Analysis and control of multi-zone sound field reproduction using modal-domain approach

Wen Zhang[a)]
Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University,
127 Youyi West Road, Xi'an, Shaanxi 710072, China

Thushara D. Abhayapala
Research School of Engineering, College of Engineering and Computer Science,
The Australian National University, Canberra, Australian Capital Territory 0200, Australia

Terence Betlehem
Digital Media and Communications R&D Centre, Samsung Electronics, Seoul 06765, South Korea

Filippo Maria Fazi
Institute of Sound and Vibration Research, University of Southampton, SO17 1BJ, United Kingdom

Multi-zone sound control aims to reproduce multiple sound fields independently and simultaneously over different spatial regions within the same space. This paper investigates the multi-zone sound control problem formulated in the modal domain using the Lagrange cost function and provides a modal-domain analysis of the problem. The Lagrange cost function is formulated to represent a quadratic objective of reproducing a desired sound field within the bright zone and with constraints on sound energy in the dark zone and global region. A fundamental problem in multi-zone reproduction is interzone sound interference, where based on the geometry of the sound zones and the desired sound field within the bright zone the achievable reproduction performance is limited. The modal-domain Lagrangian solution demonstrates the intrinsic ill-posedness of the problem, based on which a parameter, the coefficient of realisability, is developed to evaluate the reproduction limitation. The proposed reproduction method is based on controlling the interference between sound zones and sound leakage outside the sound zones, resulting in a suitable compromise between good bright zone performance and satisfactory dark zone performance. The performance of the proposed design is demonstrated through numerical simulations of two-zone reproduction in free-field and in reverberant environments. © 2016 Acoustical Society of America.
[http://dx.doi.org/10.1121/1.4963084]

I. INTRODUCTION

Multi-zone reproduction aims to extend spatial sound rendering over multiple regions of space so that different listeners would enjoy their audio material simultaneously and independently of each other but without physical isolation or using headphones. The concept of multi-zone sound field control has recently drawn attention due to a whole range of audio applications, such as controlling sound radiation from a personal audio device,[1] creating independent sound zones in different kinds of enclosures (such as shared offices, private transportation vehicles, exhibition centres, etc.),[2] and generating quiet zones in a noisy environment.[3] A single array of loudspeakers is used to control the reproduction of sound fields within multiple regions. This arrangement allows sound zones to be produced at any desired location and also the listener to freely move between zones thus can provide significant freedom and flexibility. A variety of techniques have been proposed for sound field synthesis over a fairly large region of space (or a single zone) using a loudspeaker array, such as the well-known Ambisonics, Near-Field Compensated Higher-Order Ambisonics, Wave Field Synthesis, and multipoint approach.[4] One of the current research interests in spatial audio is to further improve these techniques with a thorough perceptual assessment[5] and to use them for multi-zone sound field control applications.[6]

Choi and Kim first formulated multi-zone reproduction as creating two kinds of sound zones, the bright zone within which certain sounds with high acoustic energy are reproduced and the dark zone (or the quiet zone) within which the acoustic energy is kept at a low level.[7] The proposed method is to maximise the ratio of the average acoustic energy density in the bright zone to that in the dark zone, which is known as the *acoustic contrast control* (ACC) method. Since then, different forms of contrast control based on the same principle have been proposed, including an acoustic energy difference formulation,[8] direct and indirect acoustic contrast formulations using the Lagrangian.[9] The technique has been implemented in different personal audio systems in an anechoic chamber[1] or in a car cabin;[2] over 19 dB contrast was achieved under the ideal condition while for real-time systems in the car cabin the acoustic contrast was limited to

a)Also at: Research School of Engineering, College of Engineering and Computer Science, The Australian National University, Canberra, ACT 0200, Australia. Electronic mail: wenzhang@ieee.org

a maximum value of 15 dB. This contrast control method however does not impose a constraint on the phase of the sound field and thus cannot control the spatial aspects of the reproduced sound field in the bright zone. A recent work by Coleman *et al.* proposed to refine the cost function of the ACC with the aim of optimizing the extent to which the reproduced sound field resembles a plane wave, thus optimising the spatial aspects of the sound field.[10] Another issue in ACC is the self-cancellation problem which results in a standing wave produced within the bright zone.[11]

The *pressure matching* (PM) approach aims to reproduce a desired sound field in the bright zone while producing silence in the dark zone.[12] The approach uses a sufficiently dense distribution of microphones within all the zones as the matching points and adopts the least-squares method to control the pressure at each point. A constraint on the loudspeaker weight energy (or the array effort) is added to control the sound leakage outside the sound zones and to ensure the implementation is robust against speaker positioning errors and changes in the acoustic environment.[13] When the desired sound field in the bright zone is due to a few virtual source directions, the multi-zone sound control problem can be solved using a compressive sensing idea where the loudspeaker weights are regularised with the L_1 norm. This results in that only a few loudspeakers placed closely to the virtual source directions are activated for reproduction.[14,15] More recent works have been focusing on the combination of the ACC and PM formulations using the Lagrangian with a weighting factor to tune the trade-off between the two performance measures, i.e., the acoustic contrast and bright zone error (i.e., the reproduction error within the bright zone).[16–19] The idea of performing time-domain filters for personal audio was recently investigated.[20]

Modal-domain reproduction is based on representing the sound fields within different zones through a spatial harmonic expansion. The local sound field coefficients are then transformed to an equivalent global sound field using the harmonic translation theorem, from which the loudspeaker weights are obtained using the mode matching approach.[21,22] The modal-domain approach can provide insights into the multi-zone problem. For example, through the modal domain analysis, a theoretical basis is established for creating two sound zones with no interference.[23] Modal-domain sparsity analysis shows that a significantly reduced number of microphone points could be used quite effectively for multi-zone reproduction over a wide frequency range.[24] The synthesis of sound fields with distributed modal constraints and quiet zones having an arbitrary predefined shape have also been investigated.[25,26]

In this work, we investigate the multi-zone sound control problem formulated in the modal domain using the Lagrange cost function. Based on modal-domain analysis, a parameter, the *coefficient of realisability*, is developed to indicate the achievable reproduction performance given the sound zone geometry and the desired sound field in the bright zone. We then propose an algorithm to tune the reproduction performance between the dark zone and bright zone. This is achieved by defining two Lagrange multipliers in the modal-domain Lagrangian solution to control the

interference between sound zones and sound leakages outside the sound zones.

The paper considers the two-dimensional (2D) case for simplicity; however, the theory and results can be easily extended to the three-dimensional (3D) case. The differences between two approaches are that (1) in the 2D approach the loudspeaker radiation attenuates with $1/\sqrt{r}$ while the attenuation of a 3D point source is in accordance with $1/r$; and (2) the translation relationship between coefficients of local and global sound zones are based on the addition theorems of Bessel function and spherical Bessel function, respectively. Given the same reproduction setup, i.e., the number of sound zones and their size, the 3D approach that controls sound within a sphere would require more loudspeakers/microphones and a higher computational complexity.

II. REVIEW: MULTI-ZONE SOUND CONTROL IN MODAL DOMAIN

The objective of the general multi-zone problem is to produce a desired spatial sound field in each of Q non-overlapping sound zones. As shown in Fig. 1, we assume that each sound zone q has a radius R_q and its centre is denoted by O_q with respect to the global origin O. Any observation point within a sound zone is represented by x_q with respect to O_q, or $x = x_q + O_q$ with respect to O. We use $\| \cdot \|$ and $\hat{(\cdot)}$ to denote the magnitude and phase angle of a vector. All sound zones are within a general region of interest of radius $r_0 \geq r$, and the loudspeakers are placed on the boundary of this region.

