Two-dimensional multizone sound field reproduction using a wave-domain method

Zerui Han, Ming Wu, Qiaoxi Zhu, and Jun Yang

Citation: The Journal of the Acoustical Society of America **144**, EL185 (2018); doi: 10.1121/1.5054079
View online: https://doi.org/10.1121/1.5054079
View Table of Contents: https://asa.scitation.org/toc/jas/144/3
Published by the Acoustical Society of America

Two-dimensional multizone sound field reproduction using a wave-domain method

Zerui Han,[1,a)] Ming Wu,[1] Qiaoxi Zhu,[2] and Jun Yang[1,b)]
[1]*Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China*
[2]*Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, New South Wales 2007, Australia*
hanzr.nju@gmail.com; mingwu@mail.ioa.ac.cn; qiaoxi.zhu@gmail.com; jyang@mail.ioa.ac.cn

Abstract: This paper addresses a two-dimensional multizone sound field reproduction approach using a wave-domain method. The desired sound fields in the bright and dark zones are described as orthogonal expansions of basis functions over the regions. The loudspeaker weights are obtained by maximizing the contrast among multiple zones in the wave domain. Simulation results demonstrate that compared with the conventional acoustic contrast control approach, the proposed method improves the level of acoustic contrast and array gain over the entire control region and is less sensitive to the selection of the regularization parameter.

[PG]
Date Received: May 15, 2018 **Date Accepted:** August 27, 2018

1. Introduction

Multizone sound field reproduction aims to generate the desired sound field in a specific area without disturbing people in other areas. A common method for generating multizone sound field is the acoustic contrast control (ACC)[1] approach, which maximizes the acoustic potential energy density ratio between control points in the bright zone(s) and control points in the dark zone(s). To reduce performance degradation when system errors occur, researchers have studied the robustness of ACC techniques[2–4] and applied array gain (AG) constraint or coarse acoustic modeling to increase system robustness.

The performance of the ACC method has been investigated in various applications.[5,6] However, a practical drawback of the conventional multiple points-based ACC method is that additional control microphones are required inside the regions to ensure the sound control throughout large spatial regions. Recently, the wave-domain method has been proposed in sound field control, such as active noise control,[7] echo cancellation,[8] room compensation,[9] and multizone sound field reproduction.[10,11] One advantage of the wave-domain method is that it optimizes the performance not only at the control points but also over the region(s).

In this paper, we propose a wave-domain acoustic contrast control (WDACC) method using two circular microphone arrays placed in the desired bright and dark zones to measure the acoustic transfer functions (ATFs) and a loudspeaker array to produce a two-dimensional (2D) multizone sound field. The proposed method demonstrates improved performance in terms of acoustic contrast and AG effort over the regions. The remainder of this paper is organized as follows. Section 2 presents the mathematical formulation of ACC and WDACC. In Sec. 3, simulations are conducted to evaluate the proposed WDACC in comparison with the conventional multiple points-based ACC. Conclusions are provided in Sec. 4.

2. Multizone sound field reproduction

The reproduction of a multizone sound field in the 2D horizontal plane under free field condition is considered for simplicity and clarity. Assume that the desired bright zone is a circular area with radius r_b sampled by M_b discrete control points and that the dark zone is another circular area with radius r_d sampled by M_d control points. A circular array of L loudspeakers with a radius R_L and regular spacing $\Delta = 2\pi/L$ is

[a)]Also at: School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China.
[b)]Author to whom correspondence should be addressed.

employed. All desired regions are assumed to be inside the geometry of the loudspeaker array distribution.

2.1 ACC

Assume that each sound source has a volume velocity $q_l(\omega)$. The sound pressure at the control points in the bright zone can be written in terms of the ATF,

$$\mathbf{p}_b(\omega) = \mathbf{G}_b(\omega)\mathbf{q}(\omega), \tag{1}$$

where $\mathbf{p}_b(\omega) = [p(\mathbf{r}_{b1}, \omega), p(\mathbf{r}_{b2}, \omega), \ldots, p(\mathbf{r}_{bM_b}, \omega)]^T$ is the sound pressure vector at M_b control points, $\mathbf{G}_b(\omega)$ is the ATF matrix between the loudspeakers and the control points in the bright zone, $\mathbf{q}(\omega) = [q_1(\omega), q_2(\omega), \ldots, q_L(\omega)]^T$ is the weight vector of the loudspeaker array, and $\omega = 2\pi f$ is the angular frequency.

