A framework for auralization of boundary element method simulations including source and receiver directivity

Jonathan A. Hargreaves,[a] Luke R. Rendell, and Yiu W. Lam

Acoustics Research Group, Newton Building, University of Salford, Salford, M5 4WT, United Kingdom

The Boundary Element Method (BEM) is a proven numerical prediction tool for computation of room acoustic transfer functions, as are required for auralization of a virtual space. In this paper, it is validated against case studies drawn from the "Ground Truth for Room Acoustical Simulation" database within a framework that includes source and receiver directivity. These aspects are often neglected but are respectively important to include for auralisation applications because source directivity is known to affect how a room is excited and because the human auditory system is sensitive to directional cues. The framework uses weighted-sums of spherical harmonic functions to represent both the source directivity to be simulated and the pressure field predicted in the vicinity of the receiver location, the coefficients of the former being fitted to measured directivity and those of the latter computed directly from the boundary data by evaluating a boundary integral. Three validation cases are presented, one of which includes a binaural receiver. The computed results match measurements closely for the two cases conducted in anechoic conditions but show some significant differences for the third room scenario; here, it is likely that uncertainty in boundary material data limited modelling accuracy.

[LS]

I. INTRODUCTION

Since its inception 25 years ago,[1] auralization has become an important tool for acoustic engineers to communicate the sonic benefits of designs to stakeholders—this is particularly commonplace in architectural and automotive applications. Historically, auralizations were created by playing and recording audio inside physical scale models,[2] but as technology has advanced, the simulation is now mostly performed using computer models.

In order for such an auralization to be accurate, the numerical predictions on which it is based, and the spatial audio encoding and rendering processes used to present it to the listener, must all be accurate too. To date, the majority of simulations for auralization have been conducted using Geometrical Acoustics (GA),[3] for which the spatial audio encoding aspects are straightforward; for example, a ray may be mapped to the closest direction in a Head-Related-Transfer-Function (HRTF) set,[4] or panned between nearby loudspeakers in an array. It is, however, known that these prediction algorithms are inaccurate in certain circumstances, especially at lower frequencies or in smaller rooms, and/or in cases where diffraction or interference effects are significant.[5] Algorithms that model wave effects fully,[6] such as Finite Element Method (FEM), Boundary Element Method (BEM), and Finite-Difference Time-Domain (FDTD), are more accurate and reliable, but processes for encoding their output for auralization are more involved and less well

established. A common approach has been to include a head and torso,[7] or an idealized equivalent,[4,8,9] in the model geometry so that binaural output data can be directly generated by placing receivers at the ear locations. This approach is valid but is inflexible since it fixes the listener position and does not allow for inclusion of personalized HRTFs.

A more flexible approach is to encode the sound-field around the receiver as a weighted sum of spherical harmonic or plane waves. This approach is widely accepted by the sound-field rendering community as an appropriate encoding format for both loudspeaker array-based and binaural reproduction systems,[10,11] and has the added benefit from a prediction-algorithm verification perspective of separating validation of the prediction and rendering processes. It is also consistent with an equivalent representation at high frequencies[12] and, noting that a similar approach may be used to encode source directivity, leads to point-to-point room transfer functions being thought of as having multiple input and output channels.[13] Encoding to such a format from BEM, FEM, or FDTD has to date been achieved by simulating some type of microphone array.[14–16] Encoding of this data is, however, not straightforward and is constrained by many of the factors that affect real microphone array design, with tradeoffs having to be made between array size and density and encoding accuracy.

The only exception to this encoding approach is the 2014 method of Mehra *et al.*[17] that computes the spherical harmonic coefficients from high-order spatial derivatives of the pressure field. When a boundary integral is used to compute the pressure field in the domain, as is in principle

---

[a]Electronic mail: j.a.hargreaves@salford.ac.uk

applicable to FEM and FDTD but is best suited to BEM, these spatial derivatives may be achieved by taking spatial derivatives of the kernel of the integral (i.e., the Green's function). This means that the coefficients may be found directly by a mapping from the boundary data. Since this mapping is independent of the actual boundary data—it only depends on the receiver location—it may be pre-computed to allow interactive update of the scene data e.g., due to changes in the source.

The encoding method applied in this paper achieves the same functionality as the method of Mehra et al.[17] but differs in its mathematical formulation and derivation. In particular, the mathematical formulation[18] is derived from orthogonality statements for "spherical harmonic basis functions" (defined in Sec. II) and gives closed-form statements for the integrals to be evaluated to compute each coefficient, whereas that of Mehra et al. involves evaluating a larger number of integrals and then performing a triple summation to obtain the matrix that maps from boundary data to receiver coefficients. The approach herein may be considered a generalization of the array designs of Hulsebos et al.[19] that is applicable for arbitrary array geometries. These are "open" array designs and are unusual in that they require both pressure and the surface-normal component of pressure gradient at each sensor. Since the array geometry may be chosen freely, it may be taken to be the boundary of the room, at which the pressure and its surface-normal gradient are already known. Compared to Ref. 17, this paper includes substantially more objective validation results and encodes the HRTF datasets in a manner that considers the measurement radius, whereas Mehra et al.[17] appear to transform the simulation results into a set of plane-wave amplitudes and use the HRTF data directly as if it were measured in the far-field. The accuracy of the encoding by either of these approaches will be dictated by the resolution of the boundary mesh, in the same way that for BEM it dictates the accuracy of the sound-field calculated in the domain in general. This has the benefit that there are no parameters or design tradeoffs to be decided by the user and, unlike the mic-array-based encoding methods above, no regularized matrix inversion is required.

A. Hybrid simulation algorithms

An oft-stated limitation of FEM, BEM, and FDTD in room acoustic applications is the rate at which their computational cost increases with frequency. BEM only requires meshing of the two-dimensional boundary, hence to maintain accuracy as frequency f increases the number of Degrees Of Freedom (#DOF) must grow with $\mathcal{O}(f^2)$, compared to $\mathcal{O}(f^3)$ for FEM and FDTD that discretize the domain. However, BEM produces full interaction matrices, linking every element to every other element, leading to computational cost and storage requirements that scale $\mathcal{O}(f^4)$. This has traditionally made it less efficient than FEM or FDTD in most scenarios,[20] primarily finding application in computing scattering from small objects under anechoic conditions,[21] but modern matrix compression techniques such as fast-multipole[22] and adaptive-cross-approximation[23]

can significantly compress the matrices, making BEM competitive in many more scenarios. Even with such developments, however, the scaling of computation cost and storage with frequency for these algorithms is still sufficiently unfavorable so as to preclude full audible-bandwidth simulation for most realistic-sized room acoustics problems of interest.

Auralization of a space, however, requires measured or simulated data covering the full audible frequency spectrum. Since geometrical acoustics algorithms are inaccurate at low frequencies and FEM, BEM, and FDTD are prohibitively computationally expensive at high frequencies, the only way to currently meet the requirement is to combine the output data or two or more algorithms, each run on a section of the frequency spectrum to which they are more suited. This approach was first pioneered by Granier et al.[4] in the mid-1990s using FEM and geometrical acoustics, and the same combination has been studied between 2009 and 2014 in Refs. 7 and 24–27, and by Gómez et al.[15] and Tafur et al.[8] in 2017. BEM was used as the low frequency method by Summers et al.[9] in 2004 and FDTD was used for lower frequency bands in a multiband framework proposed by Southern et al.[28] in 2013. Mehra et al.[17] also used BEM to compute results for auralization in 2014, but extrapolated results to higher frequencies rather than combining them with those of a geometrical acoustics algorithm.

While pragmatic, this hybrid approach opens up another question: how the results from the two algorithms should be combined. It was obvious from the earliest attempts that some form of crossover filter was required between the two models,[4] and that the design of this, e.g., filter lag,[9] would have an effect on the combined Room Impulse Response (RIR) generated. Another issue is how the filtered transfer functions from the two algorithms will interfere with one another once they are combined; Aretz et al.[24] considered this in the most depth and proposed two crossover methods aimed at addressing different concerns.

This paper circumvents those issues and questions by only presenting and assessing the accuracy of the BEM-simulated part of the solution. For frequency-domain results this is straightforward—only the relevant frequency range will be presented—but for time-domain results, a low-pass filter will be employed to minimize Gibbs artefacts; these will be validated against equivalently low-pass filtered measured data. This means that the results herein could be readily combined with high-pass filtered geometrical acoustics results to form a hybrid scheme, so are representative of what the method's performance would be in such a case without opening up questions pertaining to the accuracy of the high-frequency algorithm or combining approach.

B. Input data

Uncertainties, inaccuracy, unsuitability, or presence of gaps in input data is widely acknowledged to be a factor that significantly constrains the accuracy of room acoustic simulations.[5] When dealing with FEM and BEM, for which error bounds can be quantified and which usually produce accurate results for a defined problem if applied correctly, it is

2626   J. Acoust. Soc. Am. **145** (4), April 2019

Hargreaves et al.

reasonable to state that error in input data is the main source of error in output data.

For room acoustics simulations as considered herein, the input data comprises: (i) the geometry of the space and source and receiver locations, (ii) suitable data characterizing the materials present in the space, and (iii) data characterizing the source directivity. For the simulations presented herein, this data was drawn from the Ground truth for Room Acoustical Simulation (GRAS) database[29,30] that was created for the 1st International Round Robin on Auralisation. This database provides high resolution input and output data with the aim of allowing the performance of simulation algorithms to be assessed and improved. Crucially, for the purpose of validating the main contribution of the paper, being an application of the sound-field encoding technique from Ref. 18, it includes measured Binaural RIRs (BRIRs) plus a detailed HRTF set for the Head-And-Torso-Simulator (HATS) used to acquire them.

Real acoustic sources have complex frequency-dependent directivities, but the vast majority of FEM, BEM, or FDTD simulations use simple monopole directivity and it is uncommon to see anything more complicated than a dipole. Part of the reason for this is that it is not trivial to implement higher-order sources in an algorithm that discretizes the domain, though higher-order multipoles have been attempted.[31] A directional source model very similar to that used herein was implemented in FDTD by initializing a wave in the grid,[32,33] but its details appear significantly more complex than that proposed here and the proximity of the source to boundaries is presumably limited. Source strength calibration is in general also non-trivial with FDTD; much of the detail in the multiband framework of Southern *et al.*[28] is concerned with achieving this.