A. Representing sound fields in local zones and global region

The sound fields at any point $x_q \equiv (\|x_q\|, \hat{x}_q)$ within a sound zone in 2D cylindrical coordinates can be expressed in the form
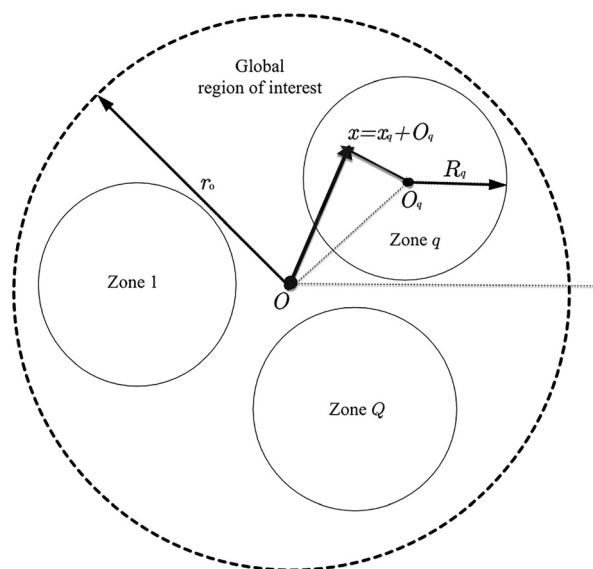

FIG. 1. Geometry of multi-zone sound control.

J. Acoust. Soc. Am. **140** (3), September 2016

Zhang *et al.* 2135

$$P^{(q)}(\boldsymbol{x}_q, k) \approx \sum_{n=-N_q}^{N_q} \alpha_n^{(q)}(k) J_n(k\|\boldsymbol{x}_q\|) e^{in\hat{x}_q}, \tag{1}$$

where $\hat{x}_q = \phi_{x_q}$, $k = 2\pi f/c$ is the wave number with f being the frequency and c the speed of sound. $J_n(k\|\boldsymbol{x}_q\|)$ is the cylindrical Bessel function of order n, and $\alpha_n^{(q)}(k)$ is the corresponding nth order sound field coefficient to describe a spatial sound field with respect to $\boldsymbol{O}_q$. Given the radius of the local sound zone R_q, wave number k, the truncation order is $N_q = \lceil ekR_q/2 \rceil$,[27] where $\lceil \cdot \rceil$ denotes the ceiling function and $e = \exp\{1\}$ is the Euler's number.

Analogous to Eq. (1), the sound field at the point $\boldsymbol{x} = (\|\boldsymbol{x}\|, \hat{x})$ in the global system can be represented by a finite number of cylindrical harmonics

$$P(\boldsymbol{x}, k) \approx \sum_{m=-N_0}^{N_0} \beta_m(k) J_m(k\|\boldsymbol{x}\|) e^{im\hat{x}}, \tag{2}$$

where $N_0 = \lceil ekr_0/2 \rceil$ and r_0 is the radius of the general region of interest including all sound zones.

Consider the sound field in zone q characterised by a set of coefficients $\alpha_n^{(q)}(k)$, these coefficients are with respect to the local origin $\boldsymbol{O}_q$ and can be related to the coefficients $\beta_m(k)$ with respect to the global origin $\boldsymbol{O}$ using the Bessel function addition theorem[28] (or the so-called translation theorem[22]), that is,

$$J_m(k\|\boldsymbol{x}\|) e^{im\hat{x}} = \sum_{n=-N_q}^{N_q} J_{m-n}(kr_q) e^{i(m-n)\phi_q} J_n(k\|\boldsymbol{x}_q\|) e^{in\hat{x}_q}, \tag{3}$$

given $\boldsymbol{x} = \boldsymbol{x}_q + \boldsymbol{O}_q$ and $\boldsymbol{O}_q \equiv \{r_q, \phi_q\}$ in the global system.

Thus, we have the following relation:

$$\alpha_n^{(q)}(k) = \sum_{m=-N_0}^{N_0} \mathcal{T}_{nm}(\boldsymbol{O}_q, k) \beta_m(k), \tag{4}$$

where

$$\mathcal{T}_{nm}(\boldsymbol{O}_q, k) = J_{m-n}(kr_q) e^{i(m-n)\phi_q}. \tag{5}$$

Using the matrix-vector notation, we can represent this relationship in a compact form as

$$\boldsymbol{A}_q = \boldsymbol{T}_q \boldsymbol{B}, \tag{6}$$

where $\boldsymbol{A}_q = [\alpha_{-N_q}^{(q)}(k), \ldots, \alpha_{N_q}^{(q)}(k)]^T$ and $\boldsymbol{B} = [\beta_{-N_0}(k), \ldots, \beta_{N_0}(k)]^T$ are column vectors of length $(2N_q + 1)$ and $(2N_0 + 1)$, respectively. $\boldsymbol{T}_q$ is the $(2N_q + 1) \times (2N_0 + 1)$ matrix representing the translation from the global system to the local system, that is $[\boldsymbol{T}_q]_{\imath,\jmath} = \mathcal{T}_{nm}(\boldsymbol{O}_q, k)$ with $\imath = m + N_q + 1$ and $\jmath = n + N_0 + 1$. Note that since each zone lies within the general region of interest (or the global sound zone), $N_0 \geq N_q$. The wave number k has been deleted for notational simplicity.

B. Problem formulation

We now want to find the global sound field coefficients $\beta_n(k)$ from Q sets of local sound field coefficients

using Eq. (6). Three objective performance measures are defined for multi-zone sound control. These are (i) the *acoustic contrast*, that is the ratio of the average acoustic energy density in the bright zone to that in the dark zone,[7] (ii) the *bright zone error*, that is the normalised spatial average error between the desired and reproduced sound fields in the bright zone,[12] and (iii) the *array effort*, that is proportional to the loudspeaker power consumption to generate the desired multi-zone field.[29]

In this work, we take into account all three objective performance measures for multi-zone reproduction and adopt the formulation using the combination of mode matching and ACC to calculate the global sound field coefficients. The multi-zone reproduction problem is formulated as finding the global sound field coefficients $\boldsymbol{B}$ to generate a desired sound field in the bright zone $\mathbb{D}_b$ characterised by its local coefficients $\boldsymbol{A}_b$ with constraints on the sound energy in the dark zone $\mathbb{D}_d$ and the energy of the entire global sound field no more than e_d and e_g, respectively. Thus, the problem we consider in this paper is as follows:

$$\min_{\boldsymbol{B}} \; \|\boldsymbol{T}_b \boldsymbol{B} - \boldsymbol{A}_b\|^2, \tag{7}$$

$$\text{subject to } \|\boldsymbol{T}_d \boldsymbol{B}\|^2 \leq e_d, \tag{7a}$$

$$\|\boldsymbol{B}\|^2 \leq e_g. \tag{7b}$$

$\boldsymbol{T}_b$ and $\boldsymbol{T}_d$ are the translation matrices of the local bright zone and dark zone sound field coefficients from the global system, respectively.

The presented formulation imposes a bound on the global sound energy $\|\boldsymbol{B}\|^2$. As detailed in Sec. IV B, this is related to the array effort constraint. Equation (17) shows that the norm of the loudspeaker weights are controlled by two elements, one is the norm of the entire global sound field coefficients $\|\boldsymbol{B}\|$ and the other is the norm of the channel transfer matrix from the produced sound field to the loudspeaker weights. Thus a constraint on $\|\boldsymbol{B}\|^2$ controls the effect of this vector on the array effort. In addition, as the channel transfer matrix can be potentially ill-conditioned depending on the loudspeaker geometries and reproduction environment, a regularisation is required to derive the loudspeaker driving signals given the global sound field coefficients.