The acoustic potential energy density in the bright zone is defined as[1]

$$E_b = \frac{1}{M_b}\sum_{i=1}^{M_b} p^*(\mathbf{r}_{bi})p(\mathbf{r}_{bi}) = \frac{1}{M_b}\mathbf{q}^H\mathbf{G}_b^H\mathbf{G}_b\mathbf{q} = \mathbf{q}^H\mathbf{R}_b\mathbf{q}, \tag{2}$$

where $(\cdot)^*$ and $(\cdot)^H$ represent the complex conjugate and the Hermitian transpose, respectively, $\mathbf{R}_b = \mathbf{G}_b^H\mathbf{G}_b/M_b$ is the correlation matrix corresponding to the bright zone, and ω is not marked for convenience in the following text. Similarly, the acoustic potential energy density in the dark zone is

$$E_d = \frac{1}{M_d}\sum_{i=1}^{M_d} p^*(\mathbf{r}_{di})p(\mathbf{r}_{di}) = \frac{1}{M_d}\mathbf{q}^H\mathbf{G}_d^H\mathbf{G}_d\mathbf{q} = \mathbf{q}^H\mathbf{R}_d\mathbf{q}, \tag{3}$$

where $\mathbf{G}_d$ is defined as the ATF matrix between the loudspeaker array and the control points in the dark zone; and $\mathbf{R}_d = \mathbf{G}_d^H\mathbf{G}_d/M_d$ is the correlation matrix corresponding to the dark zone.

The principle of the ACC method is to maximize the ratio of the acoustic potential energy density between the bright and dark zones,[1] which yields

$$\max_{\mathbf{q}} C = \frac{E_b}{E_d} = \frac{M_d}{M_b}\frac{\mathbf{q}^H\mathbf{G}_b^H\mathbf{G}_b\mathbf{q}}{\mathbf{q}^H\mathbf{G}_d^H\mathbf{G}_d\mathbf{q}} = \frac{M_d}{M_b}\frac{\mathbf{q}^H\mathbf{R}_b\mathbf{q}}{\mathbf{q}^H\mathbf{R}_d\mathbf{q}}, \tag{4}$$

where C denotes the contrast between the bright and dark zones; its level is defined as $10\log_{10}C$. The optimal weight vector that maximizes C is given by the following form:

$$\mathbf{q}_{\mathrm{ACC}} = \Phi\{(\mathbf{R}_d + \delta\mathbf{I})^{-1}\mathbf{R}_b\}, \tag{5}$$

where $\Phi\{\cdot\}$ denotes the eigenvector corresponding to the maximum eigenvalue and the diagonal loading $\delta > 0$ is a preselected regularization parameter to improve robustness, which is usually a frequency-dependent parameter.

The AG level is defined as[2]

$$\mathrm{AG} = 10\log_{10}\frac{\mathbf{p}_b^H\mathbf{p}_b}{\mathbf{q}^H\mathbf{q}} = 10\log_{10}\frac{\mathbf{q}^H\mathbf{R}_b\mathbf{q}}{\mathbf{q}^H\mathbf{q}}. \tag{6}$$

The AG is the ratio of the energy between the radiated sound field in the bright zone and the array input, which provides a measurement for the solution robustness. The AG also characterizes the radiation efficiency under a given input power condition.

2.2 WDACC

The solution of the wave equation can be written in terms of the Bessel function for cylindrical coordinates $\mathbf{r}_{bi} = (r_{bi}, \theta_{bi})$ with respect to an origin located within a source-free region. The sound pressure for regions inside the source distribution can be expressed as[12]

$$p(\mathbf{r}_{bi}) = \sum_{n=-\infty}^{\infty} a_{n,b}J_n(kr_{bi})e^{jn\theta_{bi}}, \tag{7}$$

where $a_{n,b}$ is the nth order sound field coefficient of the bright zone, $J_n(\cdot)$ is the Bessel function of order n, $k = \omega/c$ is the wavenumber, and c is the speed of sound in air. The decomposition coefficients $a_{n,b}$ represent the spatial sound field of the bright zone in the wave domain.

Since the ATF is equal to the sound field generated by a unit input, we can write it as

$$g_l(\mathbf{r}_{bi}) = \sum_{n=-\infty}^{\infty} b_{n,b}(l) J_n(kr_{bi}) e^{jn\theta_{bi}}, \tag{8}$$

where $g_l(\mathbf{r}_{bi})$ is the element of the ATF matrix $\mathbf{G}_b$, and $b_{n,b}(l)$ denotes the nth order sound field coefficient of the ATF between the lth loudspeaker and the bright zone.

