Spherical Harmonics Based Generalized Image Source Method for Simulating Room Acoustics

Prasanga N. Samarasinghe,[a)] Thushara D. Abhayapala, Yan Lu, and Hanchi Chen

Audio and Acoustic Signal Processing Group, Research School of Engineering, Australian National University, Canberra, Australia

Glenn Dickins

Dolby Laboratories, Sydney, Australia

Allen and Berkley's image source method is proven to be a very useful and popular technique for simulating the acoustic room transfer function (RTF) in reverberant rooms. It is based on the assumption that the source and receiver of interest are both omnidirectional. With the inherent directional nature of practical loudspeakers, and the increasing use of directional microphones, the above assumption is often invalid. The main objective of this paper is to generalize the frequency domain image source method in the spherical harmonics domain, such that it could simulate the RTF between practical transducers with higher-order directivity. We represent the transducer directivity patterns in terms of spherical harmonic functions and utilize the concept of image sources on spherical harmonic based propagation patterns to formulate the generalized image source method. From now on, any transducer of interest, can be modeled in the spherical harmonics domain with a realistic directivity pattern and incorporated with the proposed method to simulate room acoustics more accurately. The proposed generalization thus reconciles image source method the with the spatial soundfield theory. It also has an alternate use case of enabling RTF simulations for moving point-transducers inside pre-defined source and receiver regions.

[a)]prasanga.samarasinghe@anu.edu.au

I. Introduction

Sound propagation characteristics in reverberant environments is an important topic of research. This is due to its impact on a plethora of applications in audio signal processing. Some well known techniques for simulating and understanding room acoustics include ray/beam tracing[1–5], boundary and finite element methods[6], digital waveguide meshes[7,8], spatial sound decomposition based methods[9–12] and the well known image source method[13]. Despite the abundance of sophisticated room-acoustics simulation methods available, the relatively basic image source method proposed by Allen and Berkley[13] still remains a sought-after technique for simulating the room transfer function (RTF) and its time domain counterpart, the room impulse response (RIR).

The image source method is often utilized by researchers and engineers to simulate room characteristics for applications such as soundfield analysis and synthesis[9,14–16], generating stimuli for perceptual and psychoacoustic tests[17,18], validating algorithms or systems designed to operate in reverberant conditions[19], sound rendering and auralization in virtual auditory systems[20,21], design of acoustic spaces[22] and commercial audio device testing. The image source method is also continuously being improved to increase its efficiency and effectiveness[23–27]. The prominent nature of the image source method can be attributed to a number of strengths compared to other methods. As discussed in[28], these include (i) simplicity of algorithmic implementation; (ii) high degree of flexibility, with many simulation parameters (such as room dimensions, acoustic absorption coefficients, source and microphone positions, reverberation time) adjustable in software; and (iii) the ability to generate good approximations for realistic room impulse responses.

Inherently, the image source method simulates the room response between a point source and a point receiver with omnidirectional directivity. However in practice, acoustic transducers (speakers and mirophones) are directional due to two reasons; (i) It's impossible to realize omnidirectional or point transducers due to physical limitations and size, and (ii) With the recent advancements in design and implementation of higher-order transducers[29–31], there is an increasing interest in using transducers with pre-determined directional patterns to record/produce spatial soundfields. The application of the original image source method to emulate realistic acoustic scenarios thus introduces error as the practical transducers violate the assumption of being omnidirectional. Extension of image source method for first-order microphones has been proposed in[32] in the time domain.

In this paper, we aim to extend image source method in the spherical harmonics domain, such that it can simulate the frequency domain room response or RTF for higher-order (or directional) transducers, both sources and microphones. We first decompose the soundfield emitted/recorded by the directional transducers in terms of spherical harmonic functions. Then, the basic concept of the original image source method is utilized to derive the acoustic images for spherical harmonic shaped source emissions. These are then used to formulate the room induced coupling between the directional source and the directional receiver. Finally, the coupling coefficients are employed to derive the generalized image source method directional transducers. It is important to note that this paper is not an alternate image source method, but an expansion to the existing image source method in the spherical harmonics domain, such that it complies with directional transducers. Therefore, image source method's inherent drawbacks such as its restriction to rectangular rooms, inability to model diffraction and the presence of audible artifacts will naturally be present in the proposed generalization.

The remainder of the paper is organized as follows. Section II summarizes the original image source method while discussing the basic concept of acoustic mirroring from walls. Section III presents the formulation of problem in the spherical harmonics domain. Section IV discusses the image source concept for directional sources with known directivity patterns, followed by section V, which derives the relationship between reflected directional sources and a directional receiver. Section VI combines the aforementioned derivations to formulate the generalized image source method. Finally, section VII presents simulation results to verify the accuracy of the proposed generalization. It also briefly presents a practical application of the proposed method.

II. Summary of the Image Source Method

The image source method was originally presented to model the point-to-point RTF in rectangular enclosures, such that when multiplied with any desired input signal (in the frequency domain), simulates the room response as observed at the receiver point. This section provides a brief background review of the image source method.

Consider a shoebox room (A "Shoebox room" is a practitioning term for a typical rectangular room) with dimensions (L_x, L_y, L_z) for length, width and height, respectively. Assume a Cartesian coordinate system is defined inside this enclosure, where the origin coincides

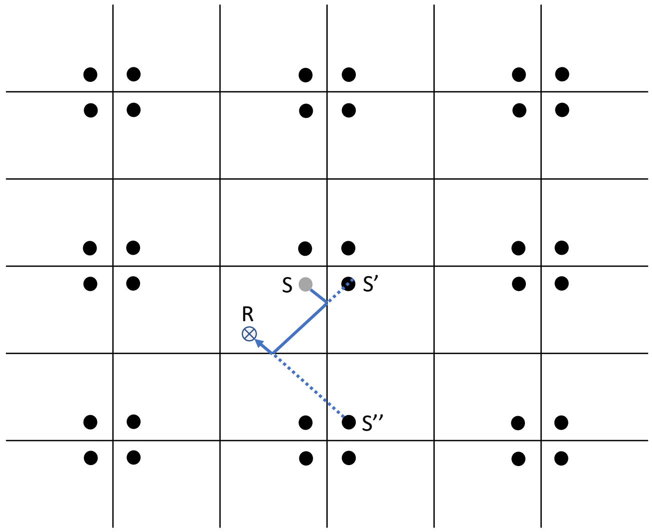


FIG. 1. (Colour online) Concept of image sources where walls are considered as mirrors.

with one of the corners of the room. Let a point source be positioned at $\boldsymbol{x}_s = (x_s, y_s, z_s)$, and a point receiver be positioned at $\boldsymbol{x}_r = (x_r, y_r, z_r)$. The direct path received at $\boldsymbol{x}_r$ is then given by the Green's function,

$$P_d(k, \boldsymbol{x}_s, \boldsymbol{x}_r) = \frac{e^{ik|\boldsymbol{x}_s - \boldsymbol{x}_r|}}{4\pi|\boldsymbol{x}_s - \boldsymbol{x}_r|} \tag{1}$$

where $k = 2\pi f/c$ with f and c representing the frequency in Hz and sound of speed in ms^{-1}, respectively. The formulation of the image source method is based on geometric room-acoustic principles. It is assumed that that the reflections characteristics of each wall can be defined in terms of a sound reflection coefficient γ, which relates to the absorption coefficient ψ through

$$\psi = 1 - \gamma^2. \tag{2}$$

In the original image source formulation[13], the reflection coefficients are assumed to be independent of both (i) sound wave incident angle, and (ii) frequency. As showed in Fig.1, the RTF from the source to the receiver can be determined by considering image sources on an infinite grid of mirror rooms expanding in all three dimensions. Note that in real-world applications this grid can be truncated to an order enclosing a sufficient number of image sources to represent the given room's inherent reverberant characteristics. The contribution from each image source to the receiver signal is a replica of the original source signal, attenuated by a certain amplitude factor and phase shifted by a certain angle. The RTF hence follows as