An alternative approach, as implemented herein and in Ref. 17, is to state the incident pressure field analytically and just compute the scattering using the numerical model. This is standard practice in BEM,[34] but is also possible for FEM and FDTD. It circumvents issues to do with the complexity and singular nature of the incident pressure field near the source because the equations for this are only ever evaluated numerically at boundaries, which are assumed to be some distance away. There remains potential for difficulties if a high order source approached a boundary—in this case, the mesh would need to be locally refined to deal with the more spatially concentrated pressure fluctuations—but in most cases, this should not be necessary.

C. Overview of this paper

The primary contribution of this paper is demonstration of the spatial audio encoding process proposed in Ref. 18. The secondary contributions are to demonstrate the effectiveness and accuracy of the full processing chain, including source directivity, BEM simulation, and binaural encoding. Section II presents the mathematical theory behind the source and receiver models and how they are interfaced to BEM. Section III gives more specific implementation details on how they were applied to the dataset used. Section IV presents results validating the simulations against measurements from the GRAS database, then Sec. V draws conclusions and discusses avenues for future research.

II. THEORY

This paper will assume that the medium of wave propagation, the air in the room, is linear, homogeneous, and isotropic, with frequency and position invariant wave speed c_0 and density ρ_0. Real-valued acoustic pressure perturbations $\varphi(\mathbf{x}, t)$, where $\mathbf{x}$ is a point in three-dimensional (3D) Cartesian space and t is time, obey the linear acoustic wave equation $\nabla^2 \varphi = c_0^{-2} \partial^2 \varphi / \partial t^2$. $\Phi(\mathbf{x}, \omega)$ is the complex-valued Fourier transform of φ, where $\omega = 2\pi f$ is angular frequency in radians per second and f is frequency in Hz, which satisfies the Helmholtz equation $\nabla^2 \Phi + k^2 \Phi = 0$, where $k = \omega / c_0$ is the wavenumber in radians per meter. In this paper, $e^{-i\omega t}$ time dependence is assumed for the inverse Fourier transform, that is, a frequency component $\Phi(\mathbf{x}, \omega)$ would produce a time-dependent pressure field $\text{Real}\{\Phi(\mathbf{x}, \omega)e^{-i\omega t}\}$. The majority of this paper will be written in terms of the latter quantity Φ, since the source and receiver descriptions are more easily stated as functions of k and the BEM algorithm used was a frequency-domain code that solves the Helmholtz equation. The measured source data and desired output data for auralization is, however, all time-domain, so the processing necessarily begins and ends with forward and inverse Fourier transforms, respectively, implemented in practice using The Fast Fourier Transform (FFT) algorithm.

A. Source and receiver models

This paper will consider sources and receivers that are spatially compact, so may reasonably be considered as centered on a point in space; these will be denoted $\mathbf{x}_s$ and $\mathbf{x}_r$, respectively. The mathematical descriptions of the pressure fields in the vicinity of each point will be based in a spherical coordinate system (r, α, β). centered on that location, with radius r and azimuthal and zenith angles α and β, respectively.

In these coordinate systems, the pressure of waves that satisfy the Helmholtz equation at frequency ω [with the exception of Eq. (1) at $\mathbf{x} = \mathbf{x}_s$] may be represented in the neighborhood of a $\mathbf{x}_s$ and $\mathbf{x}_r$ by[10,35,36]

$$\Phi_{\text{inc}}(\mathbf{x}) = \sum_{n=0}^{O_s} \sum_{m=-n}^{n} B_{m,n} H_{m,n}^{\text{out}}(\mathbf{x} - \mathbf{x}_s), \tag{1}$$

$$\Phi_{\text{total}}(\mathbf{x}) = \sum_{n=0}^{O_r} \sum_{m=-n}^{n} A_{m,n} J_{m,n}(\mathbf{x} - \mathbf{x}_r). \tag{2}$$

Here, Φ_{inc} is defined to be the incident pressure arriving from some source under anechoic conditions and Φ_{total} is the total pressure including reflections too. Equation (1) is valid when a source is present at $\mathbf{x}_s$, and Eq. (2) is valid and will converge[22] for an expansion point $\mathbf{x}_r$ that is not too close to a source or boundary. $A_{m,n}$ and $B_{m,n}$ are sets of complex frequency-dependent coefficients whose values depend on the pressure field being represented. It is intended that the $B_{m,n}$ coefficients are "input data" arising from the encoding of the directional frequency response of some source (see

J. Acoust. Soc. Am. **145** (4), April 2019

Hargreaves *et al.* 2627

Sec. III B) and that the $A_{m,n}$ coefficients are the output data of this simulation process, being the computed total pressure field encoded as directional coefficients relative to the receiver position ready for presentation by an auralization system. They may therefore respectively be used to represent the directional nature of a source and the directional nature of sound arriving at a receiver. Note that the inevitable scattering by the receiver is not included in Eq. (2); this mechanism is included either implicitly within HRTFs or physically if the sound is rendered to a listener over loudspeakers. The upper limits in n, termed O_r and O_s, are often terms the "order" of the expansion and the number of $A_{m,n}$ and $B_{m,n}$ coefficients is given by $N_r = (O_r + 1)^2$ and $N_s = (O_s + 1)^2$, respectively. The functions $H_{m,n}^{\text{out}}(\mathbf{r})$ and $J_{m,n}(\mathbf{r})$, plus another $H_{m,n}^{\text{in}}(\mathbf{r})$ that will be required later, are defined

$$H_{m,n}^{\text{out}}(\mathbf{r}) = Y_n^m(\beta, \alpha) h_n^{\text{out}}(kr), \tag{3}$$

$$H_{m,n}^{\text{in}}(\mathbf{r}) = Y_n^m(\beta, \alpha) h_n^{\text{in}}(kr), \tag{4}$$

$$J_{m,n}(\mathbf{r}) = Y_n^m(\beta, \alpha) j_n(kr). \tag{5}$$

Here, h_n^{out} and h_n^{in} are spherical Hankel functions of order n that are "outgoing" and "incoming," respectively; with $e^{-i\omega t}$ time dependence, as assumed herein, they will be of the first and second kind respectively. $j_n(kr) = \frac{1}{2} h_n^{\text{out}}(kr) + \frac{1}{2} h_n^{\text{in}}(kr)$ is a spherical Bessel function. $Y_n^m(\beta, \alpha)$ is a spherical harmonic function of order m, n. A number of marginally different normalization schemes exists, but in this paper they are defined[22]

$$Y_n^m(\beta, \alpha) = (-1)^m \sqrt{\frac{(2n+1)}{4\pi} \cdot \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos \beta) e^{im\alpha}. \tag{6}$$

Here, $P_n^m(\cdots)$ is an associated Legendre polynomial.

Spherical harmonic functions are also used to interpolate source directivities and HRTFs in geometrical methods at high frequencies.[12,37] There, the radial propagation is assumed to match that of a monopole regardless of n, so directivity becomes independent of r; this is equivalent to replacing all h_n^{out} in Eq. (3) by h_0^{out}, which is appropriate since $h_n^{\text{out}}(kr) \approx i^{-n} h_0^{\text{out}}(kr)$ for large kr.[38] Representations like Eqs. (1) and (2) are in contrast used when distance is

considered important, e.g., for near-field compensated Ambisonics[10] or HRTF range extrapolation.[39] Directivity is usually measured over a sphere, and data measured in this way may be encoded as $A_{m,n}$ and $B_{m,n}$ coefficients so long as the measurement radius is known. Alternatively, techniques that use double-layer array measurements may obtain $B_{m,n}$ coefficients directly.[36,40,41]

B. Boundary integral equations

The Kirchhoff-Helmholtz Boundary Integral Equation (KHBIE) is found by applying Green's second theorem to a pair of acoustic waves over some domain Ω_+ that contains the acoustic medium.[34] One of the waves is Φ, the pressure-field under study, which satisfies the Helmholtz equation everywhere in Ω_+. The other is the free-space acoustic Green's function $G(\mathbf{x}, \mathbf{y}) = e^{ik|\mathbf{x}-\mathbf{y}|}/4\pi|\mathbf{x} - \mathbf{y}|$, which satisfies the Helmholtz equation for all $\mathbf{y} \in \Omega_+$ $\mathbf{x}$. The result is a surface integral over the boundary Γ that contains Ω_+. For a scattering problem with $\Phi_{\text{total}} = \Phi_{\text{inc}} + \Phi_{\text{scat}}$, where Φ_{inc} is the aforementioned pressure-field radiated by the source in anechoic conditions and Φ_{scat} is the difference that occurs due to reflections from the boundary, this may be expressed as

$$\Phi_{\text{scat}}(\mathbf{x}) = \iint_{\Gamma} \left[\Phi_{\text{total}}(\mathbf{y}) \frac{\partial G}{\partial n_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\text{total}}}{\partial n_{\mathbf{y}}}(\mathbf{y}) \right] d\Gamma_{\mathbf{y}}. \tag{7}$$

Here, the notation $\partial/\partial n_{\mathbf{y}}$ is shorthand for $\hat{\mathbf{n}}_{\mathbf{y}} \cdot \nabla_{\mathbf{y}}$, where $\hat{\mathbf{n}}_{\mathbf{y}}$ is a unit vector pointing normal to Γ and into Ω_+ at point $\mathbf{y} \in \Gamma$, and subscript $\mathbf{y}$ means "with respect to or evaluated at point $\mathbf{y}$."

Equation (7) is the basis of our BEM formulation. Total pressure Φ_{total} should equal zero in a domain Ω_- that is on the opposite side of Γ to Ω_+, hence $\Phi_{\text{scat}}(\mathbf{x}) = -\Phi_{\text{inc}}(\mathbf{x})$ for $\mathbf{x} \in \Omega_-$. Taking the limit as $\mathbf{x}$ approaches Γ from within Ω_- produces an inhomogeneous Fredholm integral equation of the second kind. This may be solved by discretizing the boundary quantities Φ_{total} and $\partial\Phi_{\text{total}}/\partial n_{\mathbf{y}}$ on a boundary mesh and then solving the resulting matrix equation numerically. More details on this process are given in Sec. III C.