III. LAGRANGIAN FORMULATION AND PROBLEM ANALYSIS

We write the optimisation problem posed in Eq. (7) as a Lagrange cost function

$$\begin{aligned} \boldsymbol{B}_{\text{opt}} = \underset{\boldsymbol{B}}{\text{argmin}}\, L(\boldsymbol{B}) &= \|\boldsymbol{T}_b \boldsymbol{B} - \boldsymbol{A}_b\|^2 \\ &+ \lambda_1(\|\boldsymbol{T}_d \boldsymbol{B}\|^2 - e_d) + \lambda_2(\|\boldsymbol{B}\|^2 - e_g), \end{aligned} \tag{8}$$

where λ_1 and λ_2 are positive Lagrange multipliers. The solution that minimises $L(\boldsymbol{B})$ is obtained by setting the derivative of $L(\boldsymbol{B})$ with respect to $\boldsymbol{B}$ to zero

$$[\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{T}_{\mathrm{b}} + \lambda_1 \boldsymbol{T}_{\mathrm{d}}^{*}\boldsymbol{T}_{\mathrm{d}} + \lambda_2 \boldsymbol{I}]\boldsymbol{B} = \boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}, \tag{9}$$

which yields

$$\boldsymbol{B} = [\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{T}_{\mathrm{b}} + \lambda_1 \boldsymbol{T}_{\mathrm{d}}^{*}\boldsymbol{T}_{\mathrm{d}} + \lambda_2 \boldsymbol{I}]^{-1}\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}, \tag{10}$$

where $(\cdot)^{*}$ denotes the Hermitian transpose and $\boldsymbol{I}$ is an identity matrix of dimension $2N_0 + 1$. Note that the matrix to be inverted in Eq. (10) is of dimension $(2N_0 + 1) \times (2N_0 + 1)$ and scales with the size of the global sound reproduction region. This formulation can easily be extended to the case of Q sound zones, by augmenting additional dark zone constraints of the form of Eq. (7a) to the Lagrange cost function [Eq. (8)]. Extending the formulation as such requires solving for the greater number of Lagrange multipliers. In Sec. IV A, the Newton's method is used to solve for Lagrange multipliers that satisfy the inequality constraints.

We first have the following comments:

- *Mode matching vs ACC*: The formulation in Eq. (10) provides a great deal of flexibility to control multi-zone sound fields. The Lagrange multiplier λ_1 determines the trade-off between mode matching and ACC. For example, the case of $\lambda_1 = 1$ implies that equal effort is made to match the modes in the bright zone and to minimise the energy in the dark zone. Decreasing the values of λ_1 results in high reproduction accuracy in the bright zone while increasing this value means that more emphasis is put on creating a low level of sound in the dark zone, thus achieving an improved acoustic contrast between zones.

- *Robustness issues*: The Lagrange multiplier λ_2 introduced for limiting the global sound field energy can be interpreted as a partial constraint on the array effort. Given the array effort is related to sound levels in the room, this constraint can ensure the sound leakage outside the Q zones is not excessive. Referring to Eq. (10), the first two terms $\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{T}_{\mathrm{b}}$, $\boldsymbol{T}_{\mathrm{d}}^{*}\boldsymbol{T}_{\mathrm{d}}$ in the matrix to be inverted are not full rank; thus the multi-zone reproduction problem is intrinsically ill-conditioned. In addition, when implemented in a room environment, room compensation requires the knowledge of the acoustic transfer function (ATF) and its estimation can be potentially a problem for robustness. The Lagrange multiplier λ_2 plays a similar role as that of the regularisation parameters in the standard least square solutions to the inverse problem for improving the conditioning of the problem.

- *Dimensionality analysis*: In the proposed solution [Eq. (10)], multi-zone sound field reproduction is controlled by the sound field coefficients and translation matrices corresponding to each sound zone. Based on manipulating the sound field coefficients, the modal-domain approach can control the sound field within a region. Representing the sound field in the modal domain also shows fundamental properties of the multi-zone problem. For example, the number of parameters required to represent the sound field at wave number k within the qth sound zone, $\mathcal{N}_q = 2N_q + 1$, is termed as *dimensionality* of the region.[22] Given all sound zones are confined to a general circular region of radius r_0 with dimensionality $\mathcal{N} = 2N_0 + 1$, we need

$$2N_0 + 1 \geq \sum_{q=1}^{Q}(2N_q + 1), \tag{11}$$

so that sufficient degrees of freedoms are provided to control the sound field in each zone. Assuming that the loudspeaker has an omnidirectional radiation pattern, the value of $\mathcal{N}$ determines the minimum number of loudspeakers required for reproduction. In terms of the measurement requirements, the dimensionality and the order of each microphone together determine the number of microphones required for each zone and for the global region. A practical design has been presented for reproduction inside a room given the global sound field coefficients and the dimensionality results.[30]

- *Coefficient of realisability*: Multi-zone reproduction is fundamentally constrained whenever attempting to reproduce a sound field in the bright zone that has components directed towards the dark zone. This is known as the *occlusion problem*.[12,31] Usually, a prior knowledge of the desired sound propagation direction and the sound zone geometry is required to predict this problem. Here, we develop a parameter, the *coefficient of realisability*, as an indicator of the reproduction limitation, which is

$$\eta = 1 - \frac{\|\boldsymbol{T}_{\mathrm{d}}\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}\|}{\|\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}\|}, \tag{12}$$

where $\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}$ represents the energy projection of the desired bright zone sound field into the global region ($\boldsymbol{T}_{\mathrm{b}}$ and $\boldsymbol{T}_{\mathrm{b}}^{*}$ are one-sided inverse as shown in the Appendix), whose norm represents the energy projection. Similarly, $\|\boldsymbol{T}_{\mathrm{d}}\boldsymbol{T}_{\mathrm{b}}^{*}\boldsymbol{A}_{\mathrm{b}}\|$ represents the energy leakage of a desired bright zone sound field into the dark zone.

This coefficient ranges from 0 to 1, where 0 indicates complete bright zone sound leakage, i.e., the multi-zone filed cannot be realised, and 1 indicates no sound leakage at all, or the best possible performance. Figure 2 plots η as a function of the angle between the vector of the virtual source direction and the vector from the dark zone origin
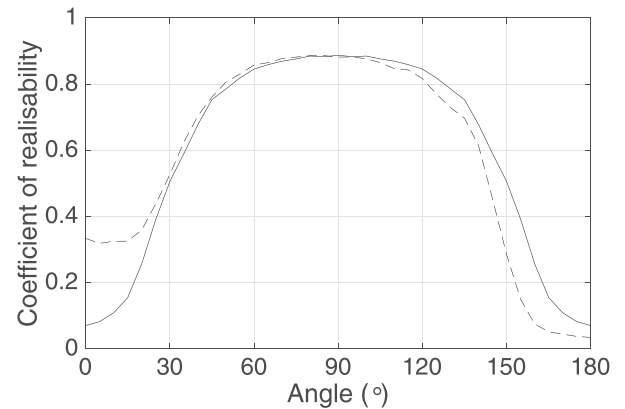

FIG. 2. The coefficient of realisability η as a function of the angle between the vector of the virtual source direction and the vector from the dark zone origin to the bright zone origin (two sound zones are assumed). Angles of $0°$ and $180°$ represent the occlusion angle in the bright zone and dark zone side. The solid line and dashed line correspond to the far-field plane wave and the near-field cylindrical wave, respectively.

to the bright zone origin (two sound zones are assumed). As shown in the figure, for the occlusion angle, the value of η is lowest. If the desired sound is a plane wave, given the same sound zone geometry, the sound leakage is identical when the bright zone and dark zone are switched. This is demonstrated by the symmetry in the coefficient of realisability. However, when the virtual source is in the near field, or the desired sound has a curved wavefront, the symmetry property does not hold especially at occlusion angles and their surroundings. It is observed that when the virtual source direction comes from the bright zone side, the coefficient tends to have a higher value indicating better achievable performance. Notice that the developed parameter, the coefficient of realisability, is only determined by the reproduced sound fields and sound zone geometry, and is not affected by the loudspeaker array geometry and ATF of the reproduction environment. Provided that the loudspeaker geometry is appropriate and active listening room compensation is introduced, reproduction performance can hence be made independent of the reverberation.[6]

In terms of reproducing multiple bright-zone sound fields, there are two options: (i) calculating the coefficient of realisability for the whole system by augmenting additional bright zone information into the translation matrix $\boldsymbol{T}_b$ and the coefficient matrix $\boldsymbol{A}_b$; (ii) calculating the coefficient of realisability for each zone, based on which the design criteria can be further refined, for example, giving the priority to the more realisable zone. Even though the second case is not the scope of this work, the presented work can easily be extended to the case of multiple bright zones.

IV. IMPLEMENTATION ISSUES

This section deals with practical implementation issues, such as (i) choosing the value of Lagrange multipliers to control the interference between sound zones and sound leakages outside the sound zones and (ii) calculating the loudspeaker weights given the global sound field coefficients.

