 1, pp. 165–172, 2011.
- [7] Y. Ogami, M. Nakayama, and T. Nishiura, "Virtual sound source construction based on radiation direction control using multiple parametric array loudspeakers," *J. Acoustical Soc. Amer.*, vol. 146, no. 2, pp. 1314–1325, 2019.
- [8] Y. Ochiai, T. Hoshi, and I. Suzuki, "Holographic whisper: Rendering audible sound spots in three-dimensional space by focusing ultrasonic waves," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 4314–4325.
- [9] 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. Acoustical Soc. Amer.*, vol. 73, no. 5, pp. 1532–1536, 1983.
- [10] K. C.-M. Lee and W.-S. Gan, "Bandwidth-efficient recursive pth-order equalization for correcting baseband distortion in parametric loudspeakers," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 706–710, Mar. 2006.
- [11] 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, Dec. 2014.
- [12] Y. Hatano, C. Shi, and Y. Kajikawa, "Compensation for nonlinear distortion of the frequency modulation-based parametric array loudspeaker," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 8, pp. 1709–1717, Aug. 2017.
- [13] F. A. Karnapi, W. S. Gan, and M.-H. Er, "Method to enhance low frequency perception from a parametric array loudspeaker," *Audio Eng. Soc. Conv.*, vol. 112, pp. 1–5, 2002.
- [14] P. Ji and J. Yang, "An experimental investigation about parameters' effects on spurious sound in parametric loudspeaker," *Appl. Acoust.*, vol. 148, pp. 67–74, 2019.
- [15] J. Zhong, R. Kirby, and X. Qiu, "The near field, Westervelt far field, and inverse-law far field of the audio sound generated by parametric array loudspeakers," *J. Acoustical Soc. Amer.*, vol. 149, no. 3, pp. 1524–1535, 2021.
- [16] P. Ji, E.-L. Tan, W.-S. Gan, and J. Yang, "A comparative analysis of pre-processing methods for the parametric loudspeaker based on the Khokhlov–Zabolotskaya–Kuznetsov equation for speech reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 4, pp. 937–946, May 2011.
- [17] M. Červenka and M. Bednářík, "Non-paraxial model for a parametric acoustic array," *J. Acoustical Soc. Amer.*, vol. 134, no. 2, pp. 933–938, 2013.
- [18] J. Zhong, R. Kirby, and X. Qiu, "A non-paraxial model for the audio sound behind a non-baffled parametric array loudspeaker (L)," *J. Acoustical Soc. Amer.*, vol. 147, no. 3, pp. 1577–1580, 2020.
- [19] J. Zhong, R. Kirby, and X. Qiu, "A spherical expansion for audio sounds generated by a circular parametric array loudspeaker," *J. Acoustical Soc. Amer.*, vol. 147, no. 5, pp. 3502–3510, 2020.
- [20] J. Zhong, R. Kirby, M. Karimi, and H. Zou, "A cylindrical expansion of the audio sound for a steerable parametric array loudspeaker," *J. Acoustical Soc. Amer.*, vol. 150, no. 5, pp. 3797–3806, 2021.
- [21] C. Shi and Y. Kajikawa, "A convolution model for computing the far-field directivity of a parametric loudspeaker array," *J. Acoustical Soc. Amer.*, vol. 137, no. 2, pp. 777–784, 2015.
- [22] M. Červenka and M. Bednářík, "An algebraic correction for the Westervelt equation to account for the local nonlinear effects in parametric acoustic array," *J. Acoustical Soc. Amer.*, vol. 151, no. 6, pp. 4046–4052, 2022.
- [23] S. I. Aanonsen, T. Barkve, J. N. Tjøtta, and S. Tjøtta, "Distortion and harmonic generation in the nearfield of a finite amplitude sound beam," *J. Acoustical Soc. Amer.*, vol. 75, no. 3, pp. 749–768, 1984.
- [24] Y. Kagawa, T. Tsuchiya, T. Yamabuchi, H. Kawabe, and T. Fujii, "Finite element simulation of non-linear sound wave propagation," *J. Sound Vib.*, vol. 154, no. 1, pp. 125–145, Apr. 1992.
- [25] M. Červenka and M. Bednářík, "A versatile computational approach for the numerical modelling of parametric acoustic array," *J. Acoustical Soc. Amer.*, vol. 146, no. 4, pp. 2163–2169, 2019.
- [26] J. Zhong, R. Kirby, M. Karimi, H. Zou, and X. Qiu, "Scattering by a rigid sphere of audio sound generated by a parametric array loudspeaker," *J. Acoustical Soc. Amer.*, vol. 151, no. 3, pp. 1615–1626, Mar. 2022.
- [27] Y. Jing, D. Shen, and G. T. Clement, "Verification of the Westervelt equation for focused transducers," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 58, no. 5, pp. 1097–1101, May 2011.
- [28] T. Kamakura, T. Ishiwata, and K. Matsuda, "Model equation for strongly focused finite-amplitude sound beams," *J. Acoustical Soc. Amer.*, vol. 107, no. 6, pp. 3035–3046, Jun. 2000.
- [29] M. Červenka and M. Bednářík, "Parametric acoustic array lensed by a gradient-index phononic crystal," *J. Acoustical Soc. Amer.*, vol. 149, no. 6, pp. 4534–4542, 2021.