Consider now the simulation process architecture shown in Fig. 1. The reader is encouraged to notice the similarities between this framework and the high-frequency geometrical acoustics framework given in Fig. 2 in Ref. 12. Blocks 1 and
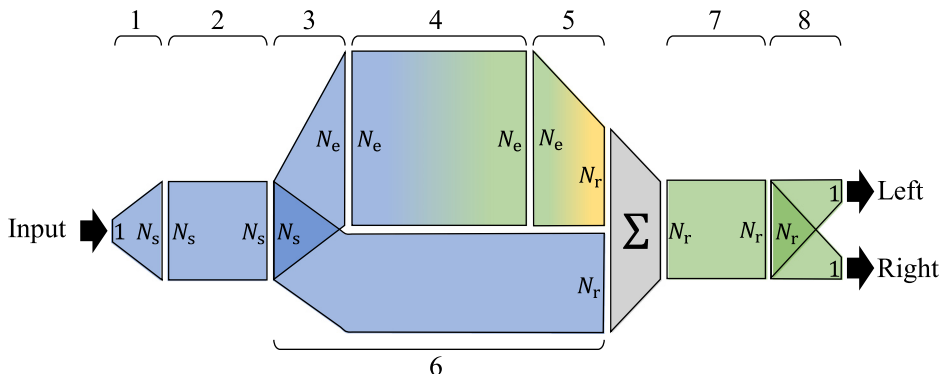

FIG. 1. (Color online) Processing blocks, indicating numbers of degrees of freedom involved: (1) Source directivity, (2) Source rotation, (3) Source to boundary Φ_{inc} mapping, (4) Boundary to boundary Φ_{inc} to Φ_{total} mapping—BEM solution, (5) Boundary to receiver Φ_{scat} mapping, (6) Source to receiver Φ_{inc} mapping, (7) Receiver rotation, (8) HRTFs.

Hargreaves *et al.*

8 are, respectively, the encoding of measured source directivity as $B_{m,n}$ coefficients following Eq. (1) and of HRTFs as $A_{m,n}$ coefficients following Eq. (2). Blocks 2 and 7 are rotations of these, and allow for changes in source and receiver orientation; this may be readily achieved in the Spherical Harmonic domain by a matrix multiplication.[22] Block 3 is simply the evaluation of Eq. (1) for points on the boundary and block 4 is the BEM solution. This leaves processes for blocks 5 and 6 to be identified. Note that these each separately encode Φ_{inc} and Φ_{scat} in the form of Eq. (2) as separate sets of coefficients $A_{m,n}^{\text{inc}}$ and $A_{m,n}^{\text{scat}}$ that are then summed.

A solution to implementing the process in block 6 and find $A_{m,n}^{\text{inc}}$ is in fact well known; it may be achieved in a straightforward way using a translation operator in the Spherical Harmonic domain.[22] How to achieve the process in block 5, mapping boundary pressure in a BEM model to scattered pressure encoded as coefficients $A_{m,n}^{\text{scat}}$, is not as well established, with simulation of virtual mics arrays previously being the norm as discussed in Sec. I and the direct approach of Mehra et al.[17] being the state of the art. In this paper, the alternative direct approach by Hargreaves and Lam[18] is implemented. This allows $A_{m,n}^{\text{scat}}$ to be found by evaluation of the following boundary integral equation,[42] where a bar over a quantity indicates conjugation

$$A_{m,n}^{\text{scat}} = ik \iint_{\Gamma} \left[\Phi_{\text{total}}(\mathbf{y}) \overline{\frac{\partial H_{m,n}^{\text{in}}}{\partial n_{\mathbf{y}}}} (\mathbf{y} - \mathbf{x}_{\text{r}}) \right.$$
$$\left. - \overline{H_{m,n}^{\text{in}}} (\mathbf{y} - \mathbf{x}_{\text{r}}) \frac{\partial \Phi_{\text{total}}}{\partial n_{\mathbf{y}}} (\mathbf{y}) \right] d\Gamma_{\mathbf{y}}. \quad (8)$$

Equation (8) possesses clear similarities to the KHBIE in Eq. (7). Noting in particular that $\overline{H_{0,0}^{\text{in}}}(\mathbf{y} - \mathbf{x}_{\text{r}}) = G(\mathbf{x}_{\text{r}}, \mathbf{y}) \times \sqrt{4\pi}/ik$ and that $J_{m,n}(0) = 1/\sqrt{4\pi}$ if $m = n = 0$ and is zero otherwise, hence $\Phi_{\text{scat}}(\mathbf{x}_{\text{r}}) = A_{0,0}^{\text{scat}}/\sqrt{4\pi}$, it is apparent that Eq. (7) is in fact a special case of Eq. (8) for $m = n = 0$. Equation (8) may be implemented numerically in a similar manner to how Eq. (7) is for omnidirectional external receivers. $H_{m,n}^{\text{in}}$ contains a higher-order singularity than G for $n > 1$, and will also contain angular oscillations that G does not, but both of these characteristics should be resolved well by standard quadrature techniques on a mesh that is fine with respect to wavelength, so long as receivers are not too close to a boundary and O_{r} is not unnecessarily high at low frequencies. More detail is given in Sec. III C.

Figure 1 also displays the #DOF present at the interfaces between the various processes; for most the computational cost of each process will scale with the product of these, though for a direct implementation of process 6, it scales an order of magnitude worse.[22] In practice, however, processes 3, 4, and 5 involving N_{e}, the #DOF in the BEM mesh, will dominate simply because N_{e} is usually a few orders of magnitude greater than N_{s} or N_{r}; for the case studies considered herein $N_{\text{s}} = 25$ and $N_{\text{r}} = 121$, whereas N_{e} was typically in the order of the tens of thousands. Process 4, the BEM solution, is therefore expected to be the most computationally intensive, since its computation cost is proportional to N_{e}^2. It will be seen from the test cases that this was, however, not

the case; the libraries used to evaluate this stage are the most optimized and, combined with the Adaptive-Cross-Approximation (ACA) solver,[23] this stage is neither the slowest nor the one with the worst computational cost scaling. Values of N_{s}, N_{e}, and N_{r} are necessary to maintain accuracy as frequency increases all scale $\mathcal{O}(f^2)$, so the total computational cost of the algorithm is expected to scale $\mathcal{O}(f^4)$.

III. IMPLEMENTATION

The above framework was implemented for a subset of the scenes from the GRAS database.[29,30] The database contains 11 scenes, some with multiple variants, with seven being fully or hemi anechoic laboratory setups and the remaining four being room acoustic scenarios. In this paper, results for the following three scenes are presented:

- Scene 1: Simple reflection (infinite plate) – hard floor variant
- Scene 3: Multiple reflection (finite plate)
- Scene 9: Small room (seminar room)

Scene 1 actually included two other variants; one with a mineral wool slab placed on the floor and another with a Medium Density Fiberboard (MDF) diffuser. These cases have also been attempted with reasonable success but are not reported here due to space limitations. It should be noted, however, that they are numerically challenging in BEM primarily due to the extreme aspect ratio of the samples,[43] requiring higher element counts and more accurate numerical integration than would normally be expected. The hard floor variant is included since its implementation is essentially image source with directivity, i.e., no BEM mesh, so it allows the accuracy of the source representation and encoding process to be independently quantified.

Scene 3 comprised two 2 m square 25 mm thick MDF panels separated by a distance of 10 m, with a source and receiver location both located on the center line between these each spaced 3 m from one of the panels. This is an interesting configuration to simulate, since it will give a flutter echo the damping of which is dictated as much by diffraction as by material absorption.[44] Scene 3 is also a good test case for binaural reproduction, since the HATS was orientated so that reflections occur from ear to ear across the head.

Scene 9 was a small seminar room with "relatively simple and easy to describe geometry, but challenging low frequency behavior," so it presents another interesting case for which to apply BEM simulation. It also typifies the input data challenges that a user of BEM encounters in reality, so it is included to demonstrate how those affect simulation accuracy.

The room was nominally 8.5 m long by 6 m wide by 3 m high, though it has various inclusions; for details, see Ref. 30. Material data was provided in the form of third-octave absorption coefficients that were established through a mixture of in situ measurement and estimation based on published database values,[29] the latter being required because the former is known to be inaccurate at low frequencies[45] and for materials with low absorption. The result was that the provided data both did not extend down to the lowest frequencies that were simulated and tended to unrealistically

J. Acoust. Soc. Am. **145** (4), April 2019

Hargreaves et al.   2629

small absorption values in this range; if not addressed, this would have led to very low modal damping and unrealistically high sound pressure levels (SPLs). Anecdotally, we were informed that the seminar room had some lightweight panel walls, and it is likely that full-panel membrane motion, which is a significant source of absorption and modal damping at low frequencies, was present but not characterized by this data. Moreover, a door was present in the room and will likely give high losses at low frequencies due to transmission, but no material data was provided for it. A process of extrapolation, modification and fabrication was therefore required to achieve an appropriate and realistic set of boundary data; this is briefly outlined in Sec. III E. The RIRs were measured with a bespoke multiway dodecahedron loudspeaker,[30] but the directional narrowband response data available for the other loudspeakers was not available for this, adding another source of uncertainty.

Temperature and humidity data was provided for each scene in the database, but it was not straightforward to use this since there was often a slight mismatch between the conditions when the source and HATS directivities were measured and those when the full scene was measured. Consequently, $c_0 = 343\text{m/s}$ and $\rho_0 = 1.21\text{kg/m}^3$ were assumed for all processes instead.

A. Fourier transforms and filtering

The simulation process depicted in Fig. 1 will be performed entirely in the frequency domain, as was also the case for all the BEM and FEM algorithms discussed in Sec. I A, meaning that an inverse Fourier transform is required in order to obtain an auralizable IR. It should also be noted that both BEM and FEM also exist as time-domain solvers,[46–48] but these are less mature and have not been applied in any hybrid framework to date.

When using an Inverse Fast Fourier Transform (IFFT) algorithm to achieve this, results for many frequencies are required at a spacing $\Delta f = 1/T$, where T is the required IR length; this must be long enough that the IR decays to a negligible level to minimize wrap-around error. For scene 1 it was taken that $\Delta f = 2$ Hz, so $T = 0.5$ seconds, and for scenes 3 and 9, where higher order reflections are expected, it was it was taken that $\Delta f = 0.5$ Hz, so $T = 2$ seconds. This was, however, found to be insufficient for scene 9, so the room transfer function was spline interpolated in the frequency domain, as suggested by Aretz et al.,[24] to give $\Delta f = 0.25$ Hz and $T = 4$ seconds. The low-pass filter was chosen to be an 8th order Butterworth filer, following method 1 of Aretz et al.,[24] and this was applied in the frequency domain pre-IFFT.

Running a BEM at so many closely spaced frequencies is not an efficient application of the algorithm since the interaction matrices must be reconstructed from scratch for each frequency; it is tolerated here for validation purposes. Multifrequency BEM[49,50] provides a solution to this by interpolating the interaction matrices between neighboring frequencies. Values from preceding frequencies were, however, used to "seed" the iterative matrix solver, meaning fewer solver iterations were required.

