Evaluation of a multi-way parametric array loudspeaker based on multiplexed double sideband modulation

Yuting GENG*, Masato NAKAYAMA†, and Takanobu NISHIURA*
* Ritsumeikan University, Shiga, Japan
E-mail: {gr0369sv@ed, nishiura@is}.ritsumei.ac.jp
† Osaka Sangyo University, Osaka, Japan
E-mail: nakayama@ise.osaka-sandai.ac.jp

***Abstract*—Parametric array loudspeakers (PALs) can achieve a sharper directivity than conventional electro-dynamic loudspeakers by utilizing ultrasonic waves. However, a PAL has difficulty reproducing low-frequency sounds because it utilizes the demodulation of ultrasound in the air. We proposed a multiplexed double sideband (M-DSB) modulation for a PAL, which can enhance the sound pressure level by utilizing harmonic distortion in demodulation. We also developed a multi-way PAL combining software and hardware approaches. Our multi-way PAL consists of a tweeter PAL (T-PAL) and woofer PALs (W-PALs), with different modulation methods and different numbers of ultrasonic transducers. In this paper, to evaluate the performance of the developed multi-way PAL, we carry out both objective and subjective experiments. In the objective experiments, the sound pressure level and frequency response are compared between target and non-target areas. In the subjective experiments, we evaluate the sound quality with sound sources of music and speech. From the experimental results, we confirmed the performance of our developed PAL in terms of frequency response, directivity, and sound quality.**

I. Introduction

Parametric array loudspeakers (PAL) [1]–[3] have been drawing attention due to their sharp directivity. They can generate a directional audible sound-beam, and the sound can be carried only to target positions. Whilst conventional electro-dynamic loudspeakers emit the target sound directly, a PAL emits an intense ultrasound synthesized by amplitude-modulating an ultrasonic carrier wave with the audible target sound. Due to the non-linear interactions in the air, the emitted sound self-demodulates into audible sound. The self-demodulation of a PAL is non-linear and frequency-dependent, so the bass reproduction of a PAL is difficult, and the sound quality is degraded by harmonic distortion [1].

Conventionally, the sound pressure and quality are considered to be affected by the modulation method. Double sideband (DSB) [4] modulation achieves higher sound pressure but causes more harmonic distortion. On the other hand, single sideband (SSB) [5] modulation achieves lower sound pressure but causes less harmonic distortion. Square root amplitude modulation (SRAM) [6] has been proposed to reduce harmonic and intermodulation distortions. Weighted double sideband (W-DSB) modulation [7] has been proposed utilizing different modulation methods in different bands with weighting factors. W-DSB can achieve a flatter frequency response, however, it still has the drawback of insufficient sound pressure in low-frequency sounds. To attain a better performance of low-frequency reproduction, we proposed multiplexed double sideband (M-DSB) modulation [8], which utilizes harmonic distortions rather than eliminating harmonic distortions to enhance the sound pressure. M-DSB modulation achieves a higher sound pressure at low frequencies than W-DSB modulation, but it is insufficient compared with the sound reproduced by electro-dynamic loudspeakers.

These previous works show that modulation methods alone are insufficient to improve the low-frequency reproduction. Psychoacoustical approaches [9], [10] have also been proposed to enhance the perceived low-frequency. However, the low-frequency components are not reproduced directly, and distortions also occurs, which degrades the sound quality. An electro-dynamic subwoofer can be combined with a PAL for a much better low-frequency reproduction [11], but leaves the problem of reproducing low-frequency sounds from PAL unsolved.

We turned our focus to multi-way design of loudspeakers. Many commercial loudspeaker systems consist of several loudspeaker units: tweeters for high-frequencies, woofers for low frequencies, mid-range drivers for mid frequencies, etc. Each reproduces a part of the frequency range, and a flat frequency response is achieved with the whole system. We introduced a multi-way design to PALs [12] and we developed a multi-way PAL comprising a tweeter PAL (T-PAL) and woofer PALs (W-PALs), with different modulation methods and numbers of ultrasonic transducers (UTs). In this paper, we focus on the evaluation of this multi-way PAL. We carry out objective and subjective experiments in terms of frequency response, directivity, and sound quality.

The remainder of this paper is organized as follows. In Section II, we describe the methodology of our proposed multi-way PAL. In Section III, we give the conditions of our developed multi-way PAL. In Sections IV and V, objective and subjective experiments are conducted, respectively. Finally, Section VI concludes the paper.

II. Proposal of Multi-Way Parametric Array Loudspeaker Based on Multiplexed Double Sideband Modulation

We proposed a multi-way PAL utilizing studies on modulation methods and hardware design for better low-frequency reproduction. Fig.1 shows an overview of the proposed PAL, and Fig.2 shows the arrangement of a trial multi-way PAL that we developed on the basis of our proposal. The proposed multi-way PAL consists of a T-PAL for low-frequency reproduction and W-PAL for high-frequency reproduction.

