7 - 12 SEPTEMBER 2014, KRAKÓW
FORUM ACUSTICUM

A Spherical Array Approach for Simulation of Binaural Impulse Responses using the Finite Difference Time Domain Method

Jonathan Sheaffer
Dept. of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel.

Maarten van Walstijn
School of Electronics, Electrical Engineering and Computer Science, Queen's University, Belfast, UK.

Boaz Rafaely
Dept. of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel.

Konrad Kowalczyk
Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany.

Summary
The finite difference time domain (FDTD) method has direct applications in musical instrument modeling, simulation of environmental acoustics, room acoustics and sound reproduction paradigms, all of which benefit from auralization. However, rendering binaural impulse responses from simulated data is not straightforward to accomplish as the calculated pressure at FDTD grid nodes does not contain any directional information. This paper addresses this issue by introducing a spherical array to capture sound pressure on a finite difference grid, and decomposing it into a plane-wave density function. Binaural impulse responses are then constructed in the spherical harmonics domain by combining the decomposed grid data with free field head-related transfer functions. The effects of designing a spherical array in a Cartesian grid are studied, and emphasis is given to the relationships between array sampling and the spatial and spectral design parameters of several finite-difference schemes.

PACS no. 43.55.Ka, 43.38.Ar, 43.55.Lb

1. Introduction

The Finite Difference Time Domain (FDTD) method is gaining popularity in room acoustics [1, 2], environmental acoustics [3] and musical acoustics [4]. Operating in the discrete time domain, FDTD is particularly suited for obtaining wideband solutions, and for solving problems involving time variability, such as moving sources and receivers. For many applications, it is common to subjectively evaluate modeled results by means of *auralization* [5]. This process is most trivially accomplished by modeling the impulse response of a system and convolving it with pre-recorded free-field stimuli in a post-processing stage. Such an auralization can be either monophonic, meaning that no spatial information is rendered, or binaural, in which case the components of the modeled soundfield are filtered with head-related transfer functions (HRTFs), which can be seen as frequency and direction dependent receivers. In geometrical methods [6], the spatial information of the soundfield is inherent in the computation, hence binaural filtering is straightforward to accomplish. In wave based methods, however, this information needs to be either explicitly modeled (for example by embedding a scattering object in the simulation) or extracted by means of array processing [7].

By approximating instantaneous intensity vectors from first order pressure gradients, Hacihabiboglu et al. [8] have suggested a method for modeling directional microphones in the closely related paradigm of digital waveguide modeling (DWM). While their results demonstrate accurate modeling of low-order directivity patterns, it is not clear how this method applies to higher-order spatial filtering, as required for binaural impulse response rendering. Southern et al. [9] have addressed the problem of spatially encoding a FDTD grid from the perspective of differential microphone arrays. By placing spaced receivers on grid nodes and applying the Blumlein difference technique, they suggested to encode the soundfield into

a *Higher-Order Ambisonics* representation. While in theory such encoded data could be transcoded into a variety of reproduction formats, the problem of directly obtaining binaural signals from the FDTD grid was not discussed. In addition, this approach - which has been demonstrated for orders up to 3 - has yet to be extended to a full 3D ambisonic encoding.

As a simplified solution for directly obtaining a binaural response, Murphy and Beeson [10] suggested placing a circular object between two spaced receivers in a 2D grid, and demonstrated that an accurate reproduction of Interaural Time Difference (ITD) cues can be achieved. Sheaffer et al. [11] extended this approach by embedding a fully featured laser-scan of a human head into a 3D finite difference grid. Their results indicated that although both ITD and Interaural Level Difference (ILD) cues can be effectively modeled, it is challenging to reproduce the fine structure of spectral cues with numerical schemes typically used for room acoustics simulation. In contrast, Xiao and Liu [12] as well as Mokhtari et al. [13], showed that an accurate representation of a full HRTF set can be modeled using specialized high-resolution FDTD schemes. Nonetheless, such schemes appear to be too computationally expensive to be applied to realistic room acoustics and environmental models.

The literature does not, so far, offer a robust solution for directly modeling binaural responses in a FDTD simulation. This paper addresses the problem of directional receiver modeling with specific applications to binaural reproduction, by introducing a volumetric spherical array which can be effectively quantized on a finite difference grid. The underlying theory of representing a binaural soundfield using spherical harmonics is presented (Section 2), followed by a discussion on realization of spherical arrays in FDTD (Section 3). To facilitate rendering of a broadband binaural response, a volumetric spherical array is suggested and studied from the perspectives of numerical dispersion and numerical robustness in Section 4. Lastly, numerical results are presented in Section 5, demonstrating the application of the proposed method to binaural simulation.