$$P(k,\boldsymbol{x}_s,\boldsymbol{x}_r)=\sum_{\boldsymbol{p}=0}^{1}\sum_{\boldsymbol{r}=-\infty}^{\infty}\gamma_{x1}^{|a-q|}\gamma_{x2}^{|a|}\gamma_{y1}^{|b-j|}\gamma_{y2}^{|b|}\gamma_{z1}^{|c-\ell|}\gamma_{z2}^{|c|}\frac{e^{ik|\boldsymbol{R_p}+\boldsymbol{R_r}|}}{4\pi|\boldsymbol{R_p}+\boldsymbol{R_r}|} \tag{3}$$

where $\boldsymbol{p}=(p_1,p_2,p_3)$ and $\boldsymbol{r}=(r_1,r_2,r_3)$ are triplet parameters controlling the indexing of the image sources in all dimensions, $\boldsymbol{R_p}=(x_r-x_s+2p_1x_s,y_r-y_s+2p_2y_s,z_r-z_s+2p_3z_s)$, and $\boldsymbol{R_r}=(2r_1L_x,2r_2L_y,2r_3L_z)$, $\gamma_{x,i},\gamma_{y,i},\gamma_{z,i}$, with $i=1,2$, are wall reflection coefficients where $i=1$ refers to the wall closest to the room origin and $i=2$ refers to walls on the opposite sides. The room origin is assumed to be at $x=y=z=0$. Note that the sum $\sum_{\boldsymbol{p}=0}^{1}$ indicates three sums, for each of the three components of $\boldsymbol{p}=(p_1,p_2,p_3)$, and similarly, the sum $\sum_{\boldsymbol{r}=-\infty}^{\infty}$ indicates three sums over $\boldsymbol{r}=(r_1,r_2,r_3)$. Physically these sums are over a 3-D lattice of image points, where $\boldsymbol{p}$ involves an eight point lattice, and $\boldsymbol{r}$ involves an infinite lattice, which can be truncated at the reflection order R. Note that this order largely depends on the room's inherent characteristics including, room size, shape and boundary materials.

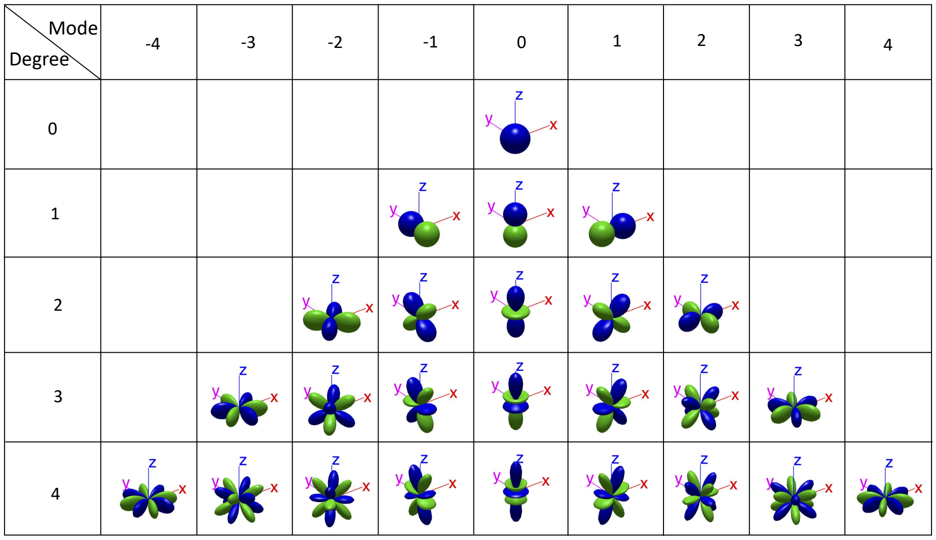


FIG. 2. (Color online) Illustration of spherical harmonic functions $Y_{nm}(\cdot)$ with different brightnesses representing the phase relationships.

With increasing order of reflections $\boldsymbol{r}$, the number of image sources included in (3) increases cubically. Therefore, even if one claims it is technically possible to represent any directional source/receiver in terms of a weighted sum of point sources/receivers, the respective calculation of the multiple RTFs can lead to a significant computational load in practice causing loss of simplicity and elegance.

III. Problem Formulation

In this section, we formulate the problem at hand in the spherical harmonics domain. The spherical harmonics (Fig.2) are a set of orthogonal spatial basis functions that can be utilized to decompose any arbitrary function defined on the sphere.

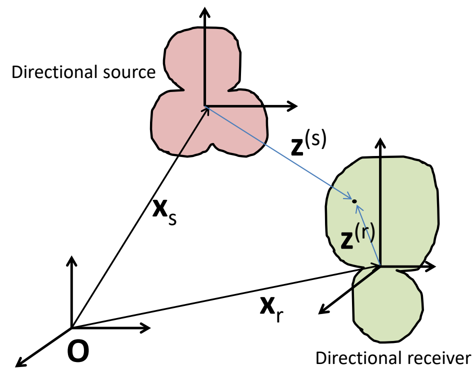


FIG. 3. Illustration of a source and receiver with directional characteristics

A. Spherical harmonics based representation of directional transducers

Here, we illustrate a realistic scenario where the source and receiver are directional. As shown in Fig.3, let there be a directional source at $\boldsymbol{x}_s$, and a directional receiver at $\boldsymbol{x}_r$. When observed on a sphere, the outgoing soundfield from the source with respect to $\boldsymbol{x}_s$ and the resulting room response arriving at the receiver with respect to $\boldsymbol{x}_r$ can both be expressed in terms of independent spherical harmonic decompositions as follows.

1. Spherical harmonics representation of the outgoing soundfield from a directional source

Consider a homogeneous outgoing soundfield from the source at $\boldsymbol{x}_s$. When observed at any arbitrary location with spherical coordinates $\boldsymbol{z}^{(s)} = (z^{(s)}, \theta_z^{(s)}, \phi_z^{(s)})$ with respect to $\boldsymbol{x}_s$, this outgoing soundfield can be represented using a spherical harmonic decomposition of the form[33]

$$S_{\text{out}}(\boldsymbol{z^{(s)}}, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \beta_{nm}(k) h_n(kz^{(s)}) Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)}) \tag{4}$$

where $h_n(\cdot)$ denotes the spherical Hankel function of first kind for order n, $Y_{nm}(\cdot)$ denotes the spherical harmonic function of order n and mode m, defined by[33]

$$Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)}) = P_{n|m|}(\cos(\theta_z^{(s)})) \frac{1}{\sqrt{2\pi}} e^{im\phi_z^{(s)}} \tag{5}$$

where $P_{n|m|}(\cos(\theta_z^{(s)})) \triangleq \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_{n|m|}(\cos(\theta_z^{(s)}))$ is the normalized associated Legendre polynomial with $P_{n|m|}(\cos(\theta_z^{(s)}))$ being the associated Legendre polynomials. The coefficients $\beta_{nm}(k)$ of (4) denote the respective spherical harmonic weighting for the order n and mode m, which in this case represents the directional characteristics of the loudspeaker. Note that depending on the source directivity pattern, the infinite summation in (4) can be truncated at order N.