A. Modulation methods

We utilized M-DSB modulation in the W-PALs for a better low-frequency reproduction, and SSB modulation in T-PAL for better high-frequency sound quality. As shown in Fig. 3, the target signal $s(t)$ is first split into low-frequency sound $s_{\mathrm{L}}(t)$ and high-frequency sound $s_{\mathrm{H}}(t)$, which can be indicated as:

$$s_{\mathrm{L}}(t) = s(t) * h_{\mathrm{LPF}_{\omega_{\mathrm{co}}}}(t), \tag{1}$$
$$s_{\mathrm{H}}(t) = s(t) * h_{\mathrm{HPF}_{\omega_{\mathrm{co}}}}(t), \tag{2}$$

where t denotes the time index, $*$ denotes the convolution operator, and $h_{\mathrm{LPF}_{\omega_{\mathrm{co}}}}(t)$ and $h_{\mathrm{HPF}_{\omega_{\mathrm{co}}}}(t)$ denote the low-pass filter and high-pass filter with a crossover frequency ω_{co}, respectively.

For high-frequency reproduction, an SSB modulated wave can be generated by filtering out the upper sideband of a DSB modulated wave, which can be indicated as:

$$u_{\mathrm{H}}(t) = \{1 + m_{\mathrm{SSB}} \cdot s_{\mathrm{H}}(t)\} \cdot c(t) * h_{\mathrm{LPF}_{f_{\mathrm{C}}}}(t), \tag{3}$$

where m_{SSB} denotes the modulation$(0 < m_{\mathrm{SSB}} \leq 1)$, $c(t)$ denotes the carrier wave with frequency f_{C}, and $h_{\mathrm{LPF}_{f_{\mathrm{C}}}}(t)$ denotes the low-pass filter with a cut-off frequency the same as the carrier frequency f_{C}. Ordinarily, a single frequency sinusoidal wave is utilized as the carrier wave, indicated as follows:

$$c(t) = A_{\mathrm{C}} \cos(2\pi f_{\mathrm{C}} t), \tag{4}$$

where A_{C} and f_{C} denote the amplitude and frequency, respectively.

For low-frequency reproduction, an M-DSB modulated wave can be generated by combining the DSB modulation and double sideband with suppressed carrier (DSB-SC) modulation [8]. As shown in Fig. 3, the low-frequency sound $s_{\mathrm{L}}(t)$ is first compressed into $s_{\mathrm{L}_{\mathrm{F}/2}}(t)$, which is one octave lower. In the case that $s_{\mathrm{L}}(t)$ is a pure tone, $s_{\mathrm{L}}(t)$ and $s_{\mathrm{L}_{\mathrm{F}/2}}(t)$ can be indicated as:

$$s_{\mathrm{L}}(t) = A_{\mathrm{L}} \cos(2\pi f_{\mathrm{L}} t), \tag{5}$$
$$s_{\mathrm{L}_{\mathrm{F}/2}}(t) = A_{\mathrm{L}} \cos(\pi f_{\mathrm{L}} t), \tag{6}$$

where A_{L} denotes the amplitude of $s_{\mathrm{L}}(t)$. Then, DSB modulation and DSB-SC modulation are carried out and the two modulated waves are added to generate the M-DSB modulated wave. The modulated waves emitted from the W-PAL can be indicated as

$$\begin{aligned} u_{\mathrm{L}}(t) =& u_{\mathrm{DSB}}(t) + u_{\mathrm{SC}}(t) \\ =& \{1 + m_{\mathrm{DSB}} \cdot s_{\mathrm{L}}(t) + m_{\mathrm{SC}} \cdot s_{\mathrm{L}_{\mathrm{F}/2}}(t)\} \cdot c(t) \\ =& \frac{m_{\mathrm{DSB}} A_{\mathrm{L}} A_{\mathrm{C}}}{2} \cos(2\pi(f_{\mathrm{C}} \pm f_{\mathrm{L}})t) \\ &+ \frac{m_{\mathrm{SC}} A_{\mathrm{L}} A_{\mathrm{C}}}{2} \cos(2\pi(f_{\mathrm{C}} \pm f_{\mathrm{L}}/2)t) \\ &+ A_{\mathrm{C}} \cos(2\pi f_{\mathrm{C}} t), \end{aligned} \tag{7}$$

where m_{DSB} and m_{SC} denote the modulation factor of DSB and DSB-SC modulation, respectively. It is known that the M-DSB modulated wave $u_{\mathrm{L}}(t)$ consists of the carrier wave (frequency: f_{C}), sidebands of the DSB modulated wave (frequency: $f_{\mathrm{C}} \pm f_{\mathrm{L}}$), and sidebands of the DSB-SC modulated wave (frequency: $f_{\mathrm{C}} \pm f_{\mathrm{L}}/2$). f_{L} can be demodulated from the difference between the DSB sideband and carrier wave, and also between the DSB-SC sidebands. On the other hand, though a half frequency distortion $f_{\mathrm{L}}/2$ is also generated, the