2. Rendering binaural responses using spherical harmonics

Consider a soundfield governed by the homogeneous acoustic wave equation,

$$\nabla^2 p(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} p(\mathbf{r},t) = 0, \tag{1}$$

where c is the speed of sound, ∇^2 is the Laplace operator and $p(\mathbf{r},t)$ is the field variable which is here assumed to be sound pressure. The wave equation can be described in Cartesian coordinates such that $\mathbf{r} \equiv (x,y,z) \in \mathbb{R}^3$, or in spherical coordinates in which case $\mathbf{r} \equiv (\theta,\phi,r) \in \mathbb{R}^2$, where θ is elevation, ϕ is azimuth and r is radial distance. While Cartesian coordinates are the natural choice for domain discretization in a FDTD model, working in spherical coordinates allows one to expand the soundfield in terms of spherical harmonics (SH) which is key to the method proposed in this paper. Let $p(k,r,\Omega) \equiv p(\mathbf{r},k)$ denote the representation of $p(\mathbf{r},t)$ in the frequency domain, where $\Omega \equiv (\theta,\phi)$ is the solid angle, $k = 2\pi f/c$ is the wavenumber and f is the frequency. Assuming that $p(k,r,\Omega)$ is square-integrable over Ω, then its spherical Fourier transform, denoted by $p_{nm}(k,r)$, and the corresponding inverse transform are given by [14]

$$p_{nm}(k,r) = \int_{\Omega\in S^2} p(k,r,\Omega) Y_n^{m^*}(\Omega) d\Omega, \tag{2}$$

$$p(k,r,\Omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} p_{nm}(k,r) Y_n^m(\Omega) \tag{3}$$

where $\int d\Omega = \int\int \sin\theta d\theta d\phi$, and the operator $(\cdot)^*$ denotes complex conjugation. The spherical harmonics function, $Y_n^m(\cdot)$, is given by [14]

$$Y_n^m(\Omega) = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \tag{4}$$

where n and m denote order and degree respectively, and $P_n^m(\cdot)$ is the associated Legendre function.

Consider now a soundfield composed of a continuum of plane waves whose amplitude density is given by the complex function $a(k,\Omega)$. When these plane waves impinge on an open sphere of radius r, the sound pressure at the surface of the sphere is given by [14]

$$\begin{aligned} p_{nm}(k,r) &= b_n(kr) \int_{\Omega\in S^2} a(k,\Omega) Y_n^{m^*}(\Omega) d\Omega \\ &= b_n(kr) a_{nm}(k), \end{aligned} \tag{5}$$

where $a_{nm}(k)$ is the spherical Fourier transform of $a(k,\Omega)$, $b_n(kr) = 4\pi i^n j_n(kr)$ is the radial function [14], $j_n(\cdot)$ is the n^{th} order spherical Bessel function and $i = \sqrt{-1}$. Following Rafaely [15], it can be shown that the soundfield can be decomposed into its plane wave components as follows:

$$a_{nm}(k) = \frac{p_{nm}(k,r)}{b_n(kr)}. \tag{6}$$

Now given an HRTF set, for example for the left ear, $H^l(k,\Omega)$, whose spherical Fourier transform is $H^l_{nm}(k)$, it is possible to render a transfer function, $p^l(k)$, which represents the filtering of the soundfield as well as the head, by [16]

$$\begin{aligned} p^l(k) &= \int_{\Omega\in S^2} a(k,\Omega) H^l(k,\Omega) d\Omega \\ &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \tilde{a}^*_{nm}(k) H^l_{nm}(k), \end{aligned} \tag{7}$$

where $\tilde{a}_{nm}(k) = (-1)^m a_n^{-m}(k)$ is the representation of $a^*(k,\Omega)$ in the SH domain. The transfer function for the right ear can be computed in a similar fashion with the right ear HRTF, $H^r(k,\Omega)$.

3. Realization of spherical arrays in finite difference simulation

To simulate wave propagation using the FDTD method, the sound field is discretized on a Cartesian grid such that

$$(x,y,z,t) \rightarrow [dX, fX, gX, uT], \tag{8}$$

where u and $[d,f,g]$ are the index positions in discrete time and space respectively, and X and T are the spatial and temporal sample periods. Correspondingly, the wave equation (1) can be modeled as [2]

$$[\delta_t^2 - \lambda^2(\delta_x^2+\delta_y^2+\delta_z^2+\dot{a}\delta_x^2\delta_y^2+ \dot{a}\delta_x^2\delta_z^2+\dot{a}\delta_y^2\delta_z^2+\dot{b}\delta_x^2\delta_y^2\delta_z^2)]p|_{d,f,g}^u = 0, \tag{9}$$

with the finite difference operators given by

$$\delta_t^2 p|_{d,f,g}^u \equiv p|_{d,f,g}^{u+1} - 2p|_{d,f,g}^u + p|_{d,f,g}^{u-1}, \tag{10}$$

$$\delta_x^2 p|_{d,f,g}^u \equiv p|_{d+1,f,g}^u - 2p|_{d,f,g}^u + p|_{d-1,f,g}^u, \tag{11}$$

$$\delta_y^2 p|_{d,f,g}^u \equiv p|_{d,f+1,g}^u - 2p|_{d,f,g}^u + p|_{d,f-1,g}^u, \tag{12}$$

$$\delta_z^2 p|_{d,f,g}^u \equiv p|_{d,f,g+1}^u - 2p|_{d,f,g}^u + p|_{d,f,g-1}^u. \tag{13}$$