By substituting Eqs. (7) and (8) into Eq. (1), the coefficients of the reproduced sound field in the bright zone and the transfer functions are related to the loudspeaker weights by

$$a_{n,b} = \sum_{l=1}^{L} b_{n,b}(l) q_l. \tag{9}$$

The series in Eqs. (7) and (8) have an infinite number of terms, which can be truncated to a finite order within a given region of interest. The sound field inside the circular region $r \leqslant r_b$ can be represented by a summation over n truncated to

$$N_b = \lceil ekr/2 \rceil \approx \lceil kr_b \rceil, \tag{10}$$

for a truncation error of less than approximately 10%,[13] where $\lceil \cdot \rceil$ denotes the ceiling function, and e is Euler's number. We can rewrite Eq. (9) in matrix form,

$$\mathbf{a}_b = \mathbf{B}_b \mathbf{q}, \tag{11}$$

where $\mathbf{a}_b = [a_{-N,b}, a_{-N+1,b}, \ldots, a_{N,b}]^T$ is the expansion coefficient vector of the sound field in the bright zone and $\mathbf{B}_b$ is the matrix of the sound field coefficient of the ATFs between the loudspeaker array and the bright zone,

$$\mathbf{B}_b = \begin{bmatrix} b_{-N,b}(1) & b_{-N,b}(2) & \cdots & b_{-N,b}(L) \\ \vdots & \vdots & \ddots & \vdots \\ b_{N,b}(1) & b_{N,b}(2) & \cdots & b_{N,b}(L) \end{bmatrix},$$

whose elements can be calculated by theoretical solutions or precalibration, such as the spatial Fourier transform[12] and matrix inversion method.[14]

The acoustic potential energy density in the wave domain over the bright zone is defined as

$$\tilde{E}_b = \frac{1}{S} \int\int_S |p(\mathbf{r}_b)|^2 r dr d\theta = \frac{2}{(kr_b)^2} \sum_{n=-\infty}^{\infty} |a_{n,b}|^2 \int_0^{kr_b} |J_n(r)|^2 r dr = \sum_{n=-\infty}^{\infty} w_n(kr_b) |a_{n,b}|^2, \tag{12}$$

which is derived using the orthogonality property of exponential functions, where $w_n(kr_b) = 2/(kr_b)^2 \int_0^{kr_b} |J_n(r)|^2 r dr$ is the coefficient weighting function. Different from Eq. (2) in the multiple points-based method, the acoustic potential energy density is defined over the controlled region by the integration in the wave domain. Figure 1 shows that the integrals in $w_n(kR)$ vary with k and $|n|$ for R set to 0.5 m. We note that the integrals rapidly decrease to zero once a threshold is reached, which indicates that only some of the basis functions contribute primary energy to the sound field. Thus Eq. (12) can be written in matrix form by truncating the order of the basis functions,

$$\tilde{E}_b = \mathbf{a}_b^H \mathbf{W}_b \mathbf{a}_b = \mathbf{q}^H \mathbf{B}_b^H \mathbf{W}_b \mathbf{B}_b \mathbf{q} = \mathbf{q}^H \tilde{\mathbf{R}}_b \mathbf{q}, \tag{13}$$

where $\mathbf{W}_b = \mathrm{diag}\{w_{-N}(kr_b), w_{-N+1}(kr_b), \ldots, w_N(kr_b)\}$ is a coefficient weighting matrix corresponding to the bright zone and $\tilde{\mathbf{R}}_b = \mathbf{B}_b^H \mathbf{W}_b \mathbf{B}_b$ is the wave-domain correlation matrix corresponding to the bright zone. Similarly, we can obtain the acoustic potential energy density of the wave domain over the dark zone. Therefore, the reproduced sound field over the entire region can be optimized by maximizing the ratio of the contrast among multiple zones in the wave domain as

$$\max_{\mathbf{q}} \tilde{C} = \frac{\tilde{E}_b}{\tilde{E}_d} = \frac{\mathbf{q}^H \mathbf{B}_b^H \mathbf{W}_b \mathbf{B}_b \mathbf{q}}{\mathbf{q}^H \mathbf{B}_d^H \mathbf{W}_d \mathbf{B}_d \mathbf{q}} = \frac{\mathbf{q}^H \tilde{\mathbf{R}}_b \mathbf{q}}{\mathbf{q}^H \tilde{\mathbf{R}}_d \mathbf{q}}, \tag{14}$$