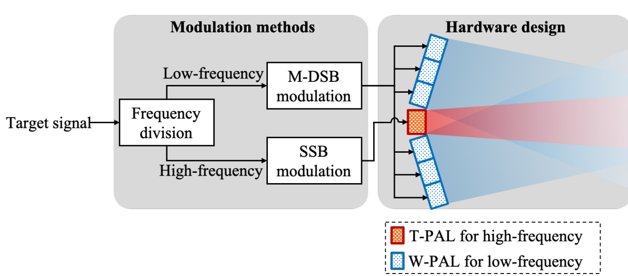

Fig. 1. Overview of the proposed multi-way PAL.

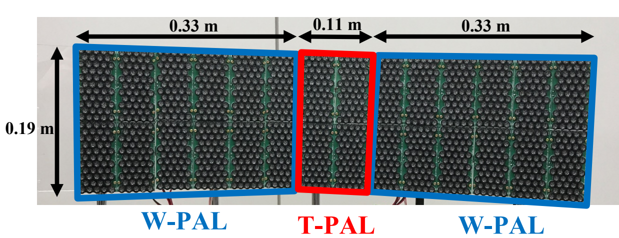

Fig. 2. Arrangement of the trial multi-way PAL.

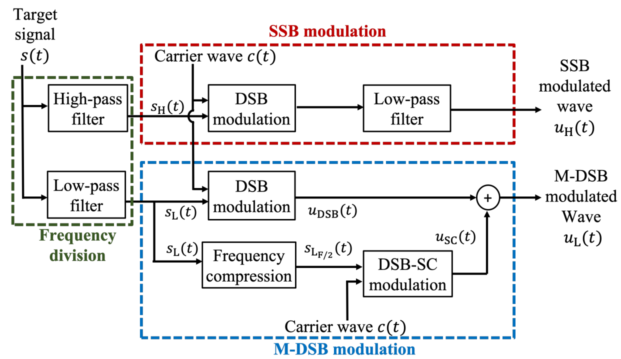

Fig. 3. Modulation methods and signal processing of the proposed multi-way parametric array loudspeaker.

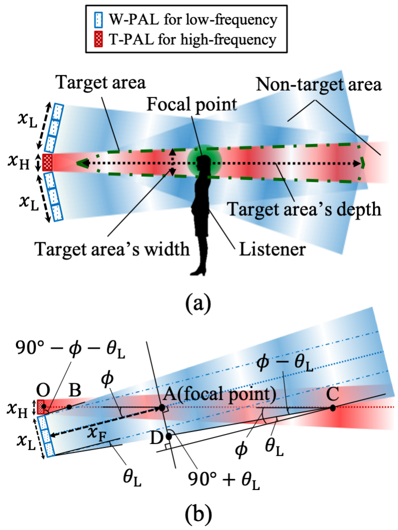

Fig. 4. Hardware arrangement of the proposed multi-way PAL: (a)overview; (b)geometric notations.

second harmonic distortion of $f_{\mathrm{L}}/2$ is utilized here to enhance the $f_{\mathrm{L}}/2$.

B. Hardware design

We use more UTs for the W-PAL than the T-PAL because the sound pressure can be improved by increasing the number of UTs. Since PAL has a very sharp directivity, the arrangement of the T-PAL and W-PAL should be studied to guarantee that the audible areas of W-PALs and T-PAL overlap. Fig.4 shows the arrangement of a multi-way PAL. W-PALs are divided into two parts and are closely spaced to create a focal point.

In Fig. 4(b), x_{F} denotes the distance between the multi-way PAL and a focal point, x_{L} and x_{H} denote the array length of W-PALs and T-PAL, respectively, and θ_{L} and θ_{H} denote the directional angle of W-PALs and T-PAL, respectively. The angle between W-PALs and T-PAL is denoted by ϕ and can be indicated as follows:

$$\phi = \arctan\left(\frac{x_{\mathrm{L}}}{2x_{\mathrm{F}}}\right) + \arctan\left(\frac{x_{\mathrm{H}}}{2x_{\mathrm{F}}}\right). \tag{8}$$

The distance between the T-PAL and target area can be indicated as follows:

$$l_{\mathrm{OB}} = \frac{x_{\mathrm{H}}}{2}\tan(90^\circ - \phi - \theta_{\mathrm{L}}). \tag{9}$$

The width of the target area at the focal point can be indicated as follows:

$$w_{\mathrm{ta}} = 2\left(\frac{x_{\mathrm{L}}}{2} + x_{\mathrm{F}}\tan\theta_{\mathrm{L}}\right) = x_{\mathrm{L}} + 2x_{\mathrm{F}}\tan\theta_{\mathrm{L}}. \tag{10}$$

By utilizing the law of sines in the triangle ACD, the distance between the focal point and the end of the target area can be calculated as:

$$l_{\mathrm{AC}} = \left(\frac{x_{\mathrm{L}}}{2} + x_{\mathrm{F}}\tan\theta_{\mathrm{L}}\right)\frac{\sin(90^\circ + \theta_{\mathrm{L}})}{\sin(\phi - \theta_{\mathrm{L}})}. \tag{11}$$

TABLE I
EXPERIMENTAL CONDITIONS.