The Courant number, $\lambda = cT/X$, and the free parameters $\dot{a}$ and $\dot{b}$, are chosen according to the desired *numerical scheme* of the FDTD model, see e.g. Table I in [2]. Consider now impulse responses recorded at Q grid nodes, arbitrarily distributed around some point on the grid, which is further referred to as the array's center position. The pressure at each of these receivers is transformed into the frequency domain, resulting in a frequency-dependent vector

$$\mathbf{p} = [p_1(k)\ ,\ p_2(k)\ ,\ \cdots\ ,\ p_Q(k)]^T, \tag{14}$$

where $(\cdot)^T$ denotes transposition. Each receiver has its own angle and radial distance with respect to the center of the array,

$$r_q = \sqrt{(d_q X)^2 + (f_q X)^2 + (g_q X)^2} \tag{15}$$

$$\theta_q = \text{acos}\frac{g_q X}{r_q} \tag{16}$$

$$\phi_q = \text{atan}\frac{f_q}{d_q}, \tag{17}$$

with $[d_q, f_q, g_q]$ denoting the grid index of the q^{th} receiver, respectively. By substituting Eqn. (6) into (3) and truncating the infinite series at some finite order, N, the pressure at each receiving node can be approximated by [17]

$$p(k, r_q, \Omega_q) \approx \sum_{n=0}^{N}\sum_{m=-n}^{n} a_{nm}(k) b_n(kr_q) Y_n^m(\Omega_q), \tag{18}$$

with $\Omega_q \equiv (\theta_q, \phi_q) \quad 1 \le q \le Q$, and the approximation becoming equality as $N \rightarrow \infty$ or if the soundfield is known to be order-limited at some finite order N. This relationship can be expressed in matrix form as follows:

$$\mathbf{p} = \mathbf{B}\mathbf{a_{nm}}, \tag{19}$$

where $\mathbf{a_{nm}}$ is a $(N+1)^2 \times 1$ vector representing the complex plane wave coefficients of the soundfield,

$$\mathbf{a_{nm}} = \left[a_{00}, a_{1(-1)}, a_{10}, a_{11}, \ldots, a_{NN}\right]^T, \tag{20}$$

and the $Q \times (N+1)^2$ matrix $\mathbf{B}$ is given by

$$\mathbf{B}^T = \begin{bmatrix} b_0(kr_1)Y_0^0(\Omega_1) & \cdots & b_0(kr_Q)Y_0^0(\Omega_Q) \\ b_1(kr_1)Y_1^{-1}(\Omega_1) & \cdots & b_1(kr_Q)Y_1^{-1}(\Omega_Q) \\ \vdots & \ddots & \vdots \\ b_N(kr_1)Y_N^N(\Omega_1) & \cdots & b_N(kr_Q)Y_N^N(\Omega_Q) \end{bmatrix} \tag{21}$$

where $(\cdot)^T$ denotes transposition. If one samples the sound field at a sufficient number of receiving nodes, then Eqn. (19) represents an overdetermined system of linear equations with a solution in a least-squares sense, in which case plane wave decomposition can be accomplished by [17]

$$\mathbf{a_{nm}} = \mathbf{B}^\dagger \mathbf{p}, \tag{22}$$

where $\mathbf{B}^\dagger = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ is the Moore-Penrose pseudoinverse of $\mathbf{B}$, and $(\cdot)^H$ denotes conjugate transpose. To obtain the left-ear transfer function, one may use the coefficients contained in $\mathbf{a_{nm}}$ to compute Eqn. (7), which is here truncated at the same order N.

4. Array design considerations

In the array processing literature, a general distinction is made between the *spatial sampling scheme* (not to be confused with the *numerical scheme* of the FDTD model), and the *array configuration*. The *sampling scheme* determines the number of receivers which are employed in sampling the sound field, as well as their positions on the surface of a sphere. Typical sampling schemes include equal-angle, Gaussian and (nearly) uniform, each having their own merits and drawbacks [18]. When designing a microphone

array according to a given scheme, it is important that transducers are placed exactly at their designated positions, as mispositioned receivers introduce processing errors resulting from phase mismatches. In an FDTD model, pressure data is only directly available at grid nodes, which complicates conforming to a specific sampling scheme. For example, choosing the receiver positions as nearest grid points would entail spatial quantization on the angular as well as the radial dimensions.