2. Spherical harmonics representation of incident soundfield at the directional receiver

Consider a homogeneous soundfield incident at the directional receiver at $\boldsymbol{x}_r$. This soundfield when observed at any arbitrary location with spherical coordinates $\boldsymbol{z^{(r)}} = (z^{(r)}, \theta_z^{(r)}, \phi_z^{(r)})$ with respect to $\boldsymbol{x}_r$, can be represented in terms of a spherical harmonic decmposition of the form[33]

$$S(\boldsymbol{z^{(r)}}, k) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \alpha_{vu}(k) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}) \tag{6}$$

where $j_v(\cdot)$ denotes the spherical Bessel function of order v, V is the respective truncation limit determined by $V = \lceil kz^{(r)} \rceil$ due to the presence of spherical Bessel functions[34]. A V^{th} order microphone located at $\boldsymbol{x}_r$ would be capable of successfully extracting the soundfield components $\alpha_{vu}(k)$ for $v = 0 : V$ and $u = -v : v$ with respect to its local origin[29]. If the higher-order microphone has beamforming capability (i.e., similar to the directional receiver shown Fig.3), then each recorded soundfield coefficient will be scaled as $\alpha_{vu}(k) \times \delta_{vu}(k)$, where $\delta_{vu}(k)$ are the beamformer coefficients or the harmonic domain coefficients of the beampattern when described using spherical harmonic decomposition similar to (6).

B. Summary of the problem

Note that for a given loudspeaker, its order can be determined by $N = \lceil k\hat{R} \rceil$, where $\hat{R}$ is the radius of the smallest sphere enclosing the physical speaker[35]. We assume the order N and outgoing soundfield coefficients $\beta_{nm}(k)$ are known for the loudspeaker of interest. Based on spatial soundfield theory, the spherical harmonic coefficients beyond this order can be assumed to be negligible. We also assume that the order V of the directional microphone is known, and it's capable of recording all the soundfield coefficients up to order V. If the directional microphone has beamforming capabilities, then the corresponding beamformer coefficients $\delta_{nm}(k)$ are also assumed to be known. The objective of this paper is to generalize the image source method to directional transducers. For this purpose, it is required to (i) apply the image source concept to directional sources and (ii) parameterize the room response between directional transducers in terms of a single closed form equation. In the remainder of this paper, we address these problems one by one, and formulate a generalized image source method for a rectangular (or shoebox) room.

IV. Acoustic Image of a directional source

In this section, we extend the image source concept to directional sources, whose outgoing soundfield can be decomposed in terms of spherical harmonics (4).

By definition, the image source method for point sources repetitively place each image of the original source on the far side of the respective wall. As expressed in (4), the outgoing soundfield from a directional source as observed at a point $\boldsymbol{z^{(s)}}$ can be decomposed in terms of spherical harmonics where each unit amplitude outgoing mode is of the form form $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$. Intuitively, extending the image source concept to each unit amplitude outgoing pattern of the above form seems straightforward. However, this is not a simple task because when performing the reflection operation along a particular wall, the positive direction of the Cartesian axes local to the directional source effectively rotates. As shown in Figure 4, this problem will not pose negative influence on point sources (or the zero$^{\text{th}}$ order source pattern $h_0(kz^{(s)})Y_{00}(\theta_z^{(s)}, \phi_z^{(s)})$) as their outgoing field is rotationally invariant. However, for all other spherical harmonic excitation patterns $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ when $n > 0$, the outgoing field gets mirrored due to the intrinsic shape of spherical harmonic functions. Thus, the reflected image (see Fig.5 for an example) has to be carefully modeled for all spherical harmonic domain excitation patterns.

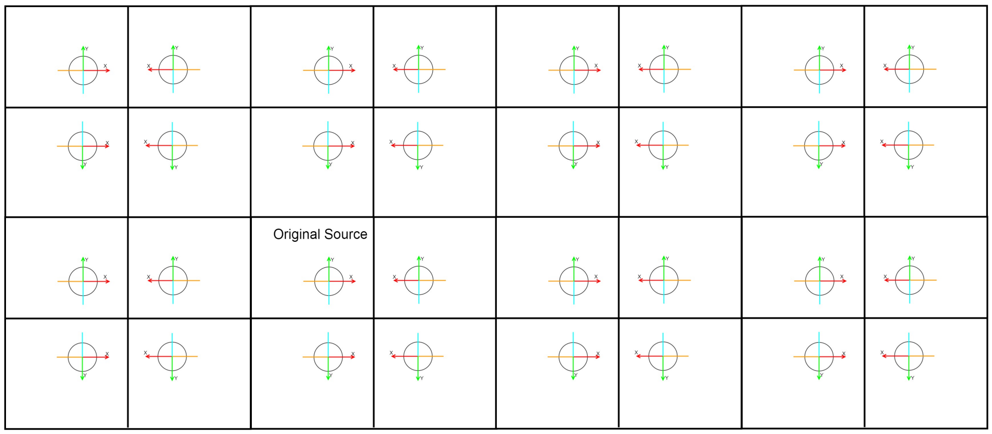


FIG. 4. (Color online) Reflection from the X-Z and Y-Z planes for an omnidirectional source

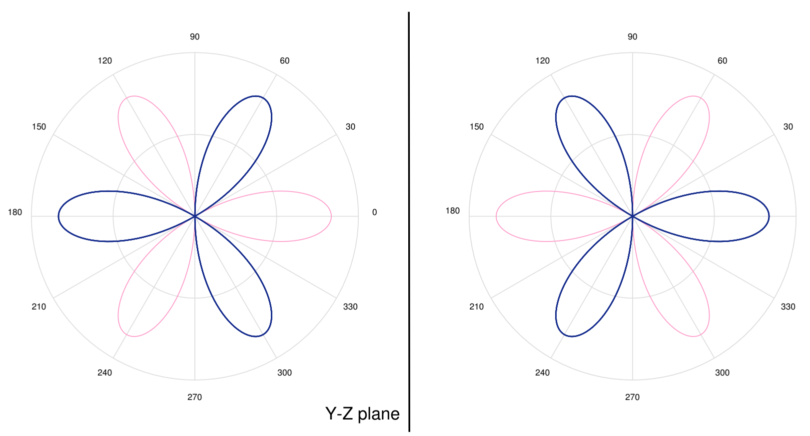


FIG. 5. (Color online) Reflection from Y-Z plane for an outgoing mode of the form $Y_{33}(\theta_z^{(s)}, \phi_z^{(s)})$. The two brightnesses represent phase relationships.

A. Acoustic image of a spherical harmonic based excitation pattern

Let us consider a unit amplitude outgoing mode of order n and mode m from the directional source. As shown in (4), each unit amplitude outgoing mode carries two functions $h_n(kz^{(s)})$ and $Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$, where $h_n(kz^{(s)})$ is not affected by the mirrored axes due to its independence of the angles θ and ϕ. For the term $Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$, it is required to incorporate an appropriate mirror operation to offset the influence from the change of axis positive direction.

Let us discuss the effect on $Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ as the original source is reflected from the Cartesian planes $X-Z$, $Y-Z$ and $X-Y$ adjacent to the room origin. As showed in Fig.4, when a directional source is reflected from an $X-Z$ plane, the azimuth angle with respect to $\boldsymbol{x}_s$ experiences a rotational shift of $\phi_{\text{rotate}} = -\phi_{\text{original}}$. Similarly, for the $X-Z$ plane, the azimuth angle experiences a rotational shift of $\phi_{\text{rotate}} = \pi - \phi_{\text{original}}$, and for the $X-Y$ plane, the elevation angle experiences a rotational shift of $\theta_{\text{rotate}} = -\theta_{\text{original}}$. These effects can be incorporated in the spherical harmonic excitation pattern ($Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$) of a directional source to summarize its adjacent image sources as follows. Note that we utilize the rotational properties of spherical harmonics[36,37] to perform an extra simplification step where the rotations on azimuth and elevation angles are transferred to degree n and mode m.