Environment	Office room (T_{60} = 650 ms)
Ambient noise level	L_A = 32.4 dB
Temperature / Humidity	24.0°C / 33.7 %
Recording distance	2.0 m
Sampling frequency / Quantization	192 kHz / 16 bits
Sound source	White noise (0.1∼8 kHz)

TABLE II
EXPERIMENTAL EQUIPMENT.

Ultrasonic transducer	SPL Limited, UT1007-Z325R
Power amplifier	JVC, PS-A2002
Microphone	SENNHEISER, MKH 416-P48
A/D, D/A converter	RME, FIREFACE UFX

The depth of the target area d_{ta} can be calculated as:

$$\begin{aligned} d_{\mathrm{ta}} =& l_{\mathrm{OA}} + l_{\mathrm{AC}} - l_{\mathrm{OB}} \\ =& x_{\mathrm{F}} + \left(\frac{x_{\mathrm{L}}}{2} + x_{\mathrm{F}}\tan\theta_{\mathrm{L}}\right)\frac{\sin(90^\circ + \theta_{\mathrm{L}})}{\sin(\phi - \theta_{\mathrm{L}})} \\ &- \frac{x_{\mathrm{H}}}{2}\tan(90^\circ - \phi - \theta_{\mathrm{L}}). \end{aligned} \tag{12}$$

In the case of $\phi \le \theta_{\mathrm{L}}$, the point C no longer exists, and the demodulated sounds of the W-PALs overlaps with the demodulated sound of the T-PAL endlessly.

III. Conditions of the Developed Multi-way Parametric Array Loudspeaker

We developed a trial of multi-way under the design described in Section II, as shown in Fig. 2, with 1400 UTs (SPL Limited, UT1007-Z325R). The crossover frequency ω_{co} is set to 1 kHz. 200 UTs are used in the T-PAL for high-frequency reproduction, and 1200 UTs are used in the W-PAL for low-frequency. The array lengths of the W-PAL and T-PAL are $(x_{\mathrm{L}}, x_{\mathrm{H}})$ = (0.33 m, 0.11 m), and we set the position of the focal point as x_{F} = 2.0 m. Under this arrangement, the average sound pressure level of the W-PAL (0.1 ∼ 1 kHz) is approximately equal to that of the T-PAL (1 ∼ 8 kHz) at the focal point.

We carried out an experiment to measure the demodulated sound pressure level (SPL) distribution of the developed multi-way PAL. The experimental conditions and equipment are listed in Tabs. I and II, respectively. As designed in Section II, we utilize M-DSB modulation (f_{C} = 40 kHz, m_{DSB} = 0.5, m_{SC} = 0.5) in the W-PAL and SSB modulation (f_{C} = 40 kHz, m_{SSB} = 1.0) in the T-PAL. The measured SPL distribution

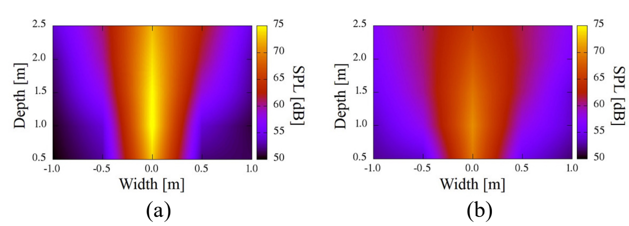

Fig. 5. Demodulated sound pressure distribution of developed multi-way PAL: (a)tweeter PAL for high-frequency reproduction; (b)woofer PAL for low-frequency reproduction.

of the T-PAL and W-PAL are shown in Fig. 5 (a) and (b), respectively. We then calculated the directivity angle of each PAL, which is defined as the angle to the point where the SPL of the demodulated sound is attenuated at 6.0 dB in the horizontal direction of the focal point. The directional angle of the W-PAL is $\theta_{\mathrm{L}} = 15.8^{\circ}$ and that of the T-PAL is $\theta_{\mathrm{H}} = 7.9^{\circ}$. Substituting these measured parameters into Eq. (8), $\phi = 6.3^{\circ}$. Therefore, on the basis of our theoretical analysis in Section II-B, the condition $\phi \leq \theta_{\mathrm{L}}$ is satisfied, and it is known that the demodulated sounds of W-PALs and T-PAL will overlap all the way in the propagation direction. Moreover, the width of target area w_{ta} is measured to be about 0.7 m, which is considered large enough for the size of a human head.