The *array configuration* refers to the type of sensors and properties of the array surfaces. In a real-world array design, the most straightforward approach is to employ an array of pressure transducers (having omni-directional pickup patterns) on the surface of a (virtual) sphere, also known as an *open sphere* design. This configuration features the radial function, $b_n(kr) = 4\pi i^n j_n(kr)$, as defined in Section 2. The most prominent shortcoming of this design is that the radial function is zeroed at different combinations of kr, which affects processing stability due to the division in Eqn. (6). One strategy commonly employed to avoid this problem, is to measure the pressure over the surface of a rigid sphere, in which case zeros in the spherical Bessel function do not result in zeroing of the radial function [18]. In context of an FDTD model, a rigid sphere array configuration requires embedding a spherical object in the grid, which for any non-conformal boundary model would result in further modeling errors due to the staircase representation of the boundary [19]. Another alternative is to use an open-sphere with cardioid microphones [20], however, within an FDTD model, it is not clear how one might employ such receivers in directions other than the three principal axes for which particle velocity can be trivially computed. Furthermore, both open sphere and rigid sphere configurations have a constrained bandwidth as they are numerically unstable at low frequencies, in which case the magnitude of the radial function diminishes as kr approaches zero.

In order to cover a wider range of frequencies, a dual open-sphere design can be employed. In this configuration, pressure is sampled at two spheres having some optimal radius ratio [21], which in practice requires twice as many microphones to accomplish. To minimize the number of required microphones in such open sphere designs, spherical shell [17] and spindle-torus [22] arrays have been suggested. In a numerical model, however, the number of receiving positions is only constrained by the available system memory, which relaxes the limitation on the total array size. As the goal in this study is to design a robust receiver array, which is scalable to various grid parameters and numerical schemes, it does not seem feasible to conform to a single-sphere sampling scheme. Rather than that, a volumetric approach is taken here, in which all nodes within a predefined volume act as receivers, as illustrated in Figure 1.

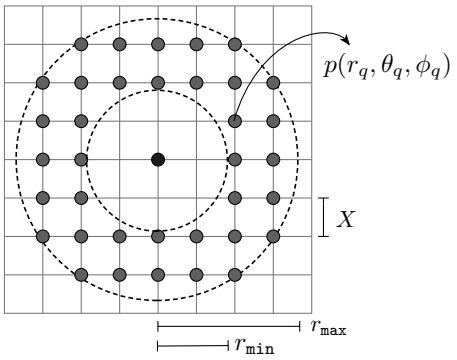

Figure 1. A 2D section of a volumetric shell array. $r_{\max}$ and $r_{\min}$ are the shell's outer and inner radii respectively, and the FDTD spatial period, X, controls the density of the receivers within the array.

While the nodal density on the grid is *a priori* dictated by the design parameters of the FDTD scheme, the total number of receivers is also dependent on the volume of the array. Thus, it is useful to quantify the effect of the parameters X, $r_{\min}$ and $r_{\max}$ (see Figure 1) on the overall performance of the array. These considerations are further discussed in the remainder of this section.

4.1. Numerical robustness

As previously noted, the PWD process described in Eqn.(6) relies on division of pressure coefficients by a radial function, whose magnitude in this case is proportional to the magnitude of the spherical Bessel function. As k or r approach zero (and therefore $|b_{n>0}(kr)| \to 0$), the PWD process has a poor numerical robustness, meaning that any input errors present in $p(k, r, \Omega)$ will be significantly amplified. In an FDTD model, such errors could be attributed to numerical dispersion as well as to the finite computation precision.

Following Rafaely [17], we evaluate numerical robustness by computing the *condition number* of the matrix $\mathbf{B}$, which is defined by

$$\kappa(\mathbf{B}) = \frac{\overline{\sigma}(\mathbf{B})}{\underline{\sigma}(\mathbf{B})}, \tag{23}$$

where $\overline{\sigma}(\mathbf{B})$ and $\underline{\sigma}(\mathbf{B})$ are the maximum and minimum singular values of $\mathbf{B}$, respectively. Figure 2 shows the condition number for the array designs detailed in Table I. All configurations are computed such that the positions of spatial samples are quantized to nearest grid nodes, except for the rigid sphere which serves as a reference design. Non-volumetric arrays employ a nearly-uniform sampling scheme tailored for $N = 12$.