The adjacent image source reflected from the X-Z plane will emit spherical harmonic patterns of the form

$$Y_{nm}(\theta_z^{(s)}, -\phi_z^{(s)}) = (-1)^m Y_{n,-m}(\theta_z^{(s)}, \phi_z^{(s)}). \tag{7}$$

The adjacent image source reflected from the Y-Z plane will emit spherical harmonic patterns of the form

$$Y_{nm}(\theta_z^{(s)}, \pi - \phi_z^{(s)}) = Y_{n,-m}(\theta_z^{(s)}, \phi_z^{(s)}). \tag{8}$$

The adjacent image source reflected from the X-Y plane will emit spherical harmonic patterns of the form

$$Y_{nm}(-\theta_z^{(s)}, \phi_z^{(s)}) = (-1)^{n+m} Y_{n,m}(\theta_z^{(s)}, \phi_z^{(s)}). \tag{9}$$

The above results are summarized below in Table I.

Cartesian Plane	Image
X-Z	$(-1)^m Y_{n,-m}(\theta_z^{(s)}, \phi_z^{(s)})$
Y-Z	$Y_{n,-m}(\theta_z^{(s)}, \phi_z^{(s)})$
X-Y	$(-1)^{n+m} Y_{n,m}(\theta_z^{(s)}, \phi_z^{(s)})$

TABLE I. Acoustic image of the spatial excitation pattern $Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ when mirrored from Cartesian planes.

Above results depict the reflection operation related to each plane adjacent to the room origin. Similar operations can be carried out to all first order and higher order images.

Therefore, analogous to the image source method for a point source (3), the room response for a unit amplitude source excitation pattern of the form $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ originated at $\boldsymbol{x}_s$, as observed at the receiver origin $\boldsymbol{x}_r$ is

$$\begin{aligned} P_{nm}(k, \boldsymbol{x}_s, \boldsymbol{x}_r) = &\sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|} (-1)^{(j+\ell)m+\ell n} h_n(k|\boldsymbol{R_p} + \boldsymbol{R_r}|) \\ &Y_{n,((-1)^{q+j}m)}(\theta_{\vec{z}}^{(s)}, \phi_{\vec{z}}^{(s)}). \end{aligned} \tag{10}$$

Note that the above expression is an important result as it can be defined as the image source method between a single-mode source (i.e., emitting $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$) and a point receiver. This will act as the basic building block of the proposed generalized image source method. Also note that when $n = 0$ and $m = 0$, (10) simplifies to the original image source method (3).

V. Coupling between a directional source and a directional receiver

A. Room response as observed by a directional receiver

Here, we look at the directionality of the receiver in more detail. As described earlier, a V^{th} order incident soundfield can be expressed by (6), and a V^{th} order microphone is capable of recording all the corresponding soundfield coefficients. These microphone recordings enable the prediction of sound at any arbitrary location $\boldsymbol{z}^{(\boldsymbol{r})}$ away from its local origin $\boldsymbol{x}_r$ given $\lceil k\boldsymbol{x}_r \rceil \leq V$.

Let's consider the incident spatial soundfield at a V^{th} order microphone due to a unit amplitude outgoing mode $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ from the source position $\boldsymbol{x}_s$. We express the soundfield observed at $\boldsymbol{z}^{(\boldsymbol{r})}$, a point away from the microphone origin, in terms a spherical harmonic decomposition similar to (6) as

$$P_{nm}(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \alpha_{vu}^{nm}(k) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}) \tag{11}$$

where $\alpha_{vu}^{nm}(k)$ denotes the v^{th} order, u^{th} mode soundfield coefficient of the room response incident at the reciver caused by a unit amplitude n^{th} order and m^{th} mode outgoing soundfield from the source. From now on, we refer to $\alpha_{vu}^{nm}(k)$ as the *mode coupling coefficients* as they represent the coupling between the outgoing modes from the directional source and the incident modes at the directional receiver for the room enclosure of interest.

B. Spherical harmonic domain mode coupling between a directional source and receiver

Section IV A describes the room response with respect to the source origin whereas Section V A describes the room response with respect to the receiver origin. In this section, we compare both expressions, and derive a closed form expression for the mode coupling parameters $\alpha_{vu}^{nm}(k)$.

Note that in Section IV A we derived the room response at the receiver origin $\boldsymbol{x}_r$ not at $\boldsymbol{z}^{(\boldsymbol{r})}$, a point away from $\boldsymbol{x}_r$. For direct comparison with the results of Section V A, this expression can be slightly modified to observe the soundfield incident at $\boldsymbol{z}^{(\boldsymbol{r})}$. That is, the image source method for a unit amplitude spherical harmonic excitation pattern of the form $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ as observed at the receiver location $\boldsymbol{z}^{(\boldsymbol{r})}$ is

$$P_{nm}(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|} (-1)^{(j+\ell)m+\ell n} h_n(k|\boldsymbol{R_p} + \boldsymbol{R_r} + \boldsymbol{z}^{(\boldsymbol{r})}|)$$
$$Y_{n,((-1)^{q+j}m)}(\theta_{\tilde{z}}^{(s)}, \phi_{\tilde{z}}^{(s)}) \tag{12}$$

Now (12) and (11) both express the soundfield at $\boldsymbol{z}^{(\boldsymbol{r})}$ due to a unit amplitude outgoing mode $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ from $\boldsymbol{x}_s$. Equation (12) expresses it in terms of a collection of mirrored outgoing modes of order n and m with respect to their respective image source origins, where as equation (11) expresses it in terms of an incident soundfield as observed by a V^{th} order microphone. We directly compare them to derive the image source method based mode coupling coefficients and introduce the below theorem.

Theorem 1 *Given an N^{th} order source and a V^{th} order receiver inside a shoe-box room, the spherical harmonic domain mode coupling between them based on the concept of image sources is*

$$\alpha_{vu}^{nm}(k) = \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \gamma_{x1}^{|a-q|}\gamma_{x2}^{|a|}\gamma_{y1}^{|b-j|}\gamma_{y2}^{|b|}\gamma_{z1}^{|c-\ell|}\gamma_{z2}^{|c|}(-1)^{(j+\ell)m+\ell n} S_{nv}^{((-1)^{q+j}m)\mu}(\boldsymbol{R_p}+\boldsymbol{R_r}) \tag{13}$$

where

$$S_{nv}^{m\mu}(\boldsymbol{x_o}) = 4\pi i^{v-n} \sum_{l=0}^{n+v} i^l (-1)^{2m-\mu} h_l(k|\boldsymbol{x_0}|) Y_{l(\mu-m)}^{*}(\theta_{x0},\phi_{x0}) W_1 W_2 \xi \tag{14}$$

with

$$W_1 = \begin{pmatrix} n & v & l \\ 0 & 0 & 0 \end{pmatrix} \quad \textit{and} \quad W_2 = \begin{pmatrix} n & v & l \\ m & -\mu & (\mu-m) \end{pmatrix} \tag{15}$$

denoting Wigner $3-j$ symbols[38] *and $\xi = \sqrt{(2n+1)(2v+1)(2l+1)/4\pi}$.*

Please refer to the appendix for a detailed proof of the above theorem. From (13) it is clear that for a given enclosure, the mode coupling relationship between an n^{th} order, m^{th} mode outgoing soundfield and a v^{th} order, u^{th} mode incoming soundfield only depends on the source/receiver local origin and the room characteristics (wall reflections, room dimensions etc.). This is an important result, because it can be incorporated with any arbitrary directional transducer when expressed in terms of spherical harmonics. In the following section, we use (13) to derive a generalized image source method between arbitrary directional transducers.