IV. Objective evaluation experiments

A. Conditions of objective experiments

In our developed multi-way PAL, we utilized M-DSB modulation and SSB modulation in the W-PAL and T-PAL with different numbers of UTs to achieve a flatter frequency response. Therefore, we carried out objective experiments on the frequency response to confirm the effectiveness. As shown in Fig. 6, we observed the demodulated sounds at three positions: focal point (0 m, 2.0 m), (0 m, 1.0 m) for the target area and (1.0 m, 2.0 m) for the non-target area. We also compare the performance of the developed multi-way PAL when utilizing different modulation methods as shown in Tab. III. Here, in the cases of "SSB" "DSB," and "M-DSB," we set no crossover so the same signal is reproduced from the W-PAL and T-PAL. In the case of "Proposed," we set the crossover frequency $\omega_{\mathrm{co}} = 1$ kHz. A signal with a frequency higher than ω_{co} is emitted from the T-PAL, modulated by SSB modulation, and a signal with a frequency lower than ω_{co} is emitted from the W-PAL, modulated by M-DSB modulation. A time-stretched pulse (TSP) signal ($0 \sim 8$ kHz) is utilized in the measurement of the frequency response [13]. A modulated TSP signal is emitted from the PAL, and the demodulated sounds are recorded at observation positions. Other experimental conditions and equipment are the same as those shown in Tabs. I and II, respectively.

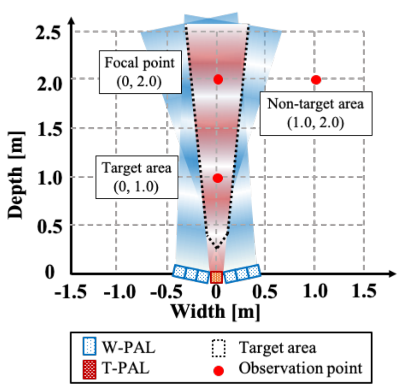

Fig. 6. Experimental arrangement (vertical view).

TABLE III
Comparison in evaluation experiments.

Index	Modulation in W-PAL (Low-frequency)	Modulation in T-PAL (High-frequency)
SSB	SSB	SSB
DSB	DSB	DSB
M-DSB	M-DSB	M-DSB
Proposed	M-DSB	SSB

To evaluate the flatness of the frequency response, we calculated the average error of observed power spectra P_{err}, which is defined as follows:

$$P_{\mathrm{err}} = \frac{1}{f_{\max} - f_{\min}} \sum_{f=f_{\min}}^{f_{\max}} |P(f) - P_{\mu}|, \tag{13}$$

where P_{μ} denotes the average of the observed power spectrum. A smaller P_{err} suggests a flatter frequency response.

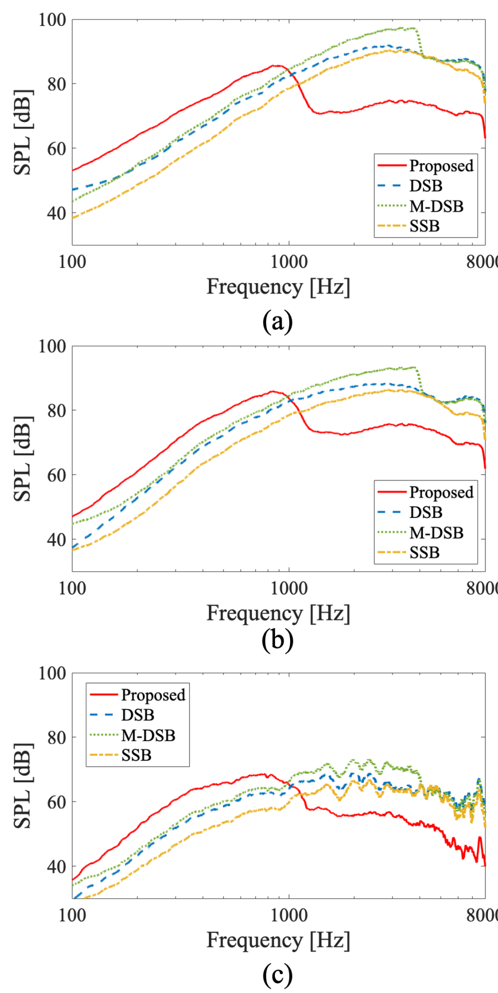

Fig. 7. Power spectra of demodulated sounds observed at different positions: (a) focal point (0 m, 2.0 m); (b) target area (0 m, 1.0 m); (c) non-target area (1.0 m, 2.0 m).