As one would expect, the open sphere design (OS-Q) is the least robust, with the condition number peaking at frequencies corresponding to zeros in the spherical Bessel function. The irregular distribution of

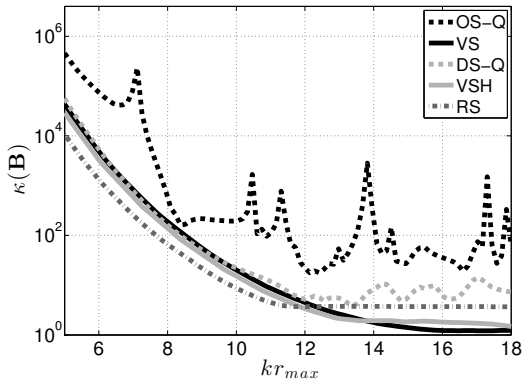

Figure 2. Condition number of a matrix $\mathbf{B}$ as function of $kr_{\mathtt{max}}$ shown for a number array configurations designed on an FDTD grid of sample period $X = 10$mm. In all cases, the matrix $\mathbf{B}$ is computed with $N = 12$.

Table I. Description of modeled arrays. N_p refers to the total number of samples used.

ID	**Type**	$r_{\mathtt{max}}$	$r_{\mathtt{min}}$	N_p
OS-Q	Open, quantized	0.1	N/A	230
VS	Volumetric	0.1	0	4169
VSH	Volumetric (shell)	0.1	0.075	2066
DS-Q	Dual, quantized	0.1	0.075	230
RS	Rigid, non-quantized	0.1	N/A	230

these frequencies may be attributed to quantization of receiver position to nearest grid nodes. When comparing the volumetric arrays (VS, VSH) to the rigid sphere (RS) reference, it can be seen that the RS has better performance at lower frequencies and VS/VSH have lower condition numbers at high frequencies, with the transition frequency being $kr_{\mathtt{max}} = N$.

4.2. Dispersion errors

One form of input error that might be amplified by the PWD process is numerical dispersion, which is both direction and frequency dependent. For the general family of compact explicit FDTD schemes used in this study, the dispersion relation is given by [2]

$$\sin^2(\pi f T) = \lambda^2[(s_x + s_y + s_z) - 4\dot{a}(s_x s_y + s_x s_z + s_y s_z) + 16\dot{b} s_x s_y s_z], \quad (24)$$

with

$$s_x = \sin^2\left(\frac{\tilde{k}X\cos\phi\cos\theta}{2}\right) \quad (25)$$
$$s_y = \sin^2\left(\frac{\tilde{k}X\sin\phi\cos\theta}{2}\right)$$
$$s_z = \sin^2\left(\frac{\tilde{k}X\sin\theta}{2}\right),$$

where

$$\tilde{k} = \frac{2\pi f}{\tilde{c}(\theta,\phi)}, \quad (26)$$

is the numerical wavenumber, and $\tilde{c}$ is the numerical wave propagation speed. Note that here (θ, ϕ) refer to the wave propagation direction. It is also worthwhile noting that if a (near) isotropic FDTD scheme is used, then one normally employs frequency warping to reverse the effects of dispersion [23]. Accordingly, the following numerical analysis for the IISO2 scheme employs a post-warped numerical wavenumber in the place of $\tilde{k}$, such that

$$\tilde{k}_w = \tilde{k}\frac{c_w}{c}, \quad (27)$$

where $\tilde{k}_w$ is the post-warped wavenumber and c_w is the numerical wave velocity at one of the extremes of the FDTD scheme.

In an FDTD simulation, numerical dispersion generally results in field errors which are amplified in the PWD process. Apart from that, the decomposition process itself assumes an ideal wavenumber, which further affects the results. To better understand this effect, we consider a single unit-amplitude plane wave arriving from the direction $\Omega_0 \equiv (\theta_0, \phi_0)$ and impinging on the surface of a sphere of radius r. In the SH domain, the pressure at the sphere's surface is given by [14]

$$p_{nm}(k) = b_n(kr)Y_n^{m^*}(\Omega_0). \quad (28)$$

The PWD process can be seen as a maximum-directivity beamformer whose look direction, $\Omega_L \equiv (\theta_L, \phi_L)$, is tuned to the direction of the arriving plane wave [15]. Accordingly, the output of the array can be described as

$$\begin{aligned} y(k,\Omega_L) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} a_{nm}Y_n^m(\Omega_L) \\ &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{b_n(kr)Y_n^{m^*}(\Omega_0)}{b_n(kr)}Y_n^m(\Omega_L) \\ &= \delta(\cos\theta_L - \cos\theta_0)\delta(\phi_L - \phi_0). \end{aligned} \quad (29)$$