VI. The generalized image source method

Here, we derive a closed form expression for the generalized image source method for a directional source emitting multiple soundfield modes (4) and a directional receiver recording multiple soundfield modes (6). An N^{th} order source emits multiple soundfield modes of the form $h_n(kz^{(s)})Y_{nm}(\theta_z^{(s)}, \phi_z^{(s)})$ scaled with respective modal weights $\beta_{nm}(k)$(4). In this case, the total RTF as observed at the directional receiver is

$$P(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \sum_{n=0}^{N} \sum_{m=-n}^{n} \beta_{nm}(k) \alpha_{vu}^{nm}(k) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}). \tag{16}$$

By substituting (13) in (16), we derive the generalized image source method for directional sources and receivers as

$$P(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \beta_{nm}(k)(-1)^{(j+\ell)m+\ell n} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|}$$
$$S_{nv}^{((-1)^{q+j}m)\mu}(\boldsymbol{R_p} + \boldsymbol{R_r}) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}). \tag{17}$$

If the directional receiver has beamforming capabilities with beamformer coefficients $\delta_{vu}(k)$, the generalized image source may be slightly modified as

$$P(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \beta_{nm}(k)\delta_{vu}(k)(-1)^{(j+\ell)m+\ell n} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|}$$
$$S_{nv}^{((-1)^{q+j}m)\mu}(\boldsymbol{R_p} + \boldsymbol{R_r}) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}). \tag{18}$$

Let us summarize the significance of the above result.

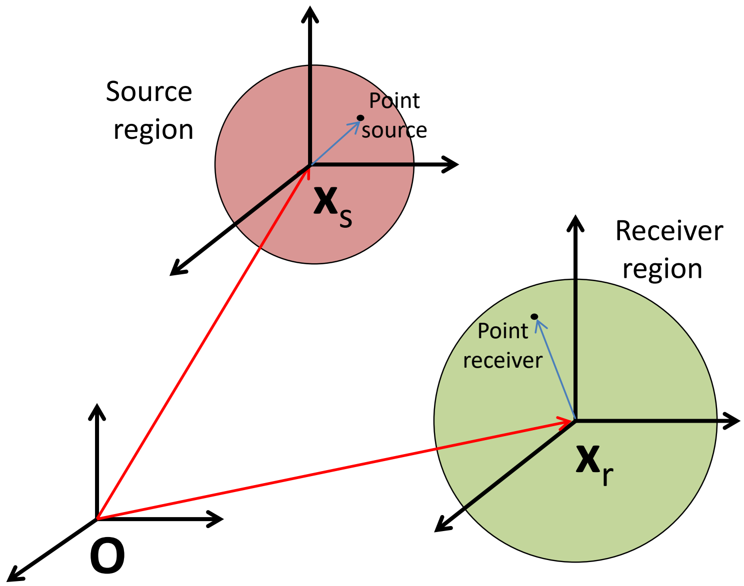


FIG. 6. (Color online) Geometrical representation of the concept of region-region RTF where a point source and a point receiver is assumed to be arbitrarily placed inside a pre-defined source region and a pre-defined receiver region

- For a given source of order N and directivity $\beta_{nm}(k)$, and a given microphone of order V, the room response can be simulated using (17). If the microphone has beamforming capability, the corresponding room response can be simulated using (18)
- When $N = 0$ and $V = 0$, the source and receiver represents ideal point transducers, thus (17) simplifies to the original image source method.
- Due to the rotational properties of spherical harmonics, the proposed model can also be applied to simulate the RTF between rotating directional transducers.

Alternate use of the proposed model; Region to region of RTF concept for omnidirectional transducers

Apart from the application to directional transducers, the generalized image source method has another important use case. That is, if one still assumes omnidirectional transducers, the proposed model can be useful in the sense of a region-to-region image source model (see Fig. 6). This is possible because the spatial soundfield due to a directional transducer at the origin (Fig. 3) and an omnidirectional point transducer away from the origin (Fig. 6) can both be represented using a similar higher order spehrical harmonic representation. Therefore if we define (i) a spherical source region centered at $\boldsymbol{x}_s$ enclosing an arbitrarily located point source(s), and (ii) a non-overlapping spherical receiver region centered at $\boldsymbol{x}_r$ enclosing an arbitrarily located point receiver(s), then each region will have a higher order directivity pattern with respect to their local origin (Fig. 6). Each region can be defined based on the practical application and the order of the soundfield can be determined based on the size and maximum frequency of request. Soundfield inside each region can be modeled in the spherical harmonics domain similar to equations (4) and (6), where the respective soundfield order is given by the radius of the interested region, and the respective soundfield coefficients can be derived based on the point source/receiver position. Once the outgoing source soundfield and the incoming receiver soundfield are modeled in the spherical harmonics domain, the proposed RTF model (17) is directly applicable. Note that the mode coupling parameters $\alpha_{vu}^{nm}(k)$ now describe the coupling between the source region and the receiver region. A useful advantage of this method is the simulation of RTF for moving transducers with omnidirectional characteristics.

In[11], the authors proposed a point-to-point (omnidirectional) RTF parameterization between two regions. While the work in[11] requires actual room measurements to find the coupling coefficients $\alpha_{vu}^{nm}(k)$, (13) can now be used to fully simulate the RTF between two point transducers, which can be arbitrarily moved inside a pre-defined source region and a receiver region.

VII. Simulation Results

In this section, we illustrate the accuracy of the proposed image source method for directional sources and receivers. We consider a shoe-box room of size $5 \times 3.5 \times 4$ m with its front-left-bottom corner defined as the origin. The room is assumed to have wall reflection coefficients $\gamma = [0.75, 0.65, 0.8, 0.2, 0.45, 0.7]$. The source is located at Cartesian coordinates $\boldsymbol{x}_s = (1, 1, 1)$, whereas the receiver is located at $\boldsymbol{x}_r = (1, 3, 3)$. We consider two cases of directional sources at $\boldsymbol{x}_s$ and derive the room response over a spherical receiver region of radius 0.25 m. For a given wavenumber k the soundfield order of the receiver region can be derived using $V = \lceil k \times 0.25 \rceil$. If the RTF is to be determined for a given directional microphone with order V, the radius of spatial area recorded by the microphone is $R_r = V/k$. In order to present a fair comparison with the original image source method and the proposed method, we make sure that the directional source of interest is capable of being represented by a combination of one or more point-sources distributed around the origin $\boldsymbol{x}_s$. Note that this is not a constraint to use the proposed method, which is applicable to any arbitrary directional source of the form (4).

We first consider directional source that emits a dipole outgoing soundfield, which resembles a spherical harmonic based outgoing pattern of mode $n = 1$ and order $m = 0$, $Y_{10}(\cdot)$. Such a radiation pattern can be obtained by two point sources along the z-axis spaced approximately at half wavelength of the target frequency. Considering our target frequency to be 2000 Hz, we use two unit amplitude sources at $(1, 1, 1.085)$ and $(1, 1, 0.915)$ to create the desired source. Utilizing the spherical harmonic decomposition of the Green's function[33], the outgoing modal coefficients $\beta_{nm}(k)$ of (4) caused by the above pair can be derived using

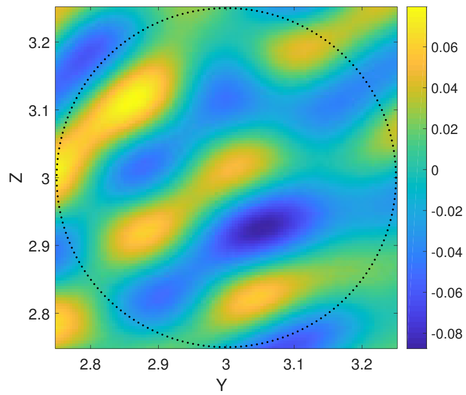
$$\beta^{(s)}_{nm}(k) = ik \sum_{d=1}^{2} w_d(k) j_n(kr_d) Y^*_{nm}(\theta^{(s)}_d, \phi^{(s)}_d) \tag{19}$$

where $w_d(k)$ is the point source weighting set at unity, $(r_1, \theta^{(s)}_1, \phi^{(s)}_1) = (0.085, 0, 0)$ and $(r_2, \theta^{(s)}_2, \phi^{(s)}_2) = (0.085, \pi, 0)$. For the given point source pair, it can be shown that $\beta^{(s)}_{nm}(k)$ is zero for all cases except for when $n = 1, m = 0$. This confirms that the directional source emits a soundfield with polar pattern $Y_{10}(\cdot)$ scaled by $\beta^{(s)}_{10}(k)$.