- [30] N. Hahn, J. Ahrens, and C. Andersson, "Parametric array using amplitude modulated pulse trains: Experimental evaluation of beamforming and single sideband modulation," in *Proc. Audio Eng. Soc. Conv.*, 2021, vol. 151, pp. 1–7.
- [31] A. Marzo, T. Corkett, and B. W. Drinkwater, "Ultraino: An open phased-array system for narrowband airborne ultrasound transmission," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 65, no. 1, pp. 102–111, Jan. 2018.
- [32] Y. N. Makov, V. J. Sánchez-Morcillo, F. Camarena, and V. Espinosa, "Nonlinear change of on-axis pressure and intensity maxima positions and its relation with the linear focal shift effect," *Ultrason.*, vol. 48, no. 8, pp. 678–686, Dec. 2008.
- [33] J. Zhong, S. Wang, R. Kirby, and X. Qiu, "Reflection of audio sounds generated by a parametric array loudspeaker," *J. Acoustical Soc. Amer.*, vol. 148, no. 4, pp. 2327–2336, 2020.
- [34] B. G. Lucas and T. G. Muir, "The field of a focusing source," *J. Acoustical Soc. Amer.*, vol. 72, no. 4, pp. 1289–1296, 1982.
- [35] F. J. Pompei, *Sound from Ultrasound: The parametric Array as an Audible Sound Source*. Cambridge, MA, USA: Massachusetts Inst. Technol., 2002.



Jiaxin Zhong (Member, IEEE) received the bachelor's and master's degrees in acoustics from Nanjing University, Nanjing, China, in 2015 and 2018, respectively, and the Ph.D. degree from the University of Technology Sydney, Ultimo, NSW, Australia, in 2022. His research interests include parametric array loudspeakers, active noise control, and audio signal processing.



Tao Zhuang received the bachelor's degrees in acoustics in 2021 from Nanjing University, Nanjing, China, where he is currently working toward the Ph.D. degree with the Key Laboratory of Modern Acoustics. His research interests include parametric array loudspeakers and audio signal processing.

Ray Kirby's photograph and biography is not available at the time of publication.

Mahmoud Karimi's photograph and biography is not available at the time of publication.

Xiaojun Qiu's photograph and biography is not available at the time of publication.



Haishan Zou received the B.S. degree in acoustics, the M.S. degree in circuits and systems, and the Ph.D. degree in acoustics from Nanjing University, Nanjing, China, in 1997, 2001, and 2007, respectively. He is currently an Associate Researcher with the Department of Acoustical Science and Engineering, Nanjing University. He has authored or coauthored more than 80 journal and conference papers, filed more than 20 patents, and completed more than 20 government and industrial projects in his research fields, which include noise control and audio signal processing.



Jing Lu (Member, IEEE) received the B.S. and Ph.D. degrees in acoustics from Nanjing University, Nanjing, China, in 1999 and 2004, respectively. He is currently a Professor and the Deputy Head with the Department of Acoustical Science and Engineering, Nanjing University. He is also the Director of Nanjing University – Horizon Intelligent Audio Lab, and the Vice Director of Audio Engineering Society of China. He has authored or coauthored more than 200 journal and proceeding papers, and completed more than 20 government and industry projects. He has been granted more than 40 patents, and his research results have been widely used in industries. His research interests include audio signal processing, machine learning, and real-time implementation of audio processing systems.