As expected, a single plane wave propagating in an ideal medium reduces to a spatial delta function at the array output. Consider now a similar situation, only where the plane wave propagates in a dispersive medium in which the numerical wavenumber and corresponding wave speed are direction dependent. This results in a similar expression, only with $\tilde{k}$ denoting the numerical (instead of analytic) wavenumber. Since the soundfield is composed of a single plane wave, it can be said that the numerical wave propagation velocity $\tilde{c}$ is homogeneous across the entire physical

domain (although it differs from its analytic counterpart). Accordingly, it is possible to express the pressure at the array in the SH domain as

$$p_{nm}(\tilde{k}) = b_n(\tilde{k}r)Y_n^{m^*}(\Omega_0), \tag{30}$$

and the corresponding array output is given by

$$\begin{aligned}\tilde{y}(k,\Omega_L) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{p_{nm}(\tilde{k},r)}{b_n(kr)}Y_n^m(\Omega_L)\\ &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\underbrace{\frac{b_n(\tilde{k}r)}{b_n(kr)}}_{A_n}Y_n^{m^*}(\Omega_0)Y_n^m(\Omega_L).\end{aligned} \tag{31}$$

Note that unlike Eqn (29), here the term "A_n" does not cancel out due to the differences in radial functions for the analytic and numerical wavenumbers. Accordingly, a spatial delta is fully recovered in a dispersive medium only in directions for which $\tilde{k} = k$. The volumetric array proposed here is based on an open-sphere design, therefore the relative error in A_n, in decibels, can be expressed as

$$E_n(kr) = 10\log_{10}\left[\frac{|j_n(\tilde{k}r)|}{|j_n(kr)|}\right]. \tag{32}$$

Figure 3 shows the values of $E_n(kr)$ for three cases of $n = 0$, $n = 2$ and $n = 8$. Results are shown for two representative angles of incidence, namely an axial direction, $\Omega_a \equiv (\frac{\pi}{2}, 0)$, and a diagonal direction, $\Omega_d \equiv (\frac{\pi}{2} - \arctan\left[\sqrt{\frac{1}{2}}\right], \frac{\pi}{4})$.

The *Interpolated Wideband* (IWB) scheme ($\dot{a} = 1/4$, $\dot{b} = 1/16$, $\lambda = 1$) features a constant relative phase velocity in the axial direction, whereas the widely used *Standard Rectilinear* (SRL) scheme ($\dot{a} = 0$, $\dot{b} = 0$, $\lambda = \sqrt{1/3}$) demonstrates similar behaviour only in the diagonal direction, albeit with considerably lower frequency bandwidth. A similar trend is evident in $E_n(kr)$, showing an error of 0dB for the IWB on the axial direction and for the SRL on the diagonal direction. The error for the *Interpolated Isotropic* (IISO2) scheme ($\dot{a} = 0$, $\dot{b} = 1/6$, $\lambda = \sqrt{3/4}$) scheme vanishes in the axial direction, owing to post-warping of the numerical wavenumber. It can also be seen that the relative error increases with frequency, which is to be expected. In addition, the error remains at 0dB for a wider range of frequencies as the mode order, n, is increased. This is proportional to the robustness of the array (see Figure 2), indicating that even though dispersion is more prominent at high frequencies, such errors will generally be less amplified in the PWD process.

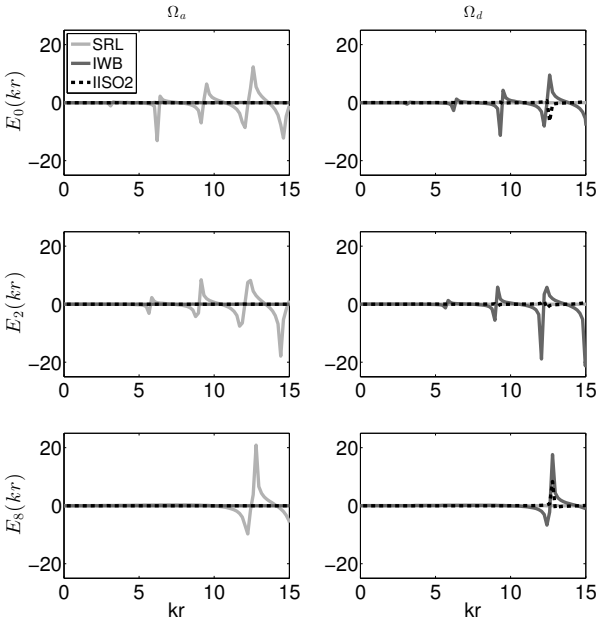

Figure 3. Dispersion error, E_n, shown for a plane wave incident axially (left column) and diagonally (right column), and for orders $n = 0$ (top row), $n = 2$ (middle row) and $n = 8$ (bottom row). SRL - Standard rectilinear, IWB - Interpolated wideband, IISO2 - Interpolated isotropic.