where $\mathbf{W}_d$ is the diagonal weighting matrix corresponding to the dark zone. The wave-domain correlation matrix corresponding to the dark zone is defined as $\tilde{\mathbf{R}}_d = \mathbf{B}_d^H \mathbf{W}_d \mathbf{B}_d$, where $\mathbf{B}_d$ is the matrix of the decomposition coefficient of the ATFs between the loudspeaker array and the dark zone. The problem has the same form of the solution as the conventional ACC approach,

$$\mathbf{q}_{\mathrm{WDACC}} = \Phi\{(\tilde{\mathbf{R}}_d + \delta \mathbf{I})^{-1} \tilde{\mathbf{R}}_b\}. \tag{15}$$

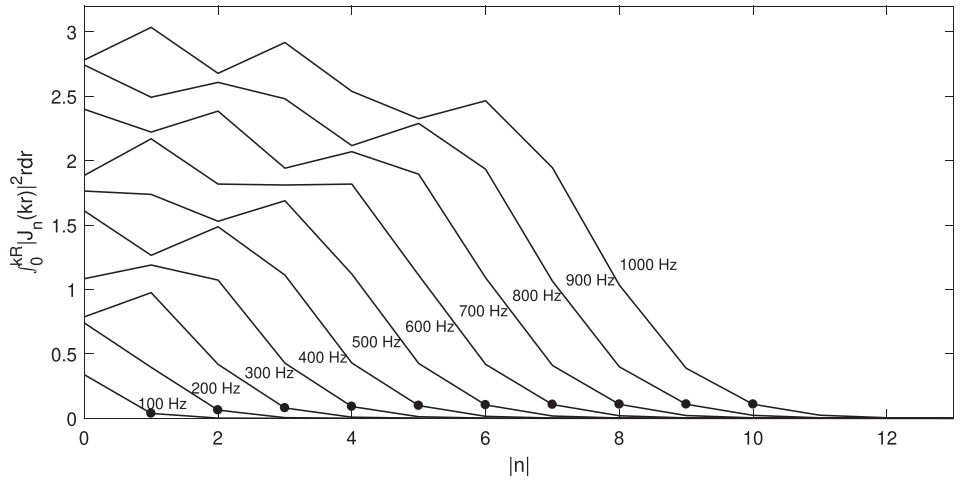

Fig. 1. Integrals $\int_0^{kR}|J_n(r)|^2rdr$ versus $|n|$ at different frequencies for a specific control region with radius $R=0.5$ m. The black dots denote the truncated order $|n|=\lceil kR\rceil$.

For a given circular array of L loudspeakers and a circular desired control region of radius R, the spatial Nyquist frequency of the array is[15]

$$f_{\mathrm{Nyq}}=\frac{c(L-1)}{4\pi R}. \tag{16}$$

This array is capable of reproducing a desired sound field up to f_{Nyq} for all radii $r\leqslant R$.

3. Simulation results

In the following examples, we placed 30 loudspeakers in a circle with radius $R_L=2$ m in a free field. The radii of a circular bright zone and a dark zone were $r_b=r_d=0.5$ m. Sixteen omni-directional microphones were used for each zone to measure the ATFs between the loudspeaker array and the desired regions. The loudspeakers were assumed to be simple point sources; thus, the ATF between the lth loudspeaker located at $\mathbf{r}_l$ and the control microphone located at $\mathbf{r}_m$ is equivalent to the steady-state Green's function in the free space: $g_l(\mathbf{r}_m)=(1/4\pi|\mathbf{r}_m-\mathbf{r}_l|)e^{-jk|\mathbf{r}_m-\mathbf{r}_l|}$. The distributions of the loudspeaker array and the microphone array were equiangular for simplicity. The center of the loudspeaker array was defined as the coordinate origin, and the centers of the microphone arrays in the bright and dark zones were located at (0.6, 0.5) m and (−0.5, −0.7) m, respectively. The observation region was a square area with a side length of 4 m. The system setup is illustrated in Fig. 2. To evaluate the performance over a spatial region, we considered the acoustic contrast and AG levels at all observation points inside the control region rather than at only 16 sampling points. To investigate the robustness of the system, additive random noise with a signal-to-noise ratio of 20 dB was considered when measuring the ATFs. For the sound reproduction, the amplitude and phase of multiplicative ATF errors that follow a Gaussian distribution with a standard deviation of 3 dB were also introduced.