It is also necessary to run the algorithm beyond the intended crossover frequency so that data is available for the frequency region where the filter "rolls-off." Aretz et al.[24] recommended this should be at least half an octave above crossover frequency, but even this appears rather optimistic when considered in terms of typical filter roll-off in dB/octave, and artefacts were reported as being visible in their RIRs. In these simulations, it was chosen that the simulations should extend one full octave above the intended crossover frequency. To mitigate the computational cost that this incurred, the mesh was not refined further beyond the crossover frequency, meaning that computational cost was fixed but accuracy reduced with increasing frequency; this is acceptable since those frequencies will be heavily attenuated by the crossover filter. The crossover frequency was chosen to be 1 kHz for scenes 1 and 3 and 400 Hz for scene 9, hence the maximum frequencies simulated were 2 kHz and 800 Hz, respectively.

B. Source directivity and HRTF encoding

Three sound sources were used in the scenarios chosen; a Genelec 8020c for scenes 1 and 3, and a QSC K8 and a custom three-way dodecahedron for scene 9, the former being used for BRIRs and the latter for RIRs. The only data available on the dodecahedron loudspeaker was that it aimed to be omnidirectional with a flat frequency response above 40 Hz, so it was assumed to follow this. For the Genelec and QSC loudspeakers extremely high-resolution data was available, being a set of 64 442 impulse responses measured at points on a sphere centered on the source; this was 2 m radius for the 8020c and 8 m for the K8. These were first zero-padded to match the intended output RIR length and then FFT'd to acquire a set of complex frequency-domain transfer functions (units Pa/V); these were taken to be the incident pressure Φ_{inc} measured at each point. To encode them to a set of coefficients $B_{m,n}$, one solution is to create a matrix equation to be inverted where each row is Eq. (1) applied at a different measurement point. However, here the number of measurement points was so great that the orthogonality of Y_n^m over a sphere could be exploited to calculate $B_{m,n}$ directly. For a given frequency, this integral is approximated by a finite sum over measurement angles

$$
\begin{aligned}
B_{m,n} &= \frac{1}{h_n^{\text{out}}(kr)} \int_0^{2\pi} \int_0^{\pi} \Phi_{\text{inc}}(r, \beta, \alpha) \overline{Y_n^m(\beta, \alpha)} \sin\beta\, d\beta\, d\alpha \\
&\approx \frac{1}{h_n^{\text{out}}(kr)} \sum_{p=1}^{64,442} \Phi_{\text{inc}}(r, \beta_p, \alpha_p) \overline{Y_n^m(\beta_p, \alpha_p)} w_p.
\end{aligned}
\tag{9}
$$

Here, w_p is equal to the area of the sphere that is closest to the pth point. Finally, the coefficients for each source were scaled by a frequency-independent calibration factor so that an SPL of 80 dB was produced at 2 m at 1 kHz, matching the procedure performed for the measurements.[30]

The measurement process used to acquire the HRTF library is detailed in Ref. 51. This also takes the form of a set of IRs acquired at different angles, but here they were measured using a loudspeaker mounted on an arc of 1.7 m

2630   J. Acoust. Soc. Am. **145** (4), April 2019

Hargreaves *et al.*

radius, and for each source location, there are two IRs: left and right. In the absence of directivity information, the source will be assumed to be a monopole; this appears not unreasonable since in Ref. 52 it is validated against BEM simulations performed by reciprocity using a source at the ear and an omnidirectional point receiver at the loudspeaker center. Measured HRTFs were normalized by removing the microphones from the ears and placing them at the origin then repeating the experiment.

Fundamentally, HRTFs are linear mappings $\mathcal{L}$ and $\mathcal{R}$ from the incoming pressure field $\Phi_{\text{total}}(\mathbf{x})$, as exists in the absence of the HATS, to the pressures at the left and right ears Φ_{L} and Φ_{R}. Assuming that Φ_{total} is represented by Eq. (2), these become discrete mappings $\Phi_{\text{L}} = \mathbf{AL}$ and $\Phi_{\text{R}} = \mathbf{AR}$, where $\mathbf{A}$ is a row vector containing the elements of $A_{m,n}^{\text{total}}$ and $\mathbf{L}$ and $\mathbf{R}$ are column vectors that are the discrete form of $\mathcal{L}$ and $\mathcal{R}$. Including the normalization by the pressure at the HATS center, this amounts to the elements of $\mathbf{A}$ being defined for the pth point by $A_{m,n}^{\text{total}} = 4\pi \overline{Y_n^m(\beta_p, \alpha_p)} h_n^{\text{out}}(kr)/h_0^{\text{out}}(kr)$, and stacking rows for all the measured points produces a matrix equation to be solved. Since the number of measurement points was so great, it was possible to solve without regularization using a standard least-squares technique.

The accuracy achieved by these encoding processes is shown for the Genelec loudspeaker and the HATS in Fig. 2; a similar figure for the QSC K8 is included in the supplementary material.[63] Here, the encoding error is quantified as normalized L^2 error, being the L^2 norm of the residual between the measured data and the encoded version evaluated on the measurement surface, divided by the L^2 norm of the measured data. The surface integrals involved in the L^2 norms were approximated by weighted sums in the same manner as was done in Eq. (9).

In both cases the approximation improves with maximum spherical harmonic order O, and a higher value of O is required to maintain the same accuracy as frequency f is increased. The lower right region of the two plots is quite similar, but with slightly smaller residual for the HRTFs. Accuracy deteriorates on the far left of Fig. 2(a) due to the singular nature of spherical Hankel functions at small kr. In contrast, a clear change in the trend-lines is seen in Fig. 2(b) at 200 Hz; it seems likely that this is caused by the transition between measured and simulated data that was necessary when creating the HRTF library.[52]
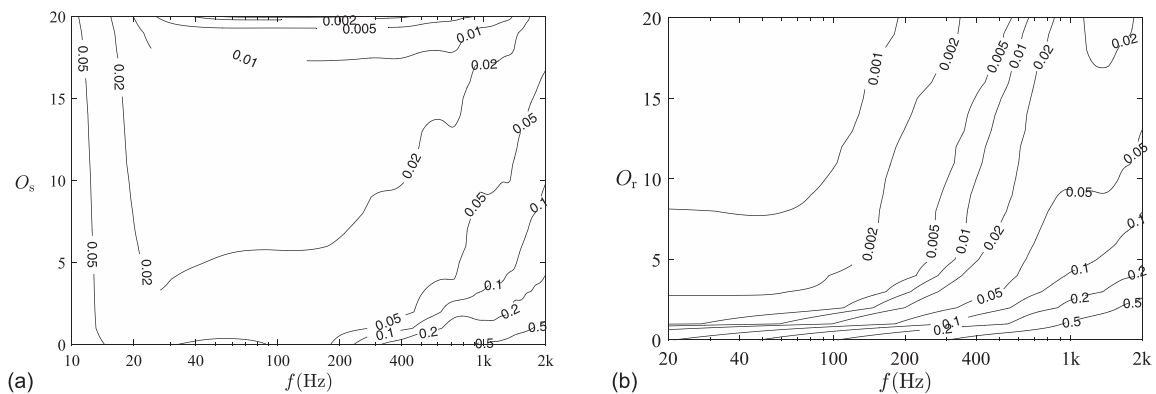
Figure 2 suggests that the optimal values of O_{s} and O_{r} should change with frequency. This was initially attempted, but sharp transitions between orders were visible in the BEM results, hence this approach was discarded. For simplicity, constant values of O_{s} and O_{r} were instead used for all frequencies; these were chosen to be $O_{\text{s}} = 4$ and $O_{\text{r}} = 10$; hence, $N_{\text{s}} = 25$ and $N_{\text{r}} = 121$. The greater value of O_{r} was chosen to allow the encoding process in Eq. (8) to be tested.

C. BEM

The BEM simulations were performed using BEM++[23,53,54] version 3.1. This is an open-source BEM library that is invoked from Python scripts, creating a flexible interface that allows the boundary integral operators provided to be assembled in customizable ways. BEM++ implements a Galerkin BEM algorithm in 3D and includes an ACA solver that accelerates matrix assembly and solution.

For the solution, the Helmholtz equation, BEM++ provides four standard boundary integral operators that each map, in a different way, a quantity Φ defined on a boundary section Γ to a location $\mathbf{x}$

$$\mathcal{S}\{\Phi\}(\mathbf{x}) = \iint_\Gamma \Phi(\mathbf{y}) G(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}}, \tag{10}$$

$$\mathcal{D}\{\Phi\}(\mathbf{x}) = \iint_\Gamma \Phi(\mathbf{y}) \frac{\partial G}{\partial n_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}}, \tag{11}$$

$$\mathcal{A}\{\Phi\}(\mathbf{x}) = \iint_\Gamma \Phi(\mathbf{y}) \frac{\partial G}{\partial n_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}}, \tag{12}$$

$$\mathcal{H}\{\Phi\}(\mathbf{x}) = \iint_\Gamma \Phi(\mathbf{y}) \frac{\partial^2 G}{\partial n_{\mathbf{x}} \partial n_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}}. \tag{13}$$

Here, $\mathcal{S}$, $\mathcal{D}$, $\mathcal{A}$ and $\mathcal{H}$ are, respectively, termed: the single-layer potential, the double-layer potential, the adjoint double-layer potential, and the hypersingular operator. Note that the definition of $\mathcal{H}$ has here been written negated compared to the convention used in the library. Additionally, an identity operator $\mathcal{I}\{\Phi\}(\mathbf{x}) = \Phi(\mathbf{x})$ is also defined. The Python objects representing these operators may be added,

FIG. 2. Normalized source and receiver directivity encoding error versus frequency f and maximum spherical harmonic order O for: (a) Genelec 8020c, (b) HATS HRTFs.

multiplied by each other or scalars, or concatenated to form blocked operators, as required by the problem being studied.