5. Binaural response simulation

To test the proposed method, a domain of $3 \times 3 \times 3$m was discretized in a grid resolution of $X = 10$mm, corresponding to a total of 27×10^6 nodes. A sound source was positioned at a radial distance of 1.5m from a volumetric array ($r_{\texttt{max}} = 0.1$m, $r_{\texttt{min}} = 0$m, total 4169 nodes), and was placed at a minimum distance of 1m from any reflecting boundaries. Simulation was executed using the *Interpolated Wideband* numerical scheme ($f_s = 34,350$Hz) until all propagated waves have fully reached all receiving grid nodes, but without introducing any reflections from the boundaries. Although the proposed method is designed to operate in free field as well as a diffuse field, such a reflection-free response facilitates a preliminary comparison with a closed form solution. The grid was excited using an impulse response of a 32^{nd} order maximally-flat lowpass filter [24, 25] with a cutoff frequency of $f_c = 0.4f_s$ (corresponding to 13,740Hz). This cutoff frequency is well above the 2% dispersion limit for the applied scheme, hence allowing for the resulting signals to be contaminated with numerical dispersion errors. Resulting signals were then processed using the method described in Section 3. To further increase the robustness of the array, all radial functions were soft-limited according to the procedure suggested in [26] providing a total dynamic range of 80dB.

Figure 4 shows a left-ear transfer function rendered using the method proposed in this paper, for a single sound source incident at 0° and 45°. The top two curves (solid gray and dashed black) represent the results of volumetric arrays modeled using FDTD and generated in closed-form, respectively. As additional reference, a single reduced-order HRTF correspond-

ing to the incident direction is also shown (lower, solid gray curve). The reduced-order HRTF was obtained by performing a forward and backward spherical Fourier transform over a left-ear HRTF, $H^l(k, \Omega)$, as in Eqns. (2) and (3), with the SH series truncated at order N instead of infinity. All HRTFs were taken off the KU-100 catalogue provided by the University of Cologne [27].

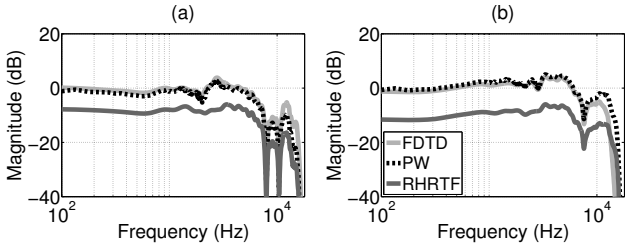

Figure 4. Magnitude of a left-ear binaural response shown for (a) incidence angle of $\Omega = (\frac{\pi}{2}, 0)$ and (b) incidence angle of $\Omega = (\frac{\pi}{2}, \frac{\pi}{4})$. FDTD - finite difference simulation, PW - plane-wave closed form solution, RHRTF - reduced order HRTF (scaled down by 6dB for visual clarity). All results computed with (or reduced to) $N = 12$.

All reference models were convolved with the same excitation signal used in the FDTD simulation, and the reduced-order HRTF was scaled down by 6dB for visual clarity. A good agreement is evident across most of the spectrum (deviation within $\pm$1.3dB upto 12kHz), with the curves diverging mostly at high frequencies. This is to be expected, as the excitation signal was designed to allow the responses to be contaminated with dispersion errors, which are mostly prominent at high frequencies.

6. Concluding remarks

In this paper, the problem of directional receiver modeling in FDTD was addressed by proposing a volumetric array. This volumetric array can be seen as a realization of a spherical shell array [17] on a finite difference grid, in which all samples within a spherical volume can be arbitrarily positioned, hence accounting for the spatial quantization inherent in FDTD. This property also makes the method generalizable to other grid based methods, such as the digital waveguide mesh [28] and the discrete Huygens' model [29].

Compared to traditional array designs, the volumetric array generally employs many more spatial samples, thus resulting in a considerably overdetermined problem and making it a numerically robust choice. This, however, may come at the expense of a considerable memory demand. For example, the largest volumetric array investigated in Sec. 4 includes 4169 samples, which for a 1 second impulse response in a sample rate of 32kHz, would require 500MB of RAM to store the array data alone. Nonetheless, the total required number of samples can be reduced by solving a numerical optimization problem [17], leading to significant memory savings.

In this paper, the volumetric array was applied to the problem of binaural impulse response modeling. In free-field conditions, a $\pm$1.3dB deviation from a reference HRTF was demonstrated. While it is assumed that free-field results are largely generalizable to a reflective environment, the effects of numerical dispersion were not investigated on the level of PWD estimation errors. This is intended to be addressed by the authors in future work. The volumetric array provides a means to describe a modeled soundfield in the spherical harmonics domain, thus it can potentially pave the way to modeling other types of directional receivers by means of known spherical beamforming techniques.

Acknowledgement

The authors would like to thank David Alon for some insightful discussions on broadband array design. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 609465 as part of the Embodied Audition for RobotS (EARS) project.