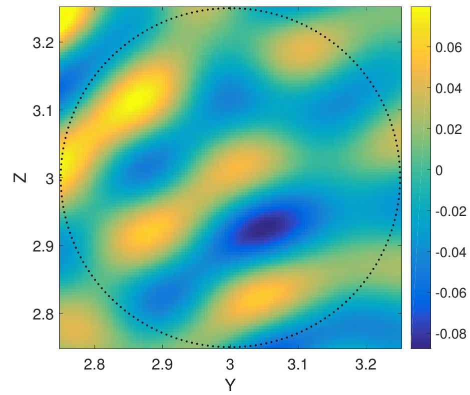
Now that the source is defined, our aim is to use the proposed and original image source methods to predict the response over a spherical receiver region at $\boldsymbol{x}_r$. At 2000 Hz the receiver region is of order 10, and therefore we are simulating the RTF between a first order source and a tenth order receiver.

We first calculate the proposed image source method (17) with the $\beta^{(s)}_{10}(k)$ derived from (19). Next, we use the equivalent point source description to predict the same incident soundfield at $\boldsymbol{x}_r$ utilizing the original image source method (3). Note that this method requires the calculations in (3) to repeat over a multiple times to account for each *point source - point receiver* pair.

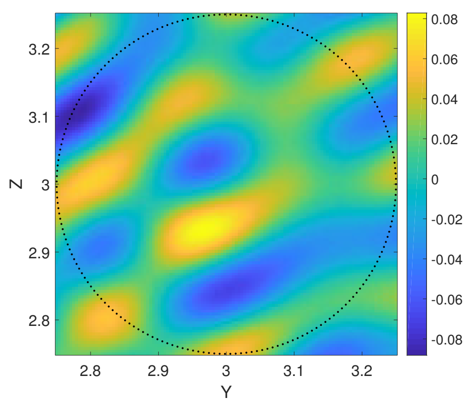
Figure (7) shows the real and imaginary parts of the two soundfields as obtained using the two image source methods. The figures depict a planar cross section parallel to the Y-Z plane across the receiver origin $\boldsymbol{x}_r$ (at $x = 1$). It is visible that the two methods



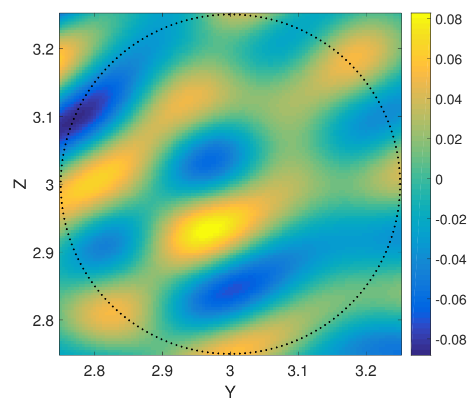
(a) Real part of the RTF: original image source method



(b) Real part of the RTF: generalized image source method



(c) Imaginary part of the RTF: original image source method



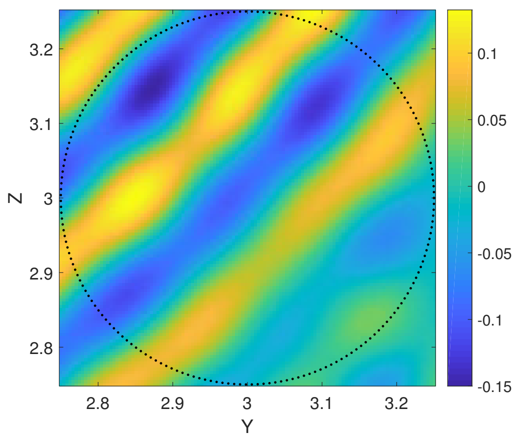
(d) Imaginary part of the RTF: generalized image source method

FIG. 7. (Color online) RTF due to a single-mode directional source as observed over a planar cross section across the receiver origin, parallel to Y-Z plane; comparison between the original image source method and the proposed method.

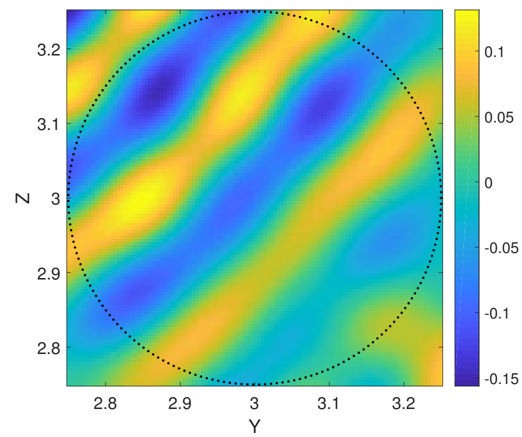
deliver similar results, which validates the accuracy of the proposed image source method for directional transducers. Note that the computational complexity of the two methods at each frequency is different with the generalized image source method being more efficient.

Through simulations we experienced that the most time consuming calculation in both methods is the image generation (the dual triple sum over $\boldsymbol{p}$ and $\boldsymbol{r}$ in (3) and (13)), which exponentially increases with image depth or reflection order R. In the conventional image source method the image generation step is repeated between each and every point source-point receiver combination, which is considerably high given the number of receiver points required to generate a spatial soundfield in the form of Fig 7. In the proposed method, the image generation is only done when calculating the mode coupling coefficients in (13), which is limited to a finite number of $(N+1)^2 \times (V+1)^2$. Once these coefficients are calculated the room response as observed over a spatial region (Fig. 7) can be calculated using (16). Through simulations, we also experienced that with increasing image depth (or reflection order R), the delay in conventional image source method increases further.

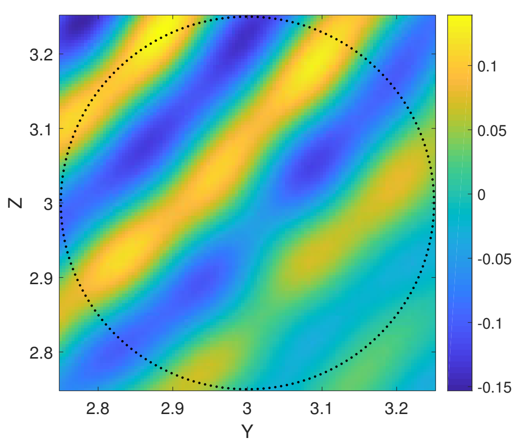
Next we consider an arbitrary directional source that emits multiple outgoing modes as shown in (4). Assuming its equivalent point source description is 3 point sources randomly distributed at $(1, 0.92, 1.085)$, $(1, 1.06, 0.915)$, and $(1.06, 1, 1)$ with respect to $\boldsymbol{x}_s$, the corresponding spherical harmonic coefficients can be obtained using (19) with the summation over d up to 3. Figure 8 shows the resulting 10^{th} order soundfield at $\boldsymbol{x}_r$ based on the proposed image source method and the original one. Similar to the first example, the results are quite similar, which re-validates the accuracy of the proposed method.