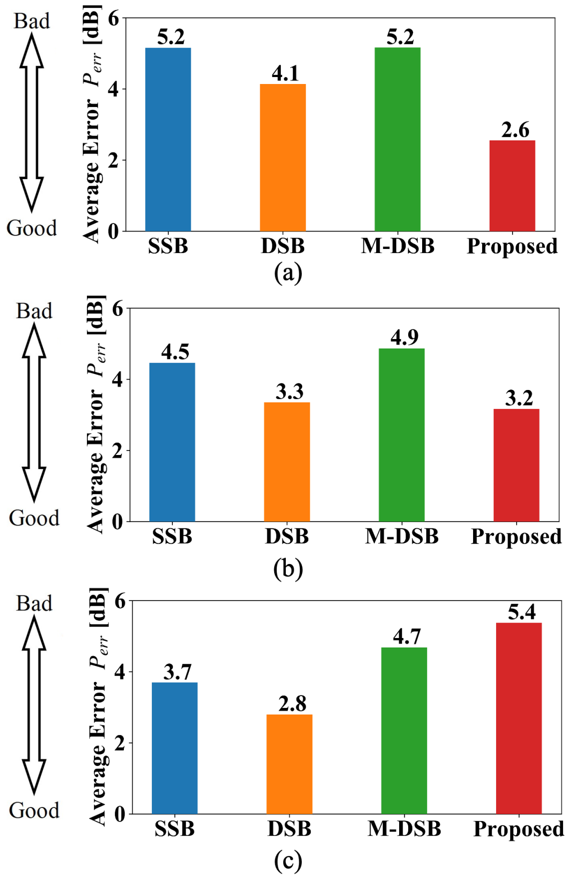

Fig. 8. Average error of power spectra observed at different positions: (a) focal point (0 m, 2.0 m); (b) target area (0 m, 1.0 m); (c) non-target area (1.0 m, 2.0 m).

B. Results of objective experiments

Figure 7 shows the power spectra and the average error of demodulated sounds at three observation positions. Figure 8 shows the average error of demodulated sounds at the same positions. Table IV shows the average demodulated SPL at the same positions.

We first focus on the focal point. From Fig. 7 (a), it is confirmed that M-DSB modulation can achieve a higher SPL than DSB modulation and SSB modulation at the focal point. It is also confirmed from Fig. 7 (a) that our multi-way design achieves an improvement of SPL at low frequencies compared with the cases utilizing only one modulation at the focal point. From Fig. 8 (a), it is confirmed that our multi-way design achieves a flatter frequency response of demodulated sound compared with the cases utilizing only one modulation. From Tab. IV, we can confirm that our proposal results in a smaller difference of SPL between low and high frequencies than others.

Next, we turn to the target area. The same conclusions can be drawn from Figs. 7 (b) and 8 (b) in that our proposal achieves better low-frequency reproduction and flatter frequency response at the target area. From Tab. IV, it is

TABLE IV
AVERAGE DEMODULATED SPL [DB].

		SSB	DSB	M-DSB	Proposed
Focal point (0 m, 2.0 m)	Low-freq.	64.8	69.9	71.0	77.0
	High-freq.	85.9	88.0	89.8	72.3
Target area (0 m, 1.0 m)	Low-freq.	65.3	70.2	71.9	76.9
	High-freq.	81.9	84.9	86.5	72.5
Non-target area (1.0 m, 2.0 m)	Low-freq.	51.4	56.2	57.8	62.9
	High-freq.	61.3	63.5	65.1	51.1

(Low-freq.: 0.1 $\sim$ 1 kHz; High-freq.: 1 $\sim$ 8 kHz;)

TABLE V
SUPPLEMENTARY EQUIPMENT IN SUBJECTIVE EXPERIMENTS.

Electro-dynamic loudspeaker	FOSTEX, FE83En
Headphone	SONY, MDR-CD900ST
Headphone amplifier	Audio-technica, AT-HA2

confirmed that the target area shows similar values of average demodulated SPL compared with the focal point. Since we set the T-PAL and W-PAL to focus on one point, it is normal that the SPL at the focal point is higher than that of other positions. On the other hand, the observation point is nearer to the T-PAL and W-PAL compared with the focal point. The distance attenuation is less in some cases, so it is normal that the SPL at this position is higher than the focal point.

Finally, we focus on the non-target area. Since we utilized the directivity of the PAL, we expect the SPL to be higher in the target area and lower in the non-target area. High frequencies are only emitted from the T-PAL in the case of "Proposed," however, they are emitted from both the W-PAL and T-PAL in other cases. So the SPL of "Proposed" at high frequencies is much lower than that of other cases. On the other hand, since "Proposed" achieves a much higher SPL at low frequencies at the focal point and target area, it is normal to result in a higher SPL at low frequencies at the non-target area.