References

[1] D. Botteldooren, "Finite-difference time-domain simulation of low-frequency room acoustic problems," *The Journal of the Acoustical Society of America*, vol. 98, no. 6, pp. 3302–3308, 1995.

[2] K. Kowalczyk and M. van Walstijn, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 19, no. 1, pp. 34–46, 2011.

[3] T. Van Renterghem, E. Salomons, and D. Botteldooren, "Parameter study of sound propagation between city canyons with a coupled FDTD-PE model," *Applied Acoustics*, vol. 67, pp. 487–510, 2006.

[4] S. Bilbao, "Numerical sound synthesis: Finite difference schemes and simulation in musical acoustics," 2009.

[5] M. Kleiner, B.-I. Dalenbäck, and P. Svensson, "Auralization-an overview," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 861–875, 1993.

[6] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[7] B. Støfringsdal and P. Svensson, "Conversion of discretely sampled sound field data to auralization formats," *Journal of the Audio Engineering Society*, vol. 54, no. 5, pp. 380–400, 2006.

[8] H. Hacihabiboglu, B. Gunel, and Z. Cvetkovic, "Simulation of directional microphones in digital waveguide mesh-based models of room acoustics," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 18, no. 2, pp. 213–223, 2010.

[9] A. Southern, D. T. Murphy, and L. Savioja, "Spatial encoding of finite difference time domain acoustic models for auralization," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 20, no. 9, pp. 2420–2432, 2012.

[10] D. T. Murphy and M. Beeson, "The KW-boundary hybrid digital waveguide mesh for room acoustics applications," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 15, no. 2, pp. 552–564, 2007.

[11] J. Sheaffer, C. Webb, and B. M. Fazenda, "Modelling binaural receivers in finite difference simulation of room acoustics," in *Proceedings of Meetings on Acoustics*, vol. 19, p. 015098, Acoustical Society of America, 2013.

[12] T. Xiao and Q. H. Liu, "Finite difference computation of head-related transfer function for human hearing," *The Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2434–2441, 2003.

[13] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Computer simulation of HRTFs for personalization of 3D audio," in *Universal Communication, 2008. ISUC'08. Second International Symposium on*, pp. 435–440, IEEE, 2008.

[14] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography.* academic press, 1999.

[15] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2149–2157, 2004.

[16] B. Rafaely and A. Avni, "Interaural cross correlation in a sound field represented by spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 823–828, 2010.

[17] B. Rafaely, "The spherical-shell microphone array," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 16, no. 4, pp. 740–747, 2008.

[18] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, 2005.

[19] S. Bilbao, "Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 7, pp. 1524–1533, 2013.

[20] F. Melchior, O. Thiergart, G. Del Galdo, D. de Vries, and S. Brix, "Dual radius spherical cardioid microphone arrays for binaural auralization," in *Audio Engineering Society Convention 127*, Audio Engineering Society, 2009.

[21] I. Balmages and B. Rafaely, "Open-sphere designs for spherical microphone arrays," *IEEE Trans. Audio Speech and Lang. Proc.*, vol. 15, no. 2, pp. 727–732, 2007.

[22] D. L. Alon and B. Rafaely, "Spindle-torus sampling for an efficient-scanning spherical microphone array," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 83–90, 2012.

[23] L. Savioja and V. Valimaki, "Interpolated rectangular 3-d digital waveguide mesh algorithms with frequency warping," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 11, no. 6, pp. 783–790, 2003.

[24] I. Khan and R. Ohba, "Explicit formulas for coefficients of maximally flat fir low/highpass digital filters," *Electronics Letters*, vol. 36, no. 23, pp. 1918–1919, 2000.

[25] J. Sheaffer, M. van Walstijn, and B. Fazenda, "Physical and numerical constraints in source modeling for finite difference simulation of room acoustics," *The Journal of the Acoustical Society of America*, vol. 135, no. 1, pp. 251–261, 2014.

[26] B. Bernschütz, C. Pörschmann, S. Spors, S. Weinzierl, and B. der Verstärkung, "Softlimiting der modalen amplitudenverstärkung bei sphärischen mikrofonarrays im plane wave decomposition verfahren," *Proceedings of the 37. Deutsche Jahrestagung für Akustik (DAGA 2011)*, pp. 661–662, 2011.

[27] B. Bernschütz, "A spherical far field hrir/hrtf compilation of the Neumann KU 100,"

[28] D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, "Acoustic modeling using the digital waveguide mesh," *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 55–66, 2007.

[29] Y. Kagawa, T. Tsuchiya, B. Fujii, and K. Fujioka, "Discrete huygens' model approach to sound wave propagation," *Journal of sound and vibration*, vol. 218, no. 3, pp. 419–444, 1998.