The multizone reproduction was illustrated at single frequencies and over a range of frequencies up to the Nyquist frequency, which is approximately 1 kHz for the specific sound field reproduction system.

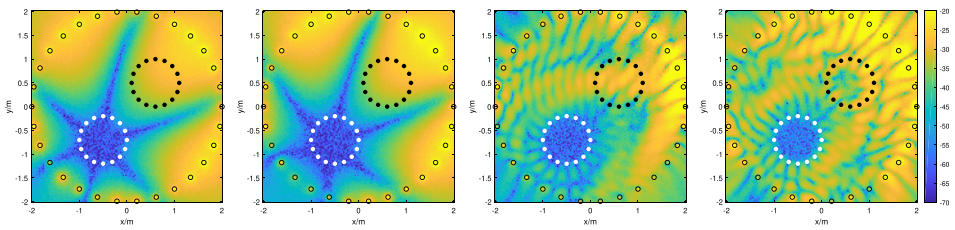

Fig. 2. (Color online) (a) and (b) present WDACC and ACC results for 100 Hz; the acoustic contrasts over the regions are 29.3 and 29.8 dB, respectively. (c) and (d) present WDACC and ACC results for 800 Hz; the acoustic contrasts are 25.9 and 20.9 dB, respectively. The open dots denote loudspeakers. The black dots and white dots denote control points in the bright and dark zones, respectively.

In the single-frequency case, 100- and 800-Hz multizone sound field reproductions were conducted. We chose the regularization parameters maximizing the average contrast performances by adopting Monte Carlo simulations which will be mentioned in the following broadband case. At 100 Hz, we chose $\delta = 3.981 \times 10^{-5}$ and 7.9433×10^{-4} for WDACC and ACC, respectively. At 800 Hz, $\delta = 1.2589 \times 10^{-4}$ and 2.5119×10^{-4} were used for the respective parameters of WDACC and ACC. The simulation results of one randomized trial are shown in Fig. 2. As shown in Figs. 2(a) and 2(b), at 100 Hz, both methods provide excellent multizone sound field separation performance for a specific regularization parameter. WDACC is more capable of maintaining an expected size dark zone than ACC, while the control goal is to maximize the contrast between the areas inside the two circular microphone arrays. According to Figs. 2(c) and 2(d), WDACC provides better results of bright and dark zone generation at 800 Hz. The acoustic contrasts of WDACC and ACC over the controlled regions are 25.9 and 20.9 dB, respectively. To achieve the same average pressure level in the bright zone, the dark zone of WDACC is 5 dB quieter than that of ACC. Furthermore, ACC fails to generate a uniform sound field in the bright zone, indicating that ACC controls multiple points instead of spatial regions.

Next, we considered a broadband robust performance evaluation with varying regularization parameters, using Monte Carlo simulations for assessment. The regularization parameter was varied from 10^{-9} to 10^{-2} at 71 logarithmically spaced values for each sampled frequency over the range of 50 Hz to 1 kHz. Since the additive noise from the microphones and the multiplicative error from the loudspeakers occur, one thousand Monte Carlo simulations were performed for each frequency and regularization parameter, and we then selected the best average contrast performance results at each frequency and the average AG levels with the same diagonal loading parameters. The resulting performances are demonstrated in Fig. 3. According to Figs. 3(a) and 3(b), the WDACC method outperforms the ACC method in terms of both acoustic contrast and AG effort with a well-chosen δ. As shown in Fig. 3, a notable feature of WDACC is that the best performing regularization parameters are numerically close in logarithm for a broadband frequency range. For two significantly different values of δ, WDACC yields a similar performance. This feature indicates the robustness of WDACC for an extensive range of regularization parameters due to the coefficient weighting functions that truncate higher-order basis functions.