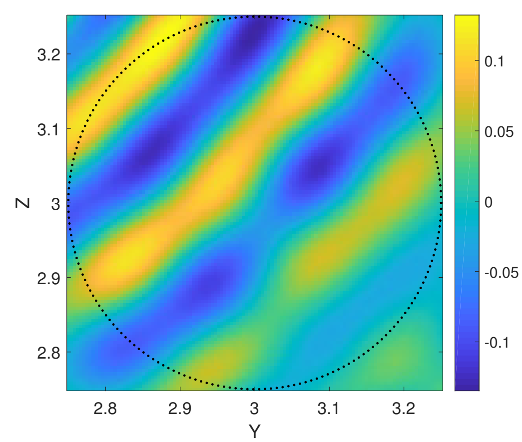
(a) Real part of the RTF; original image source method



(b) Real part of the RTF; generalized image source method



(c) Imaginary part of the RTF; original image source method



(d) Imaginary part of the RTF; generalized image source method

FIG. 8. (Color online) RTF due to a multi-mode directional source as observed over a planar cross section across the receiver origin, parallel to Y-Z plane; comparison between the original image source method and the proposed method

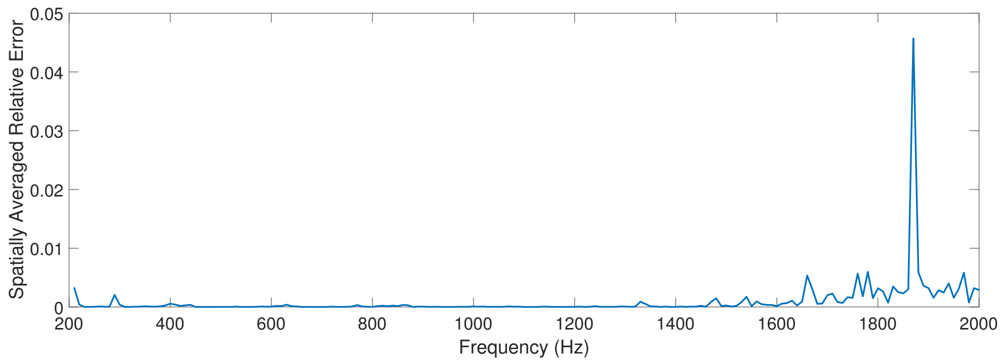


FIG. 9. (Color online) Spatially averaged relative error between the conventional and proposed image source methods

In order to analyze the performance over multiple frequencies, we study the spatially averaged relative error between the two methods, which is defined by

$$E = \frac{\sum_{i=1}^{I} |P_{ISM}(k, \boldsymbol{x}_s, \boldsymbol{z}_i^{(r)}) - P_{GISM}(k, \boldsymbol{x}_s, \boldsymbol{z}_i^{(r)})|^2}{\sum_{i=1}^{I} |P_{ISM}(k, \boldsymbol{x}_s, \boldsymbol{z}_i^{(r)})|^2} \tag{20}$$

where $P_{ISM}(k, \boldsymbol{x}_s, \boldsymbol{z}_i^{(r)})$ denotes the RTF derived by the original image source method (ISM) at the i^{th} receiver position with $i = 1, 2, , \cdots, I$, and $P_{GISM}(k, \boldsymbol{x}_s, \boldsymbol{z}_i^{(r)})$ denotes the same RTF as derived using the proposed generalized image source method (GISM). Figure 9 shows this measure averaged over 400 listening points regularly distributed over the receiver region. The error is plotted in the frequency band $200 - 2000$ Hz. It is clearly seen the error is consistently below $0.005 = 0.5\%$ (except at 1870 Hz), which clarifies the accuracy of the proposed method. The sudden rise at 1870 Hz is due to the denominator of (20) or the original RTF being too small, and the error everywhere else is mainly due to the truncation of equations (4) and (6).

A. Example application of the generalized image source method

Consider an application where the RTF between directional transducers is required for the special case when the source is rotating its look-direction.

With the generalized image source method (17), the RTF for a rotated source can be directly computed using

$$P(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{V} \sum_{u=-v}^{v} \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \rho_{nm}(k)(-1)^{(j+\ell)m+\ell n} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|}$$
$$S_{nv}^{((-1)^{q+j}m)\mu}(\boldsymbol{R_p} + \boldsymbol{R_r}) j_v(kz^{(r)}) Y_{vu}(\theta_z^{(r)}, \phi_z^{(r)}). \tag{21}$$

where ρ_{nm} are the source directivity coefficients of the rotated source. In the spherical harmonics domain, these coefficients are directly related to the original source directivity β_{nm} through the following relationship.

Rotation in the spherical harmonics domain *Let β_{nm} denote the spherical harmonic coefficients in a coordinate system $\boldsymbol{E}$ and let ρ_{nm} denote the spherical harmonic coefficients in a new coordinate system $\boldsymbol{F}$ which is a rotated version of $\boldsymbol{E}$ with the same origin. Assume $(\vartheta, \psi, \gamma)$ are the standard Euler angles*[39] *that define the rotation from $\boldsymbol{E}$ to $\boldsymbol{F}$ using the $z-y-z$ convention in a right-handed frame. That is, the rotation is first done by an angle*

ϑ about the z-axis, then by an angle ψ about the new y-axis, and finally by an angle γ about the new z-axis. Then, the relationship between ρ_{nm} and β_{nm} is given by

$$\rho_{nm}(k) = \sum_{n=0}^{N} e^{im\gamma} d_n^{m'm}(\psi) e^{im\vartheta} \beta_{nm}(k) \tag{22}$$

with

$$d_n^{m'm}(\psi) = [(n+m')!(n-m')!(n+m)!(n-m)!]^{1/2}(-1)^{m'-m} r' \cdots$$
$$\sum_s \frac{(-1)^s (\cos(\psi/2))^{2(n-s)+m-m'}}{(n+m-s)!s(m'-m+s)!(n-m'-s)!}$$

where r' is the radius determining the order of the spherical harmonic decomposition and the range of s is determined by the condition that all factorials are non-negative.

Notice that the above relationship enables RTF calculation between a directional receiver and a rotated source in a single step without having to calculate mode coupling coefficients again. It is also important to mention that a similar rotation can be introduced to a directional receiver with beamforming capabilities. This example prove one advantage of the proposed RTF method, that is not catered by any of the existing model for room response simulation.

VIII. Conclusion

Image source method is one of the most popular techniques to simulate the RTF between a point source and a point receiver. Commercial transducers (especially loudspeakers) used in practice often inherit a directivity pattern. In order to simulate the RTF between such transducers, it is required to incorporate their individual directivity patterns in the room reflection calculations. In this paper, we presented a method to achieve this in the spherical harmonics domain. We represented the directional transducers in terms of spherical harmonic decompositions and derived a compact formula for the respective room response using the image source concept. We provided a number of simulation examples to show the accuracy of the generalized image source method over narrowband and broadband frequencies. Future work involves the derivation of a generalized image source method for room impulse response generation, the time domain counterpart of the proposed method.

APPENDIX: PROOF OF THEOREM 1

Here, we derive the image source method based mode coupling coefficients $\alpha_{vu}^{nm}(k)$ of (11) by comparing (12) and (11). In order to simplify the comparison between (12) and (11), we modify (12) using the addition theorem for Hankel functions[38]. Given three vectors of the form $\boldsymbol{x_1}, \boldsymbol{x_2}$ and $\boldsymbol{x_0}$, such that $\boldsymbol{x_1} = \boldsymbol{x_2} + \boldsymbol{x_0}$, and $|\boldsymbol{x_2}| \leq |\boldsymbol{x_0}|$, the addition theorem for Hankel function is

$$h_n(k|\boldsymbol{x_1}|)Y_{nm}(\theta_{x1}, \phi_{x1}) = \sum_{v=0}^{\infty} \sum_{\mu=-v}^{v} S_{nv}^{m\mu}(\boldsymbol{x_o}) j_v(k|\boldsymbol{x_2}|)Y_{v\mu}(\theta_{x2}, \phi_{x2}) \tag{A.1}$$

$$P_{nm}(k, \boldsymbol{x}_s, \boldsymbol{z}^{(\boldsymbol{r})}) = \sum_{v=0}^{\infty} \sum_{\mu=-v}^{v} \sum_{\boldsymbol{p}=0}^{1} \sum_{\boldsymbol{r}=-\infty}^{\infty} \gamma_{x1}^{|a-q|} \gamma_{x2}^{|a|} \gamma_{y1}^{|b-j|} \gamma_{y2}^{|b|} \gamma_{z1}^{|c-\ell|} \gamma_{z2}^{|c|} (-1)^{(j+\ell)m+\ell n} S_{nv}^{((-1)^{q+j}m)\mu}(\boldsymbol{R_p} + \boldsymbol{R_r})$$
$$j_v(k|\boldsymbol{z}^{(\boldsymbol{r})}|) Y_{v\mu}(\theta_z^{(r)}, \phi_z^{(r)}) \tag{A.2}$$

Results from (A.2) and (11) can now be directly compared to derive the image source method based mode coupling coefficients as given in (13).