With the experimental results and discussions, we confirmed that our proposed multi-way PAL can achieve a better low-frequency reproduction and flatter frequency response in the target area.

V. SUBJECTIVE EVALUATION EXPERIMENTS

A. Conditions of subjective experiments

In this section, we carried out subjective experiments to confirm if the sound quality is improved. The sound quality includes many metrics, and in these experiments, we focused on low-frequency reproduction. We carried out this subjective experiment with 8 participants (3 females and 5 males). Sound sources of female speech [14], dance music, and classical music are used as stimuli. Fig. 9 shows the spectrograms and power spectra of each sound source. From Fig. 9, we confirmed that dance music includes more energy in the low-frequency range below 1 kHz.

The sounds reproduced from the PAL with different patterns, as shown in Tab. III, were compared with those reproduced from the electro-dynamic loudspeaker. All of the stimuli were recorded at the focal point in advance under the same

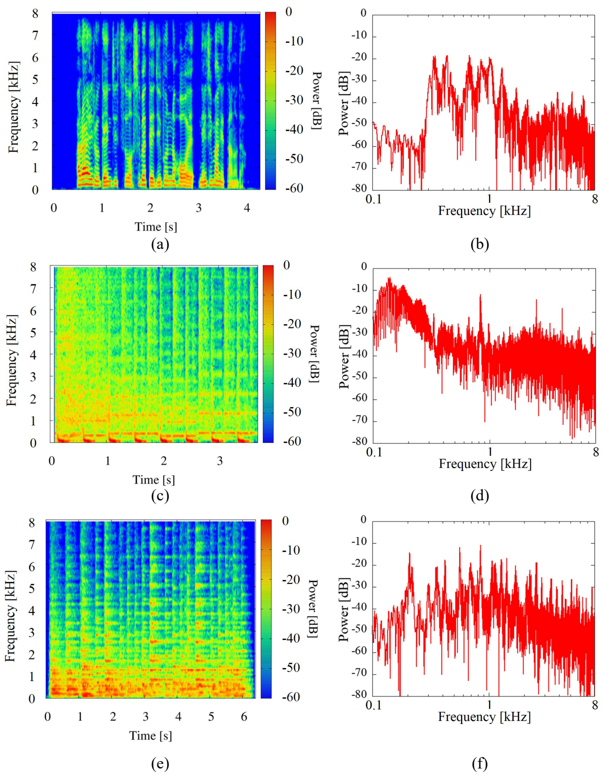

Fig. 9. Spectrograms and power spectra of sound sources: (a) spectrogram and (b) power spectrum of female speech; (c) spectrogram and (d) power spectrum of dance music; (e) spectrogram and (f) power spectrum of classical music.

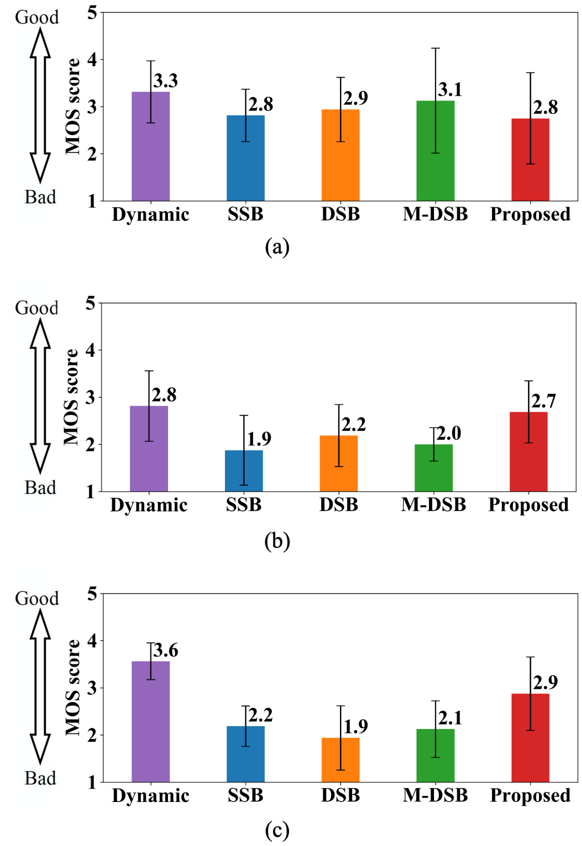

Fig. 10. Results of mean opinion score on sound quality: (a) female speech; (b) dance music; (c) classical music.

conditions shown in Tab. I, and evaluated through headphones. The equipment used is listed in Tab. II with supplementary equipment shown in Tab. V. The mean opinion score (MOS) [15] is used as the evaluation metric, graded as 1 (Bad), 2 (Poor), 3 (Fair), 4 (Good), and 5 (Excellent). The reference is the original sound reproduced directly from headphones, with a standard score of 5 (Excellent).