A. Control of reproduction performance

The Lagrange multipliers λ_1 and λ_2 are used to adjust the relative importance of constraints (7a) and (7b) for creating a very low level of sound pressure inside the dark zones and a limit on the global sound field energy. Considering the intimate relationship between the global sound field and dark zone sound field, the values of λ_1 and λ_2 should be determined jointly. Here, Newton's method is used to solve the Lagrange multipliers λ_1 and λ_2, which are guaranteed to converge at a quadratic rate.[32] In Table I, the Algorithm is proposed to satisfy the constraints (7a) and (7b) in equality so that the interference between sound zones and the sound leakages outside sound zones can be controlled even when the reproduction within the bright and dark zones is severely inconsistent. Figure 3 demonstrates its performance in the occlusion problem compared with the perfect bright zone performance, perfect dark zone performance,[23] and the reference modal-domain approach without the energy constraints.[22]

B. Loudspeaker array design

Given the global sound field coefficients, the desired global sound field is readily obtained from Eq. (2). An array of L loudspeakers is used for reproduction and its generated global sound field can be written as

$$\hat{P}(\boldsymbol{x}, k) = \sum_{\ell=1}^{L} w_\ell(k) G_\ell(\boldsymbol{x}, k), \tag{13}$$

where $G_\ell(\boldsymbol{x}, k)$ represents the ATF between the ℓth loudspeaker and the observation point $\boldsymbol{x}$ in the global system and w_ℓ is the loudspeaker weight. The ATF can also be parameterised in the modal domain as

$$G_\ell(\boldsymbol{x}, k) = \sum_{m=-N_0}^{N_0} \gamma_{m\ell}(k) J_m(k\|\boldsymbol{x}\|) e^{im\hat{x}}, \tag{14}$$

where $\gamma_{m\ell}(k)$ are ATF coefficients and assumed to be a prior knowledge obtained from theoretical solutions or pre-calibration.[30,33] For example, for 2D sound propagation in

---

TABLE I. **Algorithm**: Determine Lagrange multipliers λ_1 and λ_2 for controlling sound energy in the dark zone and global region.

1. From Eq. (10), obtain the global sound field coefficients $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{T}_b^*\boldsymbol{A}_b$, and dark zone sound field coefficients $\boldsymbol{A}_d = \boldsymbol{T}_d\boldsymbol{B} = \boldsymbol{T}_d\boldsymbol{Q}^{-1}\boldsymbol{T}_b^*\boldsymbol{A}_b$, where $\boldsymbol{Q} \triangleq \boldsymbol{T}_b^*\boldsymbol{T}_b + \lambda_1\boldsymbol{T}_d^*\boldsymbol{T}_d + \lambda_2\boldsymbol{I}$.

2. Define the dark zone sound energy $f_d = \boldsymbol{A}_d^*\boldsymbol{A}_d = \boldsymbol{A}_b^*\boldsymbol{T}_b\boldsymbol{Q}^{-1}\boldsymbol{T}_d^*\boldsymbol{T}_d\boldsymbol{Q}^{-1}\boldsymbol{T}_b^*\boldsymbol{A}_b$.

3. Define the global sound energy $f_g = \boldsymbol{B}^*\boldsymbol{B} = \boldsymbol{A}_b^*\boldsymbol{T}_b\boldsymbol{Q}^{-1}\boldsymbol{Q}^{-1}\boldsymbol{T}_b^*\boldsymbol{A}_b$.

4. Determine λ_1 and λ_2 jointly to satisfy the dark zone sound energy constraint (7a) and the global sound energy constraint (7b) with equality, that is $f_d = e_d$ and

$f_g = e_g$. The Newton's update formula is $\begin{bmatrix} \lambda_1^{n+1} \\ \lambda_2^{n+1} \end{bmatrix} = \begin{bmatrix} \lambda_1^n \\ \lambda_2^n \end{bmatrix} - \boldsymbol{J}^{-1}\begin{bmatrix} f_d - e_d \\ f_g - e_g \end{bmatrix}$, where $\boldsymbol{J} = \begin{bmatrix} \frac{\partial f_d}{\partial \lambda_1} & \frac{\partial f_d}{\partial \lambda_2} \\ \frac{\partial f_g}{\partial \lambda_1} & \frac{\partial f_g}{\partial \lambda_2} \end{bmatrix}$, and $\frac{\partial f_d}{\partial \lambda_1} = -2\boldsymbol{A}_d^*\boldsymbol{T}_d\boldsymbol{Q}^{-1}\boldsymbol{T}_d^*\boldsymbol{A}_d$, $\frac{\partial f_d}{\partial \lambda_2} = -\boldsymbol{A}_b^*\boldsymbol{T}_b\boldsymbol{Q}^{-2}\boldsymbol{T}_d^*\boldsymbol{A}_d$

$-\boldsymbol{A}_d^*\boldsymbol{T}_d\boldsymbol{Q}^{-2}\boldsymbol{T}_b^*\boldsymbol{A}_b$, $\frac{\partial f_g}{\partial \lambda_1} = -\boldsymbol{B}^*\boldsymbol{T}_d^*\boldsymbol{T}_d\boldsymbol{Q}^{-1}\boldsymbol{B} - \boldsymbol{B}^*\boldsymbol{Q}^{-1}\boldsymbol{T}_d^*\boldsymbol{T}_d\boldsymbol{B}$, $\frac{\partial f_g}{\partial \lambda_2} = -2\boldsymbol{B}^*\boldsymbol{Q}^{-1}\boldsymbol{B}'$.

The resulting update implementation for λ_1 and λ_2 is

Initialization $n = 0$, $\lambda_1^0 \geq 0$, $\lambda_2^0 \geq 0$

Repeat $\begin{bmatrix} \lambda_1^{n+1} \\ \lambda_2^{n+1} \end{bmatrix} = \begin{bmatrix} \lambda_1^n \\ \lambda_2^n \end{bmatrix} - \boldsymbol{J}^{-1}(\lambda_1^n, \lambda_2^n)\begin{bmatrix} f_d(\lambda_1^n, \lambda_2^n) - e_d \\ f_g(\lambda_1^n, \lambda_2^n) - e_g \end{bmatrix}$,

Until $\|\boldsymbol{A}_d(\lambda_1^{n+1}, \lambda_2^{n+1}) - \boldsymbol{A}_d(\lambda_1^n, \lambda_2^n)\| < \epsilon_d$ for a small ϵ_d and $\|\boldsymbol{B}(\lambda_1^{n+1}, \lambda_2^{n+1}) - \boldsymbol{B}(\lambda_1^n, \lambda_2^n)\| < \epsilon_g$ for a small ϵ_g
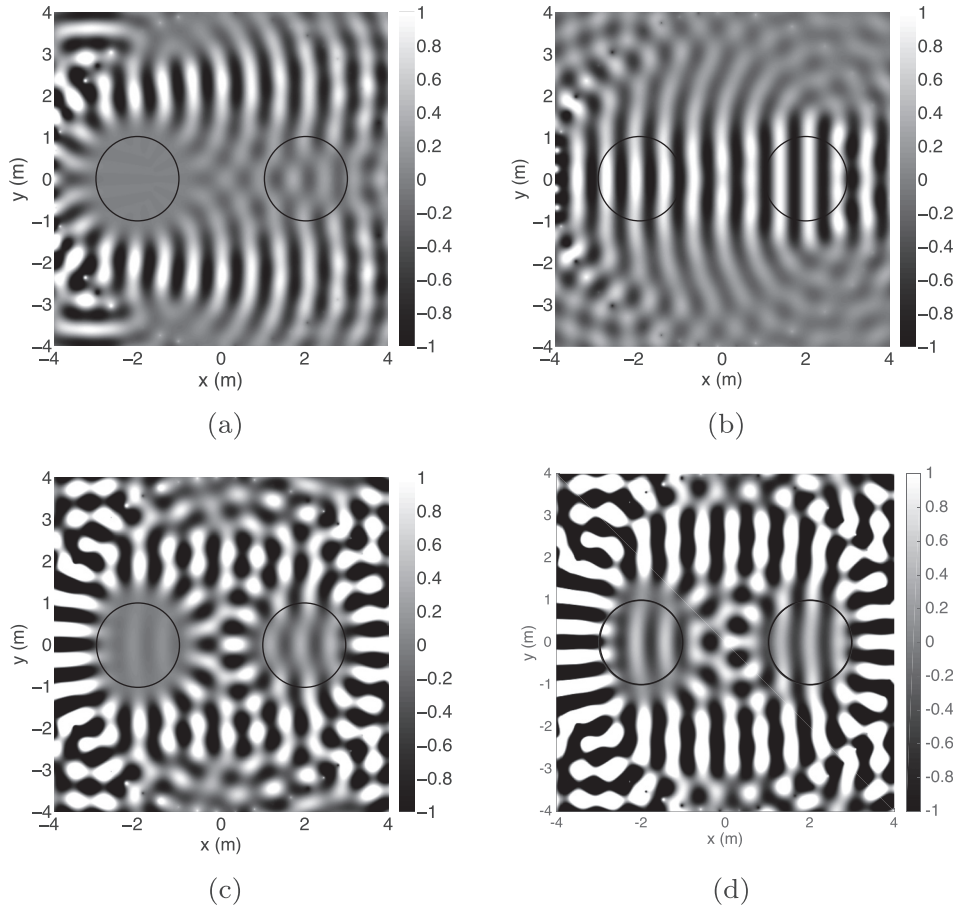