References

[1]A. Krokstad, S. Strom, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," Journal of Sound and Vibration **8**(1), 118–125 (1968).

[2]T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan, "A beam tracing method for interactive architectural acoustics," The Journal of the acoustical society of America **115**(2), 739–756 (2004).

[3]F. Antonacci, M. Foco, A. Sarti, and S. Tubaro, "Fast tracing of acoustic beams and paths through visibility lookup," IEEE Transactions on Audio, Speech, and Language Processing **16**(4), 812–824 (2008).

[4]D. Marković, F. Antonacci, A. Sarti, and S. Tubaro, "3d beam tracing based on visibility lookup for interactive acoustic modeling," IEEE transactions on visualization and computer graphics **22**(10), 2262–2274 (2016).

[5]L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," The Journal of the Acoustical Society of America **138**(2), 708–730 (2015).

[6]A. Craggs, *Acoustic modeling: finite element method in Handbook of acoustics* (M. J.Crocker, Ed. John Wiley & Sons, 1998), pp. 149–156.

[7]D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, "Acoustic modeling using the digital waveguide mesh," IEEE Signal Processing Magazine **24**(2), 55–66 (2007).

[8]S. A. Van Duyne and J. O. Smith, "Physical modeling with the 2-d digital waveguide mesh," in *Proceedings of the International Computer Music Conference*, International computer music association (1993), pp. 40–40.

[9]T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," The Journal of the Acoustical Society of America **117**(4), 2100–2111 (2005).

[10]D. Jarrett, E. Habets, M. Thomas, and P. Naylor, "Rigid sphere room impulse response simulation: Algorithm and applications," The Journal of the Acoustical Society of America **132**(3), 1462–1472 (2012).

[11]P. N. Samarasinghe, T. D. Abhayapala, M. Poletti, and T. Betlehem, "An efficient parameterization of the room transfer function," IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP) **23**(12), 2217–2227 (2015).

[12]P. N. Samarasinghe, T. D. Abhayapala, M. A. Polettfi, and T. Betlehem, "On room impulse response between arbitrary points: An efficient parameterization," in *Communications, Control and Signal Processing (ISCCSP), IEEE International Symposium on* (2014), pp. 153–156.

[13]J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," The Journal of the Acoustical Society of America **65**, 943–950 (1979).

[14]T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," IEEE Transactions on Signal Processing, **54**(10), 3790–3804 (2006).

[15]M. Kuster, "Reliability of estimating the room volume from a single room impulse response," The Journal of the Acoustical Society of America **124**(2), 982–993 (2008).

[16]T. Betlehem and M. A. Poletti, "Two dimensional sound field reproduction using higher order sources to exploit room reflections," The Journal of the Acoustical Society of America **135**(4), 1820–1833 (2014).

[17]M. A. Akeroyd, S. Gatehouse, and J. Blaschke, "The detection of differences in the cues to distance by elderly hearing-impaired listeners," The Journal of the Acoustical Society of America **121**(2), 1077–1089 (2007).

[18]S. F. Poissant, N. A. Whitmal III, and R. L. Freyman, "Effects of reverberation and masking on speech intelligibility in cochlear implant simulations," The Journal of the Acoustical Society of America **119**(3), 1606–1615 (2006).

[19]E. A. Habets, S. Gannot, I. Cohen, and P. C. Sommen, "Joint dereverberation and residual echo suppression of speech signals in noisy environments," IEEE Transactions on Audio, Speech, and Language Processing **16**(8), 1433–1451 (2008).

[20]N. H. Adams and G. H. Wakefield, "State-space synthesis of virtual auditory space," IEEE transactions on audio, speech, and language processing **16**(5), 881–890 (2008).

[21]T. Lokki, L. Savioja, R. Vaananen, J. Huopaniemi, and T. Takala, “Creating interactive virtual auditory environments,” IEEE Computer Graphics and Applications **22**(4), 49–57 (2002).

[22]D. Li and M. Hodgson, “Optimal active noise control in large rooms using a locally global control strategy,” The Journal of the Acoustical Society of America **118**(6), 3653–3661 (2005).

[23]M. Vorlander, “Simulation of the transient and steady state sound propagation in rooms using a new combined ray tracing/image source algorithm,” The Journal of the Acoustical Society of America **86**(1), 172–178 (1989).

[24]R. Duraiswami, D. Zotkin, and N. Gumerov, “Fast evaluation of the room transfer function using multipole expansion,” IEEE Transactions on Audio, Speech, and Language Processing **15**(2), 565–576 (2007).

[25]E. A. Lehmann and A. M. Johansson, “Diffuse reverberation model for efficient image-source simulation of room impulse responses,” IEEE Transactions on Audio, Speech, and Language Processing **18**(6), 1429–1439 (2010) doi: 10.1109/TASL.2009.2035038.

[26]J.-H. Pan, C.-c. Bao, B. Bu, and M.-s. Jia, “Measurement of the acoustic transfer function using compressed sensing techniques,” in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016 Asia-Pacific*, IEEE (2016), pp. 1–4.

[27]N. Antonello, E. De Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, “Room impulse response interpolation using a sparse spatio-temporal representation of the sound field,” IEEE/ACM Transactions on Audio, Speech, and Language Processing (2017).

[28]E. Lehmann and A. Johansson, “Diffuse reverberation model for efficient image-source simulation of room impulse responses,” Audio, Speech, and Language Processing, IEEE Transactions on **18**(6), 1429–1439 (2010).

[29]T. Abhayapala and D. Ward, “Theory and design of high order sound field microphones using spherical microphone array,” in *Acoustics, Speech, and Signal Processing (ICASSP), IEEE International Conference on* (2002), Vol. II, pp. 1949–1952.

[30]M. A. Poletti, T. Betlehem, and T. D. Abhayapala, “Higher-order loudspeakers and active compensation for improved 2d sound field reproduction in rooms,” Journal of the Audio Engineering Society **63**(1/2), 31–45 (2015).

[31]M. Poletti, T. Abhayapala, and P. Samarasinghe, “Interior and exterior sound field control using two dimensional higher-order variable-directivity sources,” The Journal of the Acoustical Society of America **131**(5), 3814–3823 (2012).

[32]E. A. Habets, “Room impulse response generator,” Technische Universiteit Eindhoven, Tech. Rep **2**(2.4), 1 (2006).

[33]E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*, 115–125 (Academic Press, London, UK).

[34]D. B. Ward and T. D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” IEEE Transactions on speech and audio processing **9**(6), 697–707 (2001).

[35]W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, “Insights into head-related transfer function: Spatial dimensionality and continuous representation,” The

Journal of the Acoustical Society of America **127**(4), 2347–2357 (2010).

[36]W. J. Thompson, *Angular momentum*, Chap. 1 (John Wiley & Sons).

[37]W. Lin and L. Ballentine, "Quantum tunneling and chaos in a driven anharmonic oscillator," Physical review letters **65**(24), 2927 (1990).

[38]P. Martin, *Multiple scattering: interaction of time-harmonic waves with N obstacles*, Chap. 2, 107 (Cambridge University Press).

[39]A. R. Edmonds, *Angular momentum in quantum mechanics*, Chap. 2.5 (Princeton University Press).