The discretized versions of these boundary operators, for surface to surface mappings, are found using the Galerkin method. Rather than solving a boundary integral equation (BIE) for a finite set of points on the boundary, as is done in the colocation method, this solves the "weak-form" of the BIE on average over the entire boundary, and hence involves a second surface integral.[34] A set of weighting functions are chosen to spatially weight this "testing" process and produce each row in the matrix equation; in a Galerkin scheme these are equal to the basis functions used to discretize the boundary quantities. An operator $\mathcal{K} \in \{\mathcal{S}, \mathcal{D}, \mathcal{A}, \mathcal{H}, \mathcal{I}\}$ is therefore mapped to its discrete matrix form $\mathbf{K} \in \{\mathbf{S}, \mathbf{D}, \mathbf{A}, \mathbf{H}, \mathbf{I}\}$ with entries given by

$$\mathbf{K}(i,j) = \int\int_\Gamma \overline{b_i(\mathbf{x})} \mathcal{K}\{b_j\}(\mathbf{x}) d\Gamma_\mathbf{x}. \tag{14}$$

Here, b_j is basis function drawn from the set used to representing the "radiating" quantity, and b_i is a basis function drawn from the set used to representing the "receiving" quantity that is being "tested"; these sets are not necessarily the same and may be chosen differently on different boundary sections or for representing a different quantity. A similar statement maps a user-defined Python function $f(\mathbf{x})$, typically used to compute the incident wave Φ_{inc}, onto the "receiving" set of basis functions. This produces a vector $\mathbf{f}$ with entries defined

$$\mathbf{f}(i) = \int\int_\Gamma \overline{b_i(\mathbf{x})} f(\mathbf{x}) d\Gamma_\mathbf{x}. \tag{15}$$

This interface is extremely flexible, but can be slow for complicated functions, since Python is an interpreter language and the functions are evaluated on a per-abscissa basis. In this scheme for instance, $f(\mathbf{x})$ is typically a Spherical Harmonic basis function (or its spatial derivative), in accordance with the source definition. It will be seen in Sec. IV E that this operation, which is normally assumed to trivially quick compared to assembly and solution of the linear system, actually takes the longest due to the complexity of these functions and the slow nature of native Python code compared to the compiled core libraries that BEM++ in built on.

The KHBIE in Eq. (7) can be written using $\mathcal{D}$ and $\mathcal{S}$ as $\Phi_{\text{scat}}(\mathbf{x}) = \mathcal{D}\{\Phi_{\text{total}}\}(\mathbf{x}) - \mathcal{S}\{\partial\Phi_{\text{total}}/\partial n_\mathbf{y}\}(\mathbf{x})$. In order to implement the spherical harmonic encoding in Eq. (8), two new boundary operators are defined

$$\mathcal{S}_{m,n}^{\text{in}}\{\Phi\}(\mathbf{x}) = \int\int_\Gamma \Phi(\mathbf{y}) \overline{H_{m,n}^{\text{in}}}(\mathbf{y} - \mathbf{x}) d\Gamma_\mathbf{y}, \tag{16}$$

$$\mathcal{D}_{m,n}^{\text{in}}\{\Phi\}(\mathbf{x}) = \int\int_\Gamma \Phi(\mathbf{y}) \frac{\overline{\partial H_{m,n}^{\text{in}}}}{\partial n_\mathbf{y}}(\mathbf{y} - \mathbf{x}) d\Gamma_\mathbf{y}. \tag{17}$$

This allows Eq. (8) to be written as $A_{m,n}^{\text{scat}} = ik\mathcal{D}_{m,n}^{\text{in}}\{\Phi_{\text{total}}\}(\mathbf{x}_\mathbf{r}) - ik\mathcal{S}_{m,n}^{\text{in}}\{\partial\Phi_{\text{total}}/\partial n_\mathbf{y}\}(\mathbf{x}_\mathbf{r})$. It follows that $\mathcal{S}_{0,0}^{\text{in}} = \mathcal{S} \times \sqrt{4\pi}/ik$ and $\mathcal{D}_{0,0}^{\text{in}} = \mathcal{D} \times \sqrt{4\pi}/ik$, hence

$\mathcal{D}_{m,n}^{\text{in}}$ and $\mathcal{S}_{m,n}^{\text{in}}$ may be viewed as a generalization of the standard operators $\mathcal{D}$ and $\mathcal{S}$.

The discrete form of $\mathcal{D}_{m,n}^{\text{in}}$ and $\mathcal{S}_{m,n}^{\text{in}}$ have matrix entries given by equations similar to Eqs. (16) and (17) but with b_j appearing in place of Φ. These are, unsurprisingly, not built into BEM++, but Eq. (15) provides a method to evaluate them. Python functions that compute $\overline{H_{m,n}^{\text{in}}}(\mathbf{y} - \mathbf{x}_r)$ and $\overline{\partial H_{m,n}^{\text{in}}/\partial n_\mathbf{y}}(\mathbf{y} - \mathbf{x}_r)$ may be passed to the routine that implements Eq. (15), and the result is a coefficient from the discrete matrix form of $\mathcal{S}_{m,n}^{\text{in}}$ and $\mathcal{D}_{m,n}^{\text{in}}$, respectively. Note that evaluating this in this way is an extremely slow process; the reasons stated above all still apply, but are compounded by the fact that every spherical harmonic order must be evaluated separately, and the associated Legendre polynomials therein are much more efficiently evaluated simultaneously for all m of an order n rather than separately. For the verification purposes herein, however, this inefficiency is tolerated.

The GMRES iterative matrix solver included with SciPy[55] was used to solve the matrix equations produced using the BEM++ operators. The ACA accuracy and maximum block size parameters, and the GMRES solver tolerance, were left at their default values of 1×10^{-3}, 2048 and 1×10^{-5}, respectively, for scene 3. For scene 9, the former two were reduced to 1×10^{-5} and 128, respectively, due to some convergence issues with matrix solver; this improved accuracy of the matrix approximation at the expense of increased storage and computational cost. Meshing was performed using the open-source meshing tool Gmsh.[56,57] Maximum element size at any given frequency followed $\lambda/8$, with a minimum limit chosen to match $\lambda/8$ at the crossover frequency.

D. Scene-specific implementation

1. Scene 1H

This scene comprised a loudspeaker above a hard floor in a hemi-anechoic chamber; this was assumed to be perfectly reflecting. When applying BEM in half-space problems such as this, it is common to apply an image-source principle and reflect the source and all the obstacle boundaries in the rigid ground plane. Here, there is no additional obstacle, so there is actually no BEM solution at all; there is just the source and the image source, which has reflected directivity. Finding the pressure at a point is as simple as finding Φ_{inc} from both sources and summing. Similarly, $A_{m,n}$ coefficients may be found by applying the translation techniques in Ref. 22 to each source and then summing.

2. Scene 3

This scene included panels that were much thinner in one dimension than the other two. Following standard BEM practice,[43,58] the obstacle was replaced by one of zero thickness lying on its center line. The material had a specific admittance Y that was uniform over both sides of the panel; under this condition the formulation given in Ref. 58 can be simplified to solving the following pair of coupled BIE for all $\mathbf{x} \in \Gamma$: $\mathcal{D}\{\tilde{\Phi}_{\text{total}}\}(\mathbf{x}) + [ikY\mathcal{S} - \frac{1}{2}\mathcal{I}]\{\check{\Phi}_{\text{total}}\}(\mathbf{x}) = -\Phi_{\text{inc}}(\mathbf{x})$ and $[\mathcal{H} + \frac{1}{2}ikY\mathcal{I}]\{\tilde{\Phi}_{\text{total}}\}(\mathbf{x}) + ikY\mathcal{A}\{\check{\Phi}_{\text{total}}\}(\mathbf{x}) = -\partial\Phi_{\text{inc}}/\partial n_\mathbf{x}(\mathbf{x})$. These two boundary integral equations were

combined as a blocked operator in BEM++. Here, $\breve{\Phi}_{\text{total}}$ and $\tilde{\Phi}_{\text{total}}$ are respectively the sum of, and the difference between, the pressures acting on the front and back surfaces of the panel. A discontinuous piecewise-constant space was used to discretize $\breve{\Phi}_{\text{total}}$ and to "test" the first equation. A continuous piecewise-linear space was used to discretize $\tilde{\Phi}_{\text{total}}$ and to "test" the second equation. This arrangement is consistent with the expectation that $\tilde{\Phi}_{\text{total}} \rightarrow 0$ at the edge of the panel[59] and meets the continuity requirements of $\mathcal{H}$.[60] Pressure at receiver locations can be evaluated by $\Phi_{\text{scat}}(\mathbf{x}) = \mathcal{D}\{\tilde{\Phi}_{\text{total}}\}(\mathbf{x}) + ikYS\{\breve{\Phi}_{\text{total}}\}(\mathbf{x})$. Following the same logic used to derive this statement,[61] it can be asserted that the equivalent statement for spherical harmonic coefficients will also hold: $A_{m,n}^{\text{scat}} = ik\mathcal{D}_{m,n}^{\text{in}}\{\tilde{\Phi}_{\text{total}}\}(\mathbf{x}) - k^2 Y\mathcal{S}_{m,n}^{\text{in}}\{\breve{\Phi}_{\text{total}}\}(\mathbf{x})$.

3. Scene 9

Scene 9 is a standard interior admittance problem, so is simpler in its formulation than scene 3. The boundary integral equation to be solved for all $\mathbf{x} \in \Gamma$ is $[\mathcal{D} - \frac{1}{2}\mathcal{I} + ik\mathcal{S}Y(\mathbf{y})]\{\Phi_{\text{total}}\}(\mathbf{x}) = -\Phi_{\text{inc}}(\mathbf{x})$. The scattered pressure may be found by $\Phi_{\text{scat}}(\mathbf{x}) = [\mathcal{D} + ik\mathcal{S}Y(\mathbf{y})]\{\Phi_{\text{total}}\}(\mathbf{x})$, and equivalently the spherical harmonic coefficients may be found by $A_{m,n}^{\text{scat}} = ik[\mathcal{D}_{m,n}^{\text{in}} + ik\mathcal{S}_{m,n}^{\text{in}}Y(\mathbf{y})]\{\Phi_{\text{total}}\}(\mathbf{x}_r)$. Here, the only significant complication is that Y is position dependent so cannot be brought outside $\mathcal{S}$. There are, however, a finite number of materials present in the room, each of which has a uniform value of Y. It is therefore possible to partition the boundary into segments for which Y is uniform and may be brought outside $\mathcal{S}$; this leads to a blocked operator in BEM++. Φ_{scat} and $A_{m,n}^{\text{scat}}$ can be found by evaluating the contribution from each material segment separately and then summing the results.