FIG. 3. Multi-zone sound control for the occlusion problem. The desired field in the bright zone is due to a plane wave from $0°$ for a frequency of 500 Hz. The bright zone and dark zone are located at (2,0) and (−2,0), respectively. Shown is (a) the reproduction with the sound field components that contribute to the sound in the dark zone are completely removed (Ref. 23), (b) the reproduction resulted by letting all the sound field components leak from the bright zone into the dark zone, (c) the reproduction using the proposed modal-domain control with energy constraints, and (d) the reproduction using the reference modal-domain approach without energy constraints (Ref. 22).

the free field, $G_\ell(\boldsymbol{x}, k) = (i/4)H_0^{(1)}(k\|\boldsymbol{y}_\ell - \boldsymbol{x}\|)$ and $\gamma_{m\ell}(k) = (i/4)H_m^{(1)}(k\|\boldsymbol{y}_\ell\|)e^{-im\hat{\boldsymbol{y}}_\ell}$, where $\boldsymbol{y}_\ell \equiv \{\|\boldsymbol{y}_\ell\|, \hat{\boldsymbol{y}}_\ell\}$ is the loudspeaker position and $H_m^{(1)}(\cdot)$ denotes the first kind Hankel function of order m.

Based on Eqs. (2), (13), and (14), we can design loudspeaker weights $w_\ell(k)$ to match every angular mode of a global sound field, i.e.,

$$\sum_{\ell=1}^{L} w_\ell(k)\gamma_{m\ell}(k) = \beta_m(k), \text{ for } m = -N_0, \ldots, N_0. \quad (15)$$

Representing Eq. (15) in matrix form

$$\boldsymbol{\Gamma W} = \boldsymbol{B}, \quad (16)$$

where $\boldsymbol{W} = [w_1(k), \ldots, w_L(k)]^T$ is the loudspeaker weight vector, $\boldsymbol{B}$ is the vector of the global sound field coefficients, and $\boldsymbol{\Gamma}$ is a matrix of ATF coefficients $\gamma_{m\ell}(k)$ of size $(2N_0 + 1) \times L$.

The least-squares method[34] can then be applied to calculate the loudspeaker weights, that is

$$\boldsymbol{W} = \boldsymbol{\Gamma}^\dagger \boldsymbol{B}, \quad (17)$$

where $\boldsymbol{\Gamma}^\dagger = [\boldsymbol{\Gamma}^*\boldsymbol{\Gamma} + \lambda\boldsymbol{I}]^{-1}\boldsymbol{\Gamma}^*$ is the Moore-Penrose (Pseudo) inverse of $\boldsymbol{\Gamma}$. Equation (17) shows that the array effort, or the norm of the loudspeaker weights $\|\boldsymbol{W}\|$, is controlled by two parts, the norm of the entire global sound field $\|\boldsymbol{B}\|$ and the norm of the channel transfer matrix from the generated

sound field to the driving signals, $\|\boldsymbol{\Gamma}^\dagger\|$. In this paper, we focus on controlling the norm of the generated sound field. The matrix $\boldsymbol{\Gamma}$ however can be ill-conditioned too, especially at low frequencies when the acoustic wavelength is large relative to the size of the control zones; a small amount of Tikhonov regularisation λ is applied to improve the robustness of the solution. λ is also associated with the loudspeaker array geometry and how they couple with the desired sound fields in the propagation environment. This means that the matrix $\boldsymbol{\Gamma}$ can also be ill-conditioned for irregular or poorly chosen loudspeaker array geometries or environmental conditions although this is not the scope of the work here.

V. SIMULATION RESULTS

This section provides a verification and evaluation of the proposed modal-domain multi-zone reproduction with the corresponding results introduced.

A. Simulation setup

We simulate two-zone reproduction examples in free-field and in a reverberant room of size 10×9 m. Room reverberation is simulated using the image source method[35] with the image order up to 5 (i.e., 60 image sources) for each loudspeaker. The wall reflection coefficients are 0.7 and a perfectly-absorbing surface is assumed for floor and ceiling. A circular array of 55 loudspeakers at a radius of 4 m is used to generate the two-zone sound field. The desired sound field

J. Acoust. Soc. Am. **140** (3), September 2016

Zhang et al. 2139

in the bright zone is due to a plane wave at the operating frequency of 500 Hz.

To evaluate the performance, we use the following three objective measures as introduced in Sec. II B:

- The acoustic contrast between the sound fields generated in the bright zone $\mathbb{D}_b$ and dark zone $\mathbb{D}_d$ is used to quantify the sound leakage between two zones[7]

$$\chi(k) = \frac{\frac{1}{\mathbb{V}_b}\int_{\mathbb{D}_b}|\hat{P}(\boldsymbol{x},k)|^2 d\boldsymbol{x}}{\frac{1}{\mathbb{V}_d}\int_{\mathbb{D}_d}|\hat{P}(\boldsymbol{x},k)|^2 d\boldsymbol{x}}, \quad (18)$$

where $\mathbb{V}_b$ and $\mathbb{V}_d$ represent the areas of the zones $\mathbb{D}_b$ and $\mathbb{D}_d$, respectively.

- The relative mean square error between the desired sound field $P(\boldsymbol{x},k)$ and reproduced sound field $\hat{P}(\boldsymbol{x},k)$ within the bright zone $\mathbb{D}_b$ is used as the error metric, i.e., the bright zone error[12]

$$\varepsilon(k) = \frac{\int_{\mathbb{D}_b}|\hat{P}(\boldsymbol{x},k) - P(\boldsymbol{x},k)|^2 d\boldsymbol{x}}{\int_{\mathbb{D}_b}|P(\boldsymbol{x},k)|^2 d\boldsymbol{x}}. \quad (19)$$

- The array effort or the loudspeaker weight energy is defined to quantify the total output for generating the desired sound effect[29]

$$e_w(k) = \boldsymbol{W}^H \boldsymbol{W}. \quad (20)$$

The developed parameter, the coefficient of realisability in Eq. (12), is examined as an indicator of the multi-zone reproduction performance.

B. Two-zone example

In the two-zone example, the bright zone and dark zone are located at $\boldsymbol{O}_1 = (2, 0)$ and $\boldsymbol{O}_2 = (-2, 0)$ with respect to the global origin, respectively. Each sound zone has a radius of 1 m. We first set the constraints on the sound energy in the dark zone and in the global region to $e_d = -40$ dB and $e_g = 10$ dB, respectively. These two parameters are defined in Eqs. (7a) and (7b) and in particular, e_d is determined either by a threshold on the sound level within the dark zone or based on the requirement of a certain contrast level between sound zones. For personal audio, perceptual requirements should be taken into account for setting the values of e_d and e_g, such as the noticeable level of reproduction error within the bright zone and the acceptable level of audio interference reduction within the dark zone.[5] In addition, these two parameters may be frequency and audio dependent.[36] For a practical implementation, the array effort should be considered as well with the knowledge of the ATF.