4. Conclusion

A sound field control approach was proposed to reproduce local sound in the desired bright zone while minimizing acoustic potential energy density in the dark zone; this approach is based on orthogonal expansions of basis functions in the wave domain. The proposed method was compared with a conventional multiple points-based ACC approach. Simulations in a free field were implemented to evaluate these two methods in terms of acoustic contrast and AG over the regions rather than at discrete control points. The results demonstrate that the proposed WDACC produces better sound zone separation with less control effort cost over a broadband frequency range with well-chosen regularization parameters. This is because the wave-domain approach provides insight into the sound field reproduction problem. Since a finite number of expansion coefficients can represent the sound field over a large spatial region, the multizone sound can be manipulated by controlling the coefficients obtained with a small

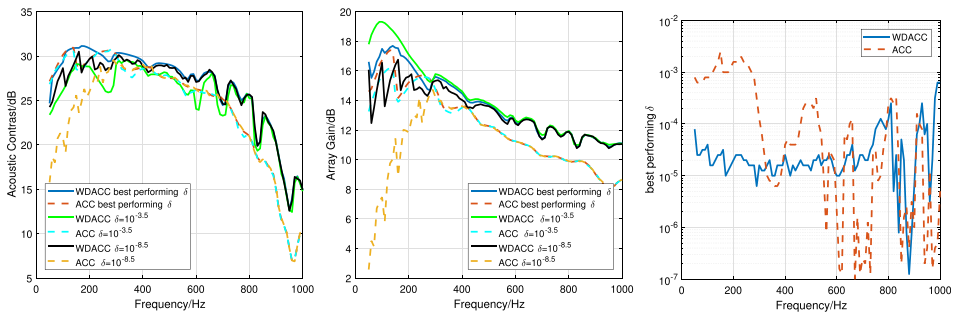

Fig. 3. (Color online) Best average performances from Monte Carlo simulations with 1000 trials and average performances for two values of δ that differ significantly in magnitude. (a) Acoustic contrast in decibels. (b) AG effort in decibels. (c) Regularization parameters that maximize the acoustic contrast at each frequency. Solid lines: WDACC, dashed lines: ACC.

number of microphones. Whereas the ACC method only maximizes the contrast between sound zones measured by multiple points, more microphones are necessary to control sound over large regions. Furthermore, WDACC demonstrates a similar performance over a range of diagonal loading parameters, indicating that the exact value of the regularization parameter does not have to be carefully determined.

Acknowledgment

This research was supported by the National Natural Science Foundation of China under Grant Nos. 11474306, 11404367, and 11474307, and the First Action Plan Program of Institute of Acoustics, Chinese Academy of Sciences.

References and links

[1]J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," J. Acoust. Soc. Am. **111**(4), 1695–1700 (2002).
[2]D. Kim, K. Kim, S. Wang, S. Q. Lee, and M. J. Crocker, "Maximization of the directivity ratio with the desired audible gain level for broadband design of near field loudspeaker arrays," J. Sound Vib. **330**(23), 5517–5529 (2011).
[3]S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," IEEE Trans. Audio, Speech, Lang. Process. **20**(7), 2123–2133 (2012).
[4]Q. Zhu, P. Coleman, M. Wu, and J. Yang, "Robust acoustic contrast control with reduced in-situ measurement by acoustic modeling," J. Audio Eng. Soc. **65**(6), 460–473 (2017).
[5]J. Cheer, S. J. Elliott, and M. F. S. Gálvez, "Design and implementation of a car cabin personal audio system," J. Audio Eng. Soc. **61**(6), 412–424 (2013).
[6]J. Cheer, S. J. Elliott, Y. Kim, and J.-W. Choi, "Practical implementation of personal audio in a mobile device," J. Audio Eng. Soc. **61**(5), 290–300 (2013).
[7]S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (2007), pp. 171–174.
[8]H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (2004), Vol. 4, pp. iv117–iv120.
[9]S. Spors, H. Buchner, and R. Rabenstien, "A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering," in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (2004), Vol. 4, pp. iv29–iv32.
[10]Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," IEEE Trans. Audio, Speech, Lang. Process. **19**(6), 1711–1720 (2011).
[11]W. Zhang, T. D. Abhayapala, T. Betlehem, and F. M. Fazi, "Analysis and control of multi-zone sound field reproduction using modal-domain approach," J. Acoust. Soc. Am. **140**(3), 2134–2144 (2016).
[12]E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, Waltham, MA, 1999).
[13]D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," IEEE Trans. Speech Audio Process. **9**(6), 697–707 (2001).
[14]I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," J. Acoust. Soc. Am. **130**(4), 2003–2015 (2011).
[15]M. Poletti and T. Abhayapala, "Interior and exterior sound field control using general two-dimensional first-order sources," J. Acoust. Soc. Am. **129**(1), 234–244 (2011).