E. Including measured material data

The boundary data provided in the GRAS database was given as third-octave random-incidence absorption coefficient α. This was converted to an admittance by assuming the material was purely resistive and locally reacting; the latter being a common assumption and the former being acceptable for materials that are fairly hard and reflective,[26] such as the MDF in scene 3. Using this, and applying the "55 degree rule,"[62] the specific admittance can be found by $Y = \cos(55°)[1 - \sqrt{1 - \alpha}]/[1 + \sqrt{1 - \alpha}]$. This was then interpolated to the required simulation frequencies using a spline fit assuming α to be constant below the lowest band provided (20 Hz).

This approach was, however, not considered adequate for some of the materials present in scene 9. In particular, it was expected that some materials, e.g., the glazing and door, would exhibit reactive behavior that gave significant losses at low frequencies, and that this was likely to be largely missing from measured material data because it would be non-locally reacting. Attempts were therefore made to fabricate plausible low-frequency resonant damping effects in their place. For example, the fundamental glazing resonant modal frequency was estimated from a plausible pane size, and the frequency of the coincidence dip visible in the measured data. This was implemented by fitting a mass-spring-damper model and the

admittance data produced was combined with that from the measured data using the non-linear crossover technique from Aretz et al.[24] Missing data for the door was drawn from standard tables[62] and proprietary field data, then embellished in a similar way. The result was a more plausible room response at low frequencies. For full details of the approaches applied see the code included in the supplementary material.[63]

IV. RESULTS

In this section, results are presented to validate the proposed approach. First, the new process in Eq. (8) that encodes the pressure around a receiver will be verified. Then the results for the three case studies are validated against measurement, including BRIRs for scene 3.

A. Verification of sound field encoding process

Here, the objective is to verify the accuracy of $A_{n,m}^{\text{inc}}$ and $A_{n,m}^{\text{scat}}$ coefficients. Evaluating a metric on these coefficients would be the ideal way of quantifying this, but this is complicated by the fact that all other known methods of obtaining them have limitations too.[14–16] So here instead, the field has been decoded at a set of points in the domain and the values obtained compared to values of Φ_{inc} and Φ_{scat} computed directly by Eqs. (2) and (7), respectively. 2427 evaluation points were used, arranged quasi-randomly within a sphere of diameter $\min(1, \lambda)$ centered on $\mathbf{x}_r$. The l^2 norm of the residual was computed and then normalized by the l^2 norm of the "correct" field; both were windowed with a Hanning window centered on $\mathbf{x}_r$. The mean and standard deviation were then computed, averaging with respect to frequency and over loudspeaker position for scene 9, to obtain the trends in Fig. 3; error bars indicate $\pm$ one standard deviation. Note that only frequencies above 343 Hz were included in these statistics; below that, the simulated array became smaller w.r.t. λ so the accuracy computed by the metric was unrepresentatively good. More detailed results plotted versus frequency are included in the supplementary material.[63]

The residual for Φ_{inc} is shown for scene 3 only since the trends for scenes 1 and 9 were identical. This continues to reduce over the full range of O_r investigated; what is seen here is just the effect of adding extra terms to the series in Eq. (2), and it appears the $A_{n,m}^{\text{inc}}$ coefficients are calculated
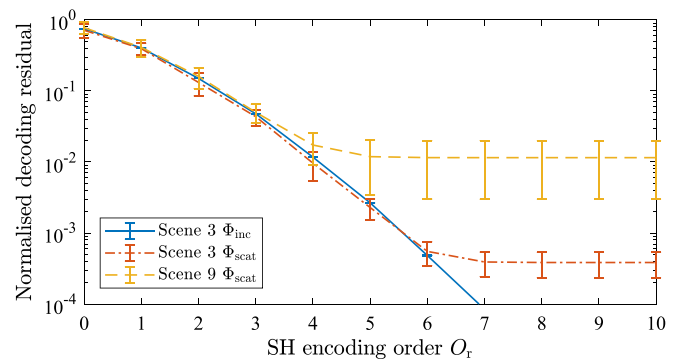

FIG. 3. (Color online) Error trends for encoding and decoding of the pressure field around a receiver, versus maximum spherical harmonic order used O_r.
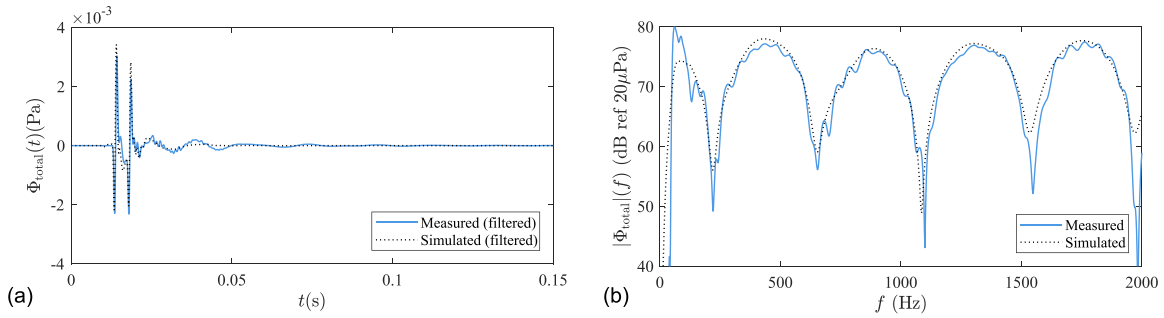
FIG. 4. (Color online) RIRs for Scene 1H: (a) versus time (low-pass filtered), (b) versus frequency.

accurately for all of these. The standard deviation over frequency is extremely small too.

In contrast, the residuals for Φ_{scat} converge up to some value and then stop; here, decoding accuracy has been limited by error in the $A_{n,m}^{scat}$ coefficients, indicating the accuracy of the encoding process; this is around 0.03% for scene 3 and around 1% for scene 9. These are average values, however, and the error bars indicate that quite a significant variation occurs; the maximum residual post-convergence was 0.1% for scene 3 and 7% for scene 9. It is clear from this that the Φ_{scat} encoding process in Eq. (8) works quite well for scene 3 but rather less so for scene 9. The reason for this requires further investigation; one possibility is that it occurs because of the sudden change in boundary condition between adjacent materials. The accuracy achieved may still be sufficient for auralization purposes, however; note that the error metric used here is rather stricter than the 'within x dB' criteria that is often used, and it includes phase error.

B. Scene 1H

The solution of scene 1H did not involve BEM, so it is included as a means of validating the source directivity model in Eq. (1) and encoding process described in Sec. III B. Measured and simulated results are compared in Fig. 4, which is plotted for loudspeaker position 3 from the database (height 2.6 m angled 60° down) and microphone position 1 (height 1.52 m, 4.1 m from loudspeaker horizontally). In both cases, it is seen that agreement is extremely good. Figure 4(a), in which the measured results have been low-pass filtered to match the processing applied to the simulations, shows correct arrival times and polarity, with onset amplitude well matched. Later, the measured result includes a low-frequency oscillation that is absent from the simulated data.

Frequency spectrum agreement in Fig. 4(b) is also good, with measured and simulation within 3 dB except at some notches and around 50–100 Hz, where there is a boost in the measured data; this is likely related to the low-frequency oscillation seen in the measured data in Fig. 4(a).

C. Scene 3

Scene 3 is an attractive verification case because its simulation involves BEM but it generated a sparse, physically insightful reflection pattern for both RIR and BRIRs. Figure 5 shows the RIRs. Again, very good agreement can be seen between simulation and measurement, possibly better than for scene 1H in fact. The onset times, phases and amplitudes of the pulses in Fig. 5(a) are all well captured. In Fig. 5(b), the interference pattern resulting from the flutter echo is very well matched up to the crossover frequency 1 kHz, after which the simulation begins to deviate from the measured result, perhaps because the BEM mesh is no longer being refined. The results from 1.5 to 2 kHz are hidden from the plot so the lower frequencies can be seen more clearly.

The BRIRs are plotted in Fig. 6. In the time-domain results in Figs. 6(a) and 6(b), an instantaneous SPL scale in dB has been used, so that the decay and relative amplitude of the pulse can be see for longer. Again, the pulse times and amplitudes are well-matched and the pattern of which channel is louder matches and makes physical sense. Reflections 1, 4, 7, 10, and 13 are louder in the right ear that faces the loudspeaker, being the original incident wave and its subsequent reflection back around the system, while others are similarly loud or louder in the left ear. The frequency domain match is less good than was seen in Fig. 5(b); the results can be seen to track each other up to 500 Hz at least. Deviations above that could occur because the simulation
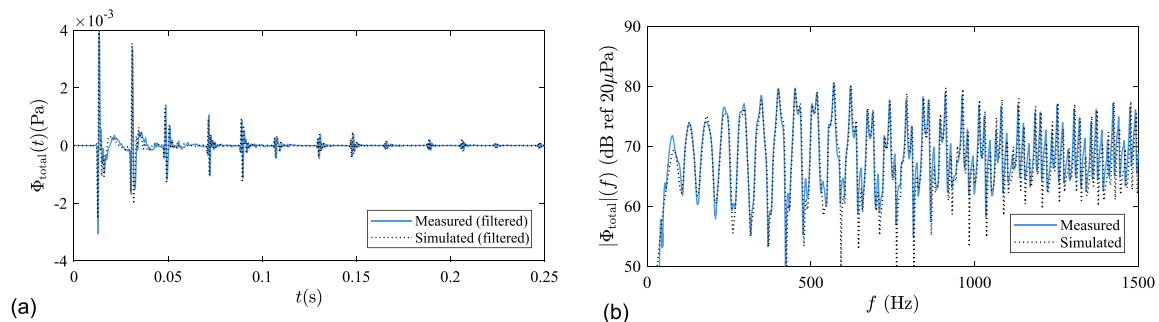