The real parts of the reproduced sound fields for the propagation direction of 0° and 90° in the free field and reverberant room are shown in Fig. 4. When the desired sound comes from the direction 90°, which is perpendicular to the line drawn through the centres of the zones, the coefficient of realisability η as shown in Fig. 2 is around 0.9 indicating that this is a well conditioned problem. Figure 4 shows that this setting produces a very small bright zone error in both free field and reverberant environments. However, the achievable ACC in the free field is about 20 dB higher than that in a reverberant room. In the other example, where the incident angle is 0°, the coefficient of realisability η is less than 0.1 (corresponding to the solid line in Fig. 2), indicating a high level of sound

(a) $\chi = 48$ dB, $\varepsilon = -28$ dB, $e_w = 21$ dB   (b) $\chi = 29$ dB, $\varepsilon = 0$ dB, $e_w = 21$ dB

(c) $\chi = 26$ dB, $\varepsilon = -28$ dB, $e_w = 45$ dB   (d) $\chi = 24$ dB, $\varepsilon = 0$ dB, $e_w = 39$ dB
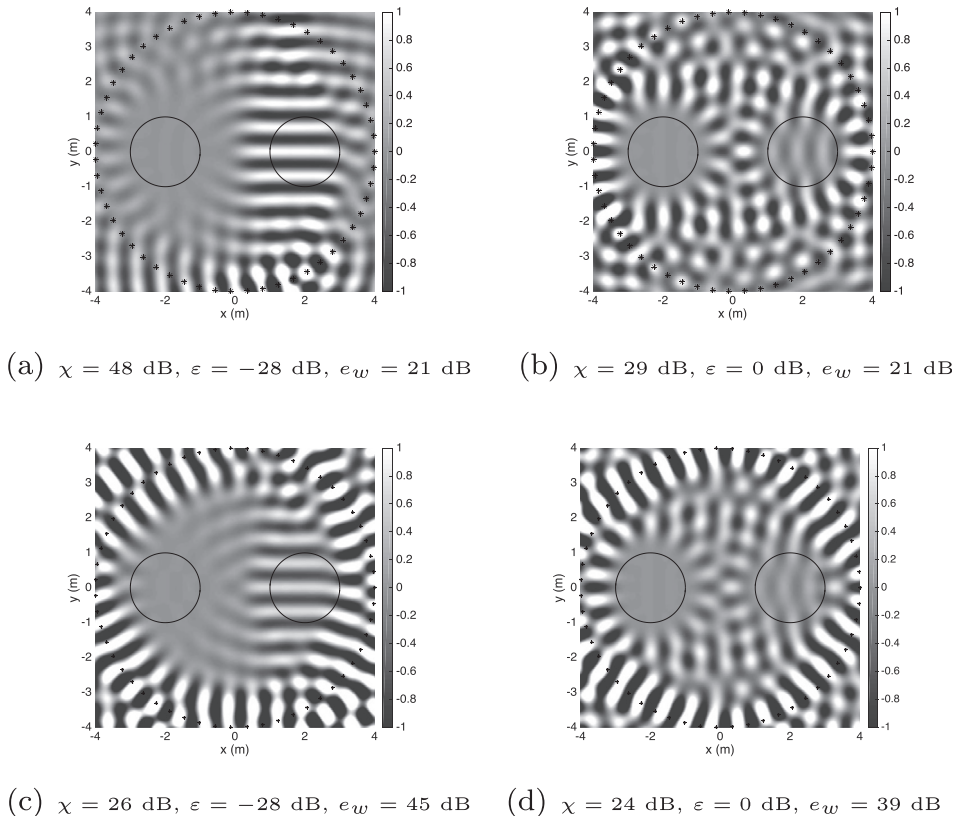
FIG. 4. Example of a two-zone reproduction where a plane wave of frequency 500 Hz is reproduced within the bright zone using a circular array of 55 loudspeakers with energy constraints $e_d = -40$ dB and $e_g = 10$ dB. The plane wave is from the direction of 90° in (a) and (c) or 0° in (b) and (d). The bright zone and dark zone of radius 1 m are located at $\boldsymbol{O}_1 = (2, 0)$ and $\boldsymbol{O}_2 = (-2, 0)$ with respect to the global origin. The encircled regions are the sound zones, the stars represent the loudspeaker positions, and the reproduction results are shown in a free field (upper row) or a reverberant room (lower row). The results of performance measures defined in Eqs. (18)–(20) are given at the subplot caption.

interference between the bright and dark zone. Setting a small value for the constraint on the dark-zone sound energy forces the creation of a quite zone; however, the wave fronts produced in the desired bright zone has artefacts.

C. Performance analysis

We first investigate the reproduction performance at different virtual source directions. Figure 5 plots the results under a variety of sound energy constraints as the virtual source incident angle in the example of Fig. 4 changes from $0°$ to $180°$. The result presented here corresponds to the free-field case. Comparing with the performance indicated by the coefficient of realisability the solid line in Fig. 2, we can observe that when the angle is in-line with both zones ($0°$ and $180°$), i.e., the angles of occlusion, the coefficient of realisability η reaches the lowest value and under the same optimisation constraint the system has the worst performance. Similarly, when the incident angle is around $90°$, η reaches the highest value; we have the best reproduction performance, i.e., the smallest bright zone error and the highest acoustic contrast simultaneously. In terms of practical implementation, the system performance depends heavily on the optimization constraints, reverberation conditions, and accuracy of ATF estimation. The proposed parameter, the coefficient of realisability, however is developed as an indicator of the multi-zone reproduction limitation based on a quantitative analysis of audio interference between sound zones. For example, for the case of smaller sized sound zones and larger spacing between zones, there would be

lower audio interference between sound zones and thus a higher value of the coefficient of realisability.

Next, we examine how sound energy constraints, which determine the value of Lagrangian multipliers λ_1 and λ_2 using the Algorithm in Table I, would affect the reproduction performance. First, we fix the global sound energy constraint e_g at $10\,dB$ and decrease the constraint e_d on the dark-zone sound energy from -20 to $-40\,dB$. The lower value of e_d means a higher priority of creating low-level sounds inside the dark zone and is normally associated with a larger value of λ_1. A higher acoustic contrast can be achieved for all cases including the occlusion angle. For example, in Fig. 5(b), more than $25\,dB$ acoustic contrast is achieved for the angle of occlusion when $e_d \leq -30\,dB$. On the other hand, when the dark zone energy e_d is fixed to $-20\,dB$, increasing the value of e_g from 5 to $15\,dB$ gives reduced bright zone error and higher acoustic contrast. This is due to the fact that a relaxed limit on the global sound energy gives a low value of λ_2 and the priority to satisfy the bright zone and dark zone design criteria. However, for a practical implementation it is necessary to limit the global sound zone energy so that several undesirable characteristics in the multi-zone reproduction system can be minimized. These are (1) the sound levels outside the sound zones and the required array effort, (2) the level of reverberation outside the global sound zone, and (3) the sensitivity to the ill-posedness of the problem.

D. Discussion

The proposed method adopts a very similar concept to PM for controlling both magnitude and phase of the reproduced sound field within the bright zone but using the modal-domain approach, i.e., the so-called "mode matching." A Lagrangian is formulated in the modal domain with additional constraints added to control the sound field energy in the dark zone and global region. This is equivalent to combining ACC and mode matching for multi-zone reproduction. The difference is that instead of maximising the acoustic contrast between the bright and dark zone, the proposed method aims to maintain a certain level of sound energy within the dark zone.

The simulation results demonstrate that by choosing appropriate values of two Lagrange multipliers we can tune the bright zone and dark zone performances. The proposed method is not restricted to two zone reproduction as in the work[23] but applicable to sound control within higher number of zones.