B. *Results of subjective experiments*

Fig. 10 shows the results of subjective experiments. It is confirmed that sounds reproduced from the PAL were evaluated lower than sounds reproduced from the electro-dynamic loudspeaker. This is because of the poor low-frequency reproduction of the PAL, as discussed in Section I. It is also confirmed that in Fig. 10 (b) and (c), our proposal evaluated higher than other sounds reproduced from the PAL. This is because our proposal can achieve a better low-frequency reproduction, as proved from the objective experiments in Section IV. In Fig. 10 (a), our proposal was evaluated lower than others from the PAL when reproducing female speech. However, the difference is relatively small. From the experimental results above, we confirmed the effectiveness of our proposal.

VI. Conclusion

In this paper, we proposed a multi-way PAL based on M-DSB modulation and carried out evaluation experiments. In our proposal, M-DSB and SSB modulation are utilized for low and high frequencies, respectively. The hardware design was also studied to achieve a better low-frequency reproduction and flatter frequency response. From the experimental results of the objective and subjective experiments, we confirmed the effectiveness of our proposal in that it can achieve a better low-frequency reproduction and flatter frequency response.

However, since we employed many more UTs in low-frequency reproduction than in high-frequency reproduction, the large scale of devices becomes a problem. Also, the frequency response shows a transient around 1 kHz. In future work, we will attempt to increase the number of frequency divisions to achieve an even flatter frequency response. We will also take other signal processing techniques in addition to modulation methods to obtain a better sound quality.

Acknowledgment

This work is supported by Ritsumeikan R-GIRO, JST COI, JSPS KAKENHI Grant Numbers JP18K11365, JP19H04142.

References

[1] P. J. Westervelt, "Parametric acoustic array," *The Journal of the Acoustical Society of America*, vol. 35, no. 4, pp. 535–537, 1963.

[2] K. Aoki, T. Kamakura, and Y. Kumamoto, "Parametric loudspeaker-characteristics of acoustic field and suitable modulation of carrier ultrasound," *Electronics and Communications in Japan*, vol. 74, no. 9, pp. 76–82, 1991.

[3] W. S. Gan, J. Yang, and T. Kamakura, "A review of parametric acoustic array in air," *Applied Acoustics*, vol. 73, no. 12, pp. 1211–1219, 2012.

[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," *The Journal of the Acoustical Society of America*, vol. 73, no. 5, pp. 1532–1536, 1983.

[5] K. Aoki, T. Kamakura, and Y. Kumamoto, "Parametric loudspeaker–characteristics of acoustic field and suitable modulation of carrier ultrasound," *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, vol. 74, no. 9, pp. 76–82, 1991.

[6] T. Kamakura, "Developments of parametric loudspeaker for practical use," in *Proc. 10th International Symposium on Non-linear Acoustics*, pp. 147–150, 1984.

[7] D. Ikefuji, M. Nakayama, T. Nishiura, and Y. Yamashita, "Weighted double sideband modulation toward high quality audible sound on parametric loudspeaker," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 843–847, 2013.

[8] Y. Nakano, T. Fukumori, M. Nakayama, and T. Nishiura, "A study on audible low-frequency sound emphasis based on multiplexed double sideband modulation in parametric loudspeaker," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 258, no. 5, pp. 2054–2065, 2018.

[9] C. Shi, H. Mu, and W. S. Gan, "A psychoacoustical preprocessing technique for virtual bass enhancement of the parametric loudspeaker," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. pp. 31–35, 2013.

[10] Y. Geng, M. Nakayama, and T. Nishiura, "Virtual bass enhancement based on harmonics control using missing fundamental in parametric array loudspeaker," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 259, no. 7, pp. 2660–2671, 2019.

[11] Y. Sugibayashi, S. Kurimoto, D. Ikefuji, M. Morise, and T. Nishiura, "Three-dimensional acoustic sound field reproduction based on hybrid combination of multiple parametric loudspeakers and electrodynamic subwoofer," *Applied Acoustics*, vol. 73, no. 12, pp. 1282–1288, 2012.

[12] Y. Geng, Y. Nakano, M. Nakayama, and T. Nishiura, "Development of multi-way parametric array loudspeaker using multiplexed double sideband modulation," in *23rd International Congress on Acoustics*, pp. 1069–1076, 2019.

[13] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *The Journal of the Acoustical Society of America*, vol. 69, no. 5, pp. 1484–1488, 1981.

[14] A. Kurematsu, K. Takeda, Y. Sagisaka, S. Katagiri, H. Kuwabara, and K. Shikano, "ATR Japanese speech database as a tool of speech recognition and synthesis," *Speech communication*, vol. 9, no. 4, pp. 357–363, 1990.

[15] Rec. ITU-R P. 800:, "Methods for subjective determination of transmission quality," 1996.