FIG. 5. (Color online) RIRs for Scene 3: (a) versus time (low-pass filtered), (b) versus frequency.
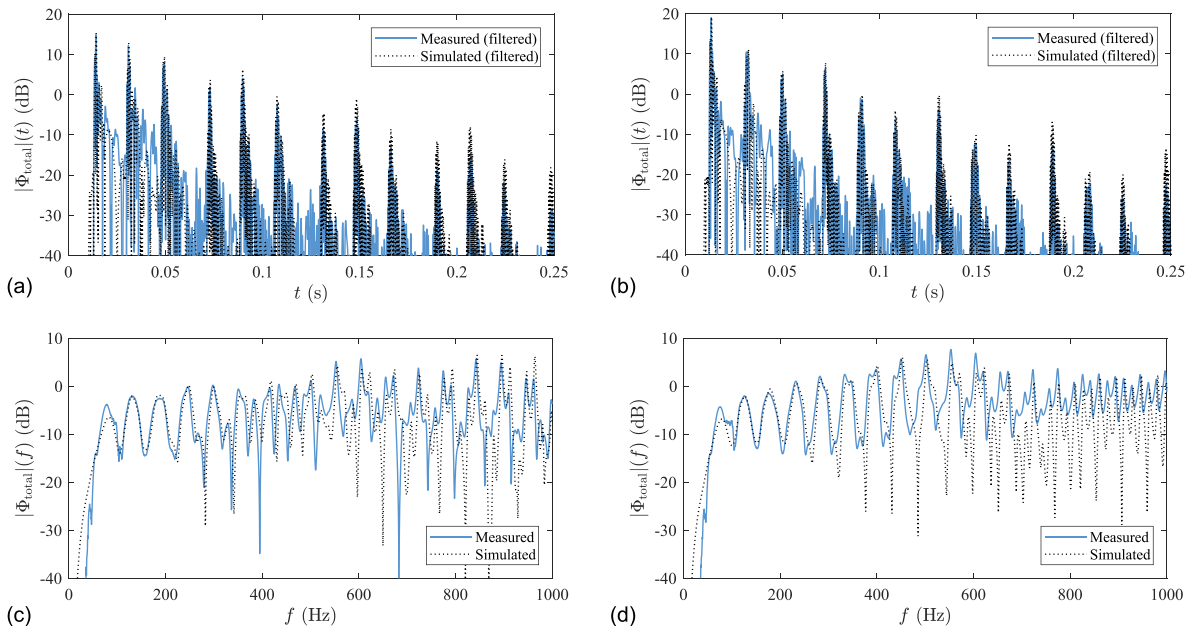
FIG. 6. (Color online) BRIRs for Scene 3: (a) left ear versus time (low-pass filtered), (b) right ear versus time (low-pass filtered), (c) left ear versus frequency, (d) right ear versus frequency.

process has combined datasets that have been measured at different times under different conditions. It should be mentioned that Fig. 6 hides a detail that the simulated and measured BRIR were negated with respect to one another. Noting, however, that the RIRs in Fig. 5 matched in sign and that the encoding and decoding was verified in Fig. 3 and Fig. 2(b), it seems most likely that this originates from the HRTF dataset. Auralizations for this scene are included in the supplementary material that accompanies this article.[63]

D. Scene 9

Scene 9 was included as a more realistic application of the simulation framework. Accuracy is, however, expected to be much poorer than for the other two scenes because: (i) as a closed geometry, the modal and reverberant damping is controlled entirely by boundary absorption mechanisms; (ii) these mechanisms were quite crudely quantified, as discussed in Secs. I B and III E. Results are shown for source position 2, coordinates [0.12, 2.88, 0.72 m], and microphone position 2, coordinates [0.44, −0.15, 0.12 m]; the origin of the coordinate system is roughly the center of the room at floor height.

The RIR is shown in the time-domain in Fig. 7(a) using an instantaneous SPL scale in dB. It was not expected that the fine detail would match and it does not, but it can be seen

that the general SPL and the decay rate match well. Figure 7(b) shows the same data in the frequency domain, displaying 0–250 Hz since the modal density above this means no discernable features are observable. The general SPL trend is captured quite well up to 170 Hz, with matches between individual modal frequencies identifiable. That some of these peaks match quite well in SPL and bandwidth is impressive, since this is largely dictated by the boundary absorption data, which was heavily extrapolated. BRIR results are not shown since little detail can be discerned graphically, but auralizations for this scene are included in the supplementary material that accompanies this article.[63]

E. Computation times

Computing times for scenes 3 and 9 are shown in Figs. 8(a) and 8(b), respectively. Here, the main observation is that all trends scale with #DOF (plotted against the right-hand axis). This is expected for the "Setup RHS" and the "Receiver encoding" tasks, being processes 3 and 5 in Fig. 1, respectively, but not for "Setup LHS" and "Solve," together being process 4 in Fig. 1. In a conventional BEM, these would scale with #DOF², so it is clear that the ACA solver has achieved significant computational savings. In contrast, "Setup RHS" and "Receiver encoding" would normally be
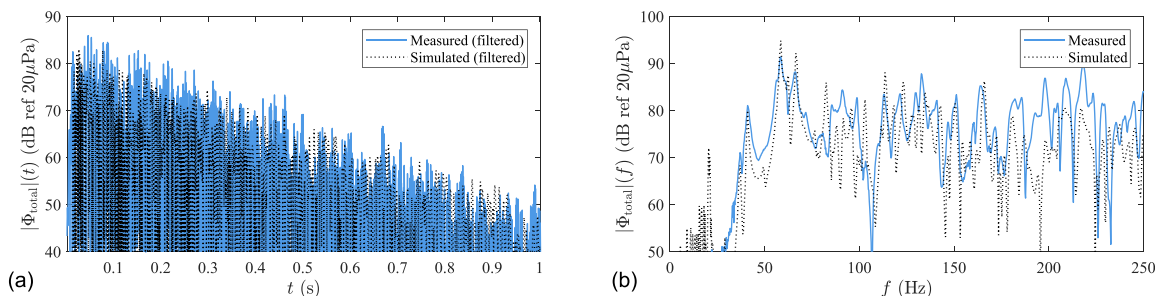

FIG. 7. (Color online) RIRs for Scene 9: (a) versus time (low-pass filtered), (b) versus frequency.

J. Acoust. Soc. Am. **145** (4), April 2019
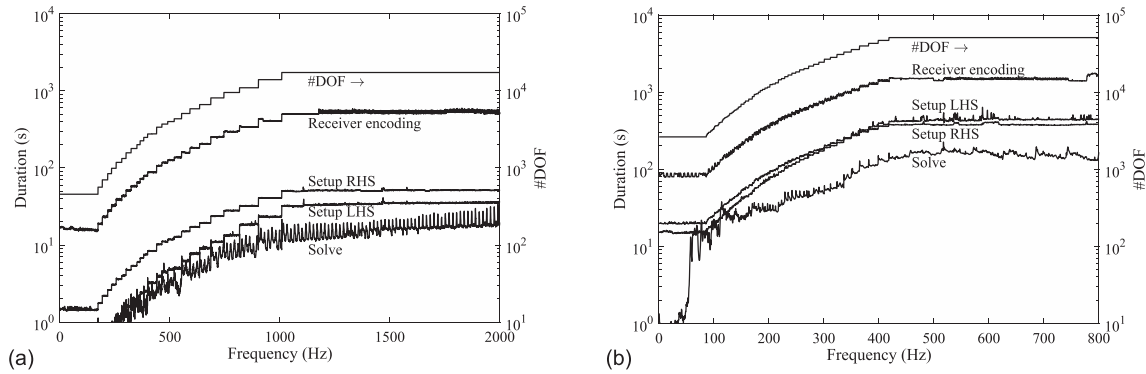
Hargreaves *et al.* 2635

FIG. 8. Computation times for: (a) scene 3, (b) scene 9.

expected to be by far the quickest steps but, due to the afore-mentioned implementation in as interpreted Python functions, they are much slower than the optimized, compiled core of the library. This should not however be taken as representative of the new source and receiver mapping techniques described herein; it is merely due to implementation compromises and an efficient, compiled implementation of them could be readily achieved.

V. CONCLUSIONS AND FURTHER WORK

This paper has proposed a framework for low-frequency room acoustic simulation, echoing similar frameworks that have been proposed for geometrical acoustics models at high-frequencies. A key component of this was the mapping proposed by Hargreaves and Lam[18] to encode boundary data to spherical harmonic descriptions of the pressure field around a receiver, and verification data, results, and auralizations using that are provided herein. The full simulation chain was validated using three case studies drawn from the GRAS database, one of which was hemi-anechoic, one fully-anechoic, and one a real room. The simulations were seen to match measurement well for the hemi-anechoic and fully-anechoic cases, but less so for the room; this was expected since standardized means of quantifying boundary material data are not sufficient for the simulation algorithms brought to bear. Clearly, this latter aspect is a limiting factor in the simulation chain, and one that must be addressed if room acoustic simulations are to move from being plausible to reliably physically accurate.

In terms of future work, it is clear that the current implementation of the new source and receiver mappings is extremely inefficient, and an optimized compiled version would be required for serious usage. It is also clear that repeated use of a frequency-domain BEM code followed by IFFT is not an efficient way of generating an impulse response; convolution quadrature[46] or multifrequency[49,50] approaches would be far more efficient. More research is required to set bounds on the accuracy of the new pressure field encoding process, since this was seen to vary with the problem modelled. Finally, these simulations have shown that numerical models can closely match reality when the input data is of a good quality but that they deviate when it is not. Hence improved techniques to characterize materials, ideally *in situ*, are required.

[1]M. Kleiner, B.-I. Dalenbäck, and P. Svensson, "Auralization—An overview," J. Audio Eng. Soc. **41**, 861–875 (1993).

[2]N. Xiang and J. Blauert, "Binaural scale modelling for auralisation and prediction of acoustics in auditoria," Appl. Acoust. **38**, 267–290 (1993).

[3]L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," J. Acoust. Soc. Am. **138**, 708–730 (2015).

[4]E. Granier, M. Kleiner, B.-I. Dalenbäck, and P. Svensson, "Experimental auralization of car audio installations," J. Audio Eng. Soc. **44**, 835–849 (1996).

[5]M. Vorländer, "Computer simulations in room acoustics: Concepts and uncertainties," J. Acoust. Soc. Am. **133**, 1203–1213 (2013).

[6]T. Sakuma, S. Sakamoto, and T. Otsuru, *Computational Simulation in Architectural and Environmental Acoustics* (Springer Japan, Tokyo, 2014).

[7]M. Aretz and M. Vorländer, "Combined wave and ray based room acoustic simulations of audio systems in car passenger compartments, Part II: Comparison of simulations and measurements," Appl. Acoust. **76**, 52–65 (2014).

[8]L. A. T. Jiménez, S. R. Sánchez, J. O. Villegas, and A. L. Ciro, "Subjective assessment of auralizations created by means of geometrical acoustics and finite element numerical methods" in *Proceedings of the International Congress on Sound and Vibration*, London, UK (23–27, 2017).

[9]J. E. Summers, K. Takahashi, Y. Shimizu, and T. Yamakawa, "Assessing the accuracy of auralizations computed using a hybrid geometrical-acoustics and wave-acoustics method," J. Acoust. Soc. Am. **115**, 2514 (2004).