In terms of the wideband design, we plot the coefficient of realisability in Fig. 6 for illustration. At low frequencies, sound zones become close together relative to the acoustic wavelength. As a result, the ATFs to the zones become highly correlated, and the maximum achievable performance drops. The exception is for the angle of occlusion ($\phi = 0°$), at high frequencies, sound propagates more in straight lines which exacerbates the occlusion problem. The computational complexity of the Newton's based algorithm also scales with frequency. This is due to the fact that at low frequencies the number of modes in the global region is small and at high frequencies it is large. The required Lagrange multipliers will vary with frequency as well. In practice these should not vary too fast with frequency, as they may produce severe transients in the filters of the loudspeaker
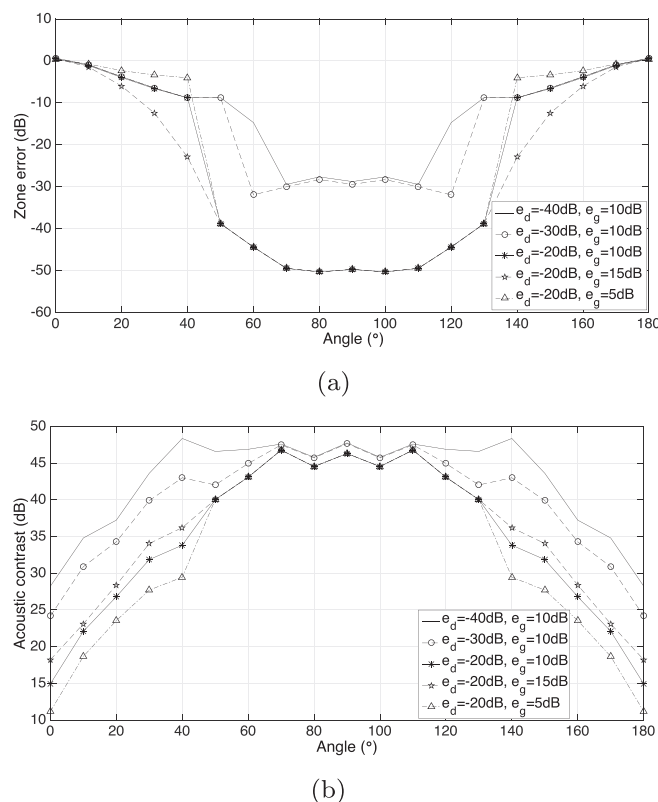

(a)

(b)

FIG. 5. Reproduction performance as a plane wave incident angle is panned with different sound energy limits. (a) and (b) are the metrics of the bright zone reproduction error and the acoustic contrast between two zones, respectively.
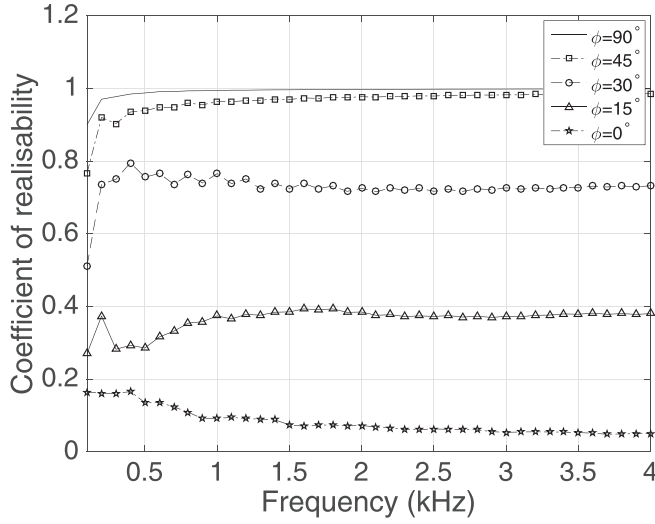
FIG. 6. Wideband system performance as indicated by the coefficient of realisability.

driving units. The global sound energy constraint, i.e., effort-based regularisation, has the effect of reducing the transients in the filters.

VI. CONCLUSION

This paper has provided theoretical insights into the multi-zone control problem, based on formulating the problem in the modal domain using cylindrical harmonic expansion of the sound fields and a translation matrix to relate the local and global sound fields. We reviewed the general framework of the multi-zone control on the sound field expansion coefficients using the Lagrange cost function that combines the principles of mode matching and ACC to recreate a desired sound field in the bright zone and to minimise, at the same time, the sound level in the dark zone and global region. The intrinsic ill-posedness of the multi-zone problem was analyzed using a Lagrangian-style solution. The coefficient of realisability was proposed as an indicator of the performance

limitation, such as the occlusion problem, with the knowledge of the bright-zone sound field coefficients and sound zone geometry. Simulation results showed that by choosing appropriate values of the Lagrange multipliers to satisfy the sound energy constraints in the dark zone and global region, the proposed method achieves a compromise between good bright zone performance and suitable dark zone performance. The efficient design of the loudspeaker filters with desirable time-domain properties is a topic of future work.

APPENDIX

The coefficients of the bright zone sound field and that of the global sound field are related by the following two equations:

$$\alpha_n^{(b)}(k) = \sum_{m=-N_0}^{N_0} J_{m-n}(kr_b)e^{i(m-n)\phi_b}\beta_m(k),$$
$$n = -N_b, \ldots, N_b, \tag{A1}$$

$$\beta_m(k) = \sum_{n=-N_b}^{N_b} J_{m-n}(kr_b)e^{-i(m-n)\phi_b}\alpha_n^{(b)}(k),$$
$$m = -N_0, \ldots, N_0. \tag{A2}$$

In matrix form, we have

$$\boldsymbol{A}_b = \boldsymbol{T}_b\boldsymbol{B}, \quad \boldsymbol{B} = \hat{\boldsymbol{T}}_b\boldsymbol{A}_b, \tag{A3}$$

where $\boldsymbol{A}_b = [\alpha_{-N_b}^{(b)}(k), \ldots, \alpha_{N_b}^{(b)}(k)]$ and $\boldsymbol{B} = [\beta_{-N_0}(k), \ldots, \beta_{N_0}(k)]$ are column vectors of bright zone sound field coefficients and global sound field coefficients, respectively. The translation matrices

$$\boldsymbol{T}_b = \begin{bmatrix} J_{-N_0+N_b}(kr_b)e^{i(-N_0+N_b)\phi_b} & \cdots & J_{N_0+N_b}(kr_b)e^{i(N_0+N_b)\phi_b} \\ \vdots & \ddots & \vdots \\ J_{-N_0-N_b}(kr_b)e^{i(-N_0-N_b)\phi_b} & \cdots & J_{N_0-N_b}(kr_b)e^{i(N_0-N_b)\phi_b} \end{bmatrix} \tag{A4}$$

and

$$\hat{\boldsymbol{T}}_b = \begin{bmatrix} J_{-N_0+N_b}(kr_b)e^{-i(-N_0+N_b)\phi_b} & \cdots & J_{-N_0-N_b}(kr_b)e^{-i(-N_0-N_b)\phi_b} \\ \vdots & \ddots & \vdots \\ J_{N_0+N_b}(kr_b)e^{-i(N_0+N_b)\phi_b} & \cdots & J_{N_0-N_b}(kr_b)e^{-i(N_0-N_b)\phi_b} \end{bmatrix}, \tag{A5}$$

are of size $\mathcal{N}_b \times \mathcal{N}_0$ and $\mathcal{N}_0 \times \mathcal{N}_b$, respectively, with $\mathcal{N}_b = 2N_b + 1$ and $\mathcal{N}_0 = 2N_0 + 1$.