[10]M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," J. Audio Eng. Soc. **53**, 1004–1025 (2005).

[11]J. Ahrens, "The single-layer potential approach applied to sound field synthesis including cases of non-enclosing distributions of secondary sources," Ph.D. thesis, TU Berlin, Berlin, 2010.

[12]S. Pelzer, M. Pollow, and M. Vorländer, "Continuous and exchangeable directivity patterns in room acoustic simulation," in *Proceedings of DAGA 38*, Darmstadt, Germany (March 19–22, 2012).

[13]P. Samarasinghe, T. Abhayapala, M. Poletti, and T. Betlehem, "An efficient parameterization of the room transfer function," IEEE/ACM Trans. Audio Speech Lang. Process. **23**, 2217–2227 (2015).

[14]A. Southern, D. T. Murphy, and L. Savioja, "Spatial encoding of finite difference time domain acoustic models for auralization," IEEE Trans. Audio. Speech. Lang. Process. **20**, 2420–2432 (2012).

[15]D. Gómez, J. Astley, and F. Fazi, "Low frequency interactive auralization based on a plane wave expansion," Appl. Sci. **7**, 558 (2017).

[16]J. Sheaffer, M. van Walstijn, B. Rafaely, and K. Kowalczyk, "Binaural reproduction of finite difference simulations using spherical array processing," IEEE/ACM Trans. Audio Speech Lang. Process. **23**, 2125–2135 (2015).

[17]R. Mehra, L. Antani, S. Kim, and D. Manocha, "Source and listener directivity for interactive wave-based sound propagation," IEEE Trans. Vis. Comput. Graph. **20**, 495–503 (2014).

[18]J. A. Hargreaves and Y. W. Lam, "An energy interpretation of the kirchhoff-helmholtz boundary integral equation and its application to sound field synthesis," Acta Acust. united Ac. **100**, 912–920 (2014).

[19]E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis," in *Proceedings of the AES Convention*, Amsterdam, the Netherlands (May 12–15, 2001).

[20]I. Harari and T. J. R. Hughes, "A cost comparison of boundary element and finite element methods for problems of time-harmonic acoustics," Comput. Methods Appl. Mech. Eng. **97**, 77–102 (1992).

[21]T. J. Cox, "Prediction and evaluation of the scattering from quadratic residue diffusers," J. Acoust. Soc. Am. **95**, 297–305 (1994).

[22]N. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*, 1st ed. (Elsevier Science, Amsterdam, the Netherlands, 2005).

[23]T. Betcke, E. van't Wout, and P. Gélat, "Computationally efficient boundary element methods for high-frequency helmholtz problems in unbounded domains," in *Modern Solvers for Helmholtz Problems* (Birkhäuser, Germany, 2017), pp. 215–243.

[24]M. Aretz, R. Nöthen, M. Vorländer, and D. Schröder, "Combined broadband impulse responses using fem and hybrid ray-based methods," in *EAA Symposium on Auralization*, Espoo, Finalnd (June 15–17, 2009).

[25]M. Aretz, "Specification of realistic boundary conditions for the FE simulation of low frequency sound fields in recording studios," Acta Acust. united Ac. **95**, 874–882 (2009).

[26]M. Aretz, "Combined wave and ray based room acoustic simulations of small rooms," Ph.D. thesis, RWTH Aachen, Aachen, Germany, 2012.

[27]M. Aretz and M. Vorländer, "Combined wave and ray based room acoustic simulations of audio systems in car passenger compartments, Part I: Boundary and source data," Appl. Acoust. **76**, 82–99 (2014).

[28]A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja, "Room impulse response synthesis and validation using a hybrid acoustic model," IEEE Trans. Audio. Speech. Lang. Process. **21**, 1940–1952 (2013).

[29]F. Brinkmann, L. Aspöck, D. Ackermann, R. Opdam, M. Vorlaender, and S. Weinzierl, "First international round robin on auralization: Concept and database," J. Acoust. Soc. Am. **141**, 3856–3856 (2017).

[30]L. Aspöck, F. Brinkmann, D. Ackerman, S. Weinzierl, and M. Vorländer, "GRAS—Ground Truth for Room Acoustical Simulation," https://dx.doi.org/10.14279/depositonce-6726 (Last viewed March 19, 2018).

[31]S. Bilbao and B. Hamilton, "Directional source modeling in wave-based room acoustics simulation," in *Proceedings of the 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY (October 15–18, 2017).

[32]S. Sakamoto and R. Takahashi, "Directional sound source modeling by using spherical harmonic functions for finite-difference time-domain analysis," Proc. Meet. Acoust. **19**, 015129 (2013).

[33]F. Georgiou and M. Hornikx, "Incorporating directivity in the Fourier pseudospectral time-domain method using spherical harmonics," J. Acoust. Soc. Am. **140**, 855–865 (2016).

[34]S. Marburg, "Boundary element method for time-harmonic acoustic problems," in *Computational Acoustics* (Springer International Publishing, New York, 2018), pp. 69–158.

[35]E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, San Diego, CA, 1999).

[36]G. Weinreich, "Method for measuring acoustic radiation fields," J. Acoust. Soc. Am. **68**, 404–411 (1980).

[37]M. J. Evans, J. A. S. Angus, and A. I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," J. Acoust. Soc. Am. **104**, 2400–2411 (1998).

[38]F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders, "NIST digital library of mathematical functions," http://dlmf.nist.gov/ (Last viewed January 24, 2018).

[39]M. Pollow, K.-V. Nguyen, O. Warusfel, T. Carpentier, M. Müller-Trapet, M. Vorländer, and M. Noisternig, "Calculation of head-related transfer functions for arbitrary field points using spherical harmonics decomposition," Acta Acust. united Ac. **98**, 72–82 (2012).

[40]M. Melon, C. Langrenne, P. Herzog, and A. Garcia, "Evaluation of a method for the measurement of subwoofers in usual rooms," J. Acoust. Soc. Am. **127**, 256–263 (2010).

[41]Klippel GmbH, "Near field scanner system," https://www.klippel.de/products/rd-system/modules/nfs-near-field-scanner.html (Last viewed August 1, 2018).

[42]J. A. Hargreaves and Y. W. Lam, "Corrigendum to 'An energy interpretation of the Kirchhoff-Helmholtz boundary integral equation and its application to sound field synthesis,'" Acta Acust. united Ac. **104**(6), 1134 (2018).

[43]R. Martinez, "The thin-shape breakdown (TSB) of the Helmholtz integral equation," J. Acoust. Soc. Am. **90**, 2728–2738 (1991).

[44]Y. Sakurai and K. Ishida, "Multiple reflections between rigid plane panels," J. Acoust. Soc. Jpn. **3**, 183–190 (1982).

[45]E. Brandão, A. Lenzi, and S. Paul, "A review of the *in situ* impedance and sound absorption measurement techniques," Acta Acust. united Ac. **101**, 443–463 (2015).

[46]S. A. Sauter and M. Schanz, "Convolution quadrature for the wave equation with impedance boundary conditions," J. Comput. Phys. **334**, 442–459 (2017).

[47]J. A. Hargreaves and T. J. Cox, "A transient boundary element method model of Schroeder diffuser scattering using well mouth impedance.," J. Acoust. Soc. Am. **124**, 2942–2951 (2008).

[48]T. Okuzono, T. Yoshida, K. Sakagami, and T. Otsuru, "An explicit time-domain finite element method for room acoustics simulations: Comparison of the performance with implicit methods," Appl. Acoust. **104**, 76–84 (2016).

[49]O. von Estorff, S. Rjasanow, M. Stolper, and O. Zaleski, "Two efficient methods for a multifrequency solution of the Helmholtz equation," Comput. Vis. Sci. **8**, 159–167 (2005).

[50]J.-P. Coyette, C. Lecomte, J.-L. Migeot, J. Blanche, M. Rochette, and G. Mirkovic, "Calculation of vibro-acoustic frequency response functions using a single frequency boundary element solution and a Padé expansion," Acta Acust. united with Ac. **85**, 371–377 (1999).

[51]F. Brinkmann, A. Lindau, and S. Weinzierl, "A high resolution head-related transfer function database including different orientations of head above the torso," in *Proceedings of the AIA-DAGA*, Merano, Italy (March 18–21, 2013).

[52]F. Brinkmann, A. Lindau, M. Müller-Trapet, M. Vorländer, and S. Weinzierl, "Cross-validation of measured and modeled head-related transfer functions," in *Proceedings of DAGA 2015*, Nürnberg, Germany (March 16–19, 2015).

[53]W. Śmigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger, "Solving boundary integral problems with BEM++," ACM Trans. Math. Softw. **41**, 1–40 (2015).

[54]S. Arridge, T. Betcke, M. Scroggs, and E. van't Wout, "bempp services," www.bempp.org, (Last viewed December 14, 2017).

[55]E. Jones, E. Oliphant, and P. Peterson, "SciPy: Open source scientific tools for Python," http://www.scipy.org/ (Last viewed December 14, 2017).

[56]C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities," Int. J. Numer. Methods Eng. **79**, 1309–1331 (2009).

[57]C. Geuzaine and J.-F. Remacle, "Gmsh: A three-dimensional finite element mesh generator with built-in pre- and post-processing facilities," http://gmsh.info/ (Last viewed December 14, 2017).

[58]Y. Yasuda and T. Sakuma, "Boundary element method," in *Computational Simulation in Architectural and Environmental Acoustics* (Springer Japan, Tokyo, 2014), pp. 79–115.

[59]D. P. Hewett and U. P. Svensson, "The diffracted field and its gradient near the edge of a thin screen," J. Acoust. Soc. Am. **134**, 4303–4306 (2013).

[60]G. Krishnasamy, L. W. Schmerr, T. J. Rudolphi, and F. J. Rizzo, "Hypersingular boundary integral equations: Some applications in acoustic and elastic wave scattering," J. Appl. Mech. **57**, 404–414 (1990).

[61]T. Terai, "On calculation of sound fields around three dimensional objects by integral equation methods," J. Sound Vib. **69**, 71–100 (1980).

[62]T. J. Cox and P. D'Antonio, *Acoustic Absorbers and Diffusers: Theory, Design and Application*, 2nd ed. (Taylor & Francis, London, 2009).

[63]See supplementary material at https://doi.org/10.1121/1.5096171 for additional figures, code used to perform the simulations, and auralizations of the results.

J. Acoust. Soc. Am. **145** (4), April 2019

Hargreaves *et al.* 2637