We observe here that

$$\boldsymbol{T}_b = \hat{\boldsymbol{T}}_b^* \tag{A6}$$

and also

$$T_b \hat{T}_b = \begin{bmatrix} \varphi(-N_b, -N_b) & \cdots & \varphi(-N_b, N_b) \\ \vdots & \ddots & \vdots \\ \varphi(N_b, -N_b) & \cdots & \varphi(N_b, N_b) \end{bmatrix}, \qquad \text{(A7)}$$

where

$$\begin{aligned}
\varphi(n_1, n_2) &\triangleq \sum_{m=-N_0}^{N_0} J_{m-n_1}(kr_b) e^{i(m-n_1)\phi_b} \\
&\qquad \times J_{m-n_2}(kr_b) e^{-i(m-n_2)\phi_b} \\
&= \sum_{m=-N_0}^{N_0} J_{m-n1}(kr_b) J_{m-n_2}(kr_b) e^{i(n_2-n_1)\phi_b} \\
&= e^{i(n_2-n_1)\phi_b} \sum_{m'=-N_0-n_1}^{N_0-n_1} J_{m'}(kr_b) J_{m'-(n_2-n_1)}(kr_b),
\end{aligned}$$

$$\text{(A8)}$$

with $m' = m - n_1$. For a bright zone radius $r_b \le \min\ (|N_0 - n_1|, |N_0 + n_1|)/k$, Eq. (A8) can be approximated as

$$\begin{aligned}
\varphi(n_1, n_2) &\cong e^{i(n_2-n_1)\phi_b} \sum_{m'=-\infty}^{\infty} J_{m'}(kr_b) J_{m'-(n_2-n_1)}(kr_b) \\
&= e^{i(n_2-n_1)\phi_b} \sum_{m'=-\infty}^{\infty} J_{m'}(kr_b) J_{n_2-n_1-m'}(-kr_b) \\
&= e^{i(n_2-n_1)\phi_b} J_{n_2-n_1}(0) = \begin{cases} 0, & n_1 \ne n_2 \\ 1, & n_1 = n_2. \end{cases}
\end{aligned}$$

$$\text{(A9)}$$

The derivation is based on the addition formula for Bessel functions[37]

$$J_n(x+y) = \sum_{m=-\infty}^{\infty} J_m(x) J_{n-m}(y). \qquad \text{(A10)}$$

Therefore, provided the global region is big enough so that

$$r_b \le \frac{\min(|N_0 - n_1|, |N_0 + n_1|)}{k} = \frac{N_0 - N_b}{k} = r - R_b,$$

i.e., $r \ge r_b + R_b$, we have

$$T_b \hat{T}_b = T_b T_b^* = I. \qquad \text{(A11)}$$

This shows that T_b and T_b^* are one-sided inverse.

Given the following relations and based on Eq. (A11):

$$T_b T_b^* A_b = A_b = T_b B,$$

we have

$$T_b^* A_b = B.$$

[1] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," J. Acoust. Soc. Am. 125(4), 2091–2097 (2009).

[2] 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), 414–424 (2013).

[3] T. D. Abhayapala and Y. J. Wu, "Spatial soundfield reproduction with zones of quiet," in Proceedings of the 127th Audio Engineering Society Convention, New York, NY (October 2009), 7 pp.

[4] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial sound with loudspeakers and its perception: A review of the current state," Proc. IEEE 101(9), 1920–1938 (2013).

[5] H. Wierstorf, "Perceptual assessment of sound field synthesis," Ph.D. dissertation, Technical University of Berlin, 2014, pp. 37–54.

[6] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, "Personal sound zones: Delivering interface-free audio to multiple listeners," IEEE Signal Proc. Mag. 32(2), 81–91 (2015).

[7] 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).

[8] M. Shin, S. Q. Lee, F. M. Fazi, P. A. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces," J. Acoust. Soc. Am. 128, 121–131 (2010).

[9] S. J. Elliott, J. Cheer, J.-W. Choi, and Y.-H. Kim, "Robustness and regularization of personal audio systems," IEEE Trans. Audio, Speech, Lang. Process. 20(7), 2123–2133 (2012).

[10] P. Coleman, P. Jackson, M. Olik, and J. A. Pederson, "Personal audio with a planar bright zone," J. Acoust. Soc. Am. 136(4), 1725–1735 (2014).

[11] P. Coleman, P. Jackson, M. Olik, M. Møller, M. Olsen, and J. A. Pederson, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," J. Acoust. Soc. Am. 135(4), 1929–1940 (2014).

[12] M. A. Poletti, "An investigation of 2D multizone surround sound systems," in Proceedings of the 125th Audio Engineering Society Convention, San Francisco, CA (October 2008), 9 pp.

[13] T. Betlehem and C. Withers, "Sound field reproduction with energy constraint on loudspeaker weights," IEEE Trans. Audio, Speech, Lang. Process. 20(8), 2388–2392 (2012).

[14] N. Radmanesh and I. S. Burnett, "Generation of isolated wideband sound field using a combined two-stage Lasso-LS algorithm," IEEE Trans. Audio, Speech, Lang. Process. 21(2), 378–387 (2013).

[15] W. Jin and W. B. Kleijn, "Multizone sound field reproduction in reverberant rooms using compressed sensing techniques," in Proceedings of the IEEE ICASSP, Florence, Italy (May 2014), pp. 4728–4732.

[16] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," J. Acoust. Soc. Am. 131(6), 4518–4525 (2012).

[17] J.-H. Chang and F. Jacobsen, "Experimental validation of sound field control with a circular double-layer array of loudspeakers," J. Acoust. Soc. Am. 133, 2046–2054 (2013).

[18] W. Jin, W. B. Kleijn, and D. Virette, "Multizone soundfield reproduction using orthogonal basis expansion," in Proceedings of the IEEE ICASSP, Vancouver, Canada (May 2013), pp. 311–315.

[19] Y. Cai, M. Wu, and J. Yang, "Sound reproduction in personal audio systems using the least-squares approach with acoustic contrast control constraint," J. Acoust. Soc. Am. 135, 734–741 (2014).

[20] S. Galvez, F. Marcos, S. J. Elliott, and C. Jordan, "Time domain optimisation of filters used in a loudspeaker array for personal audio," IEEE/ACM Trans. Audio, Speech, Lang. Process. 23(11), 1869–1878 (2015).

[21] Y. J. Wu and T. D. Abhayapala, "Multizone 2D soundfield reproduction via spatial band stop filters," in Proceedings of the IEEE WASPAA, New Paltz, NY (October 2009), pp. 309–312.

[22] 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).

[23] M. A. Poletti and F. M. Fazi, "An approach to generating two zones of silence with application to personal sound systems," J. Acoust. Soc. Am. 137(2), 598–605 (2015).

[24] W. Jin and W. B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," IEEE/ACM Trans. Audio, Speech, Lang. Process. 23(12), 2343–2355 (2015).

[25] D. Menzies, "Sound field synthesis with distributed modal constraints," Acta Acust. Acust. 98(1), 15–27 (2012).

[26] K. Helwani, S. Spors, and H. Buchner, "The synthesis of sound figures," Multidimens. Syst. Signal Process. 25(2), 379–403 (2014).

[27] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," IEEE Trans. Signal Process. 55(6), 2542–2556 (2007).

J. Acoust. Soc. Am. 140 (3), September 2016

Zhang et al.   2143

[28]G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed. (Cambridge University Press, New York, 1995), Chap. 7, 198 pp.

[29]S. J. Elliott, J. Cheer, H. Murfet, and K. R. Holland, "Minimally radiating sources for personal audio," J. Acoust. Soc. Am. **128**(4), 1721–1728 (2010).

[30]T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," J. Acoust. Soc. Am. **117**(4), 2100–2111 (2005).

[31]T. Betlehem and P. D. Teal, "A constrained optimization approach for multi-zone surround sound," in *Proceedings of the IEEE ICASSP*, Prague, Czech Republic (May 2011), pp. 437–440.

[32]N. R. Nassif and D. K. Fayyad, *Introduction to Numerical Analysis and Scientific Computing*, 1st ed. (CRC Press, Boca Raton, FL, 2013), 63 pp.

[33]M. A. Poletti, T. Betlehem, and T. D. Abhayapala, "Higher-order loudspeakers and active compensation for improved 2D sound field reproduction in rooms," J. Audio Eng. Soc. **63**, 31–45 (2015).

[34]M. A. Poletti, "Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts," J. Audio Eng. Soc. **55**, 598–610 (2007).

[35]J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**(4), 943–950 (1979).

[36]K. Baykaner, P. Coleman, R. Mason, P. J. B. Jackson, J. Francombe, M. Olik, and S. Bech, "The relationship between target quality and interference in sound zone," J. Audio Eng. Soc. **63**, 78–89 (2015).

[37]D. R. Lide, *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data*, 1st ed. (CRC Press LLC, Boca Raton, FL, 2003), A-95 pp.