Computation of the head-related transfer function via the fast multipole accelerated boundary element method and representation via the spherical harmonic spectrum

Nail A. Gumerov*

Perceptual Interfaces and Reality Laboratory,

Institute for Advanced Computer Studies,

University of Maryland, College Park. Also at Fantalgo, LLC.

Adam O'Donovan†

Perceptual Interfaces and Reality Laboratory,

Department of Computer Science and Institute for Advanced Computer Studies,

University of Maryland, College Park.

Ramani Duraiswami‡

Perceptual Interfaces and Reality Laboratory,

Department of Computer Science and Institute for Advanced Computer Studies,

University of Maryland, College Park, MD. Also at Fantalgo, LLC.

Dmitry N. Zotkin§

Perceptual Interfaces and Reality Laboratory,

Institute for Advanced Computer Studies,

University of Maryland, College Park, MD. ¶

(Dated: 15 May 2009.)

Abstract

The head-related transfer function (HRTF) is computed using the fast multipole accelerated boundary element method. For efficiency, the HRTF is computed using the reciprocity principle, by placing a source at the ear and computing its field. Analysis is presented to modify the boundary value problem accordingly. To compute the HRTF corresponding to different ranges via a single computation, a compact and accurate representation of the HRTF, termed the *spherical spectrum*, is developed. Computations are reduced to a two stage process, the computation of the spherical spectrum and a subsequent evaluation of the HRTF. This representation allows easy interpolation and range extrapolation of HRTFs.

HRTF computations are performed for the range of audible frequencies up to 20 kHz for several models including a sphere, human head models (for the "Fritz" and "Kemar"), and head and torso model (the Kemar manikin). Comparisons between the different cases and analysis of limiting cases is provided. Comparisons with the computational data of other authors and available experimental data are conducted and show satisfactory agreement for the frequencies for which reliable experimental data is available. Our results show that, given a good mesh it is feasible to compute the HRTF over the full audible range on a regular personal computer.

Published as University of Maryland, Department of Computer Science, Technical Report CS-TR-4936 and UMIACS-TR-2009-06.

PACS numbers: 43.55 Ka, 43.28.Js, 43.20.Fn

Keywords: Head Related Transfer Function, Fast Multipole Method, Boundary Element Method, Preconditioning, Scattering, Spherical Spectrum, HRTF representation

I. INTRODUCTION

The scattering of sound off the bodies of human and animal listeners, especially their heads and external ears, provides them with cues that enable localization of the sound source (Batteau (1967)[5], Hartmann (1999) [21]). The scattering process can be modeled as a linear filter applied to the sound emanating from the source. The transfer function that when applied to the Fourier

*Electronic address: {gumerov@umiacs.umd.edu.}

†Electronic address: {adamod@gmail.com}

‡Electronic address: {ramani@umiacs.umd.edu}

§Electronic address: {dz@umiacs.umd.edu}

¶URL: www.umiacs.umd.edu/~gumerov; URL: http://adamod.googlepages.com; URL: www.umiacs.umd.edu/~ramani; URL: www.umiacs.umd.edu/~dz

transform of the source sound at the head center in the absence of the head, and converts it to the Fourier transform of the received sound at the ear is called the "Head Related Transfer Function (HRTF)." It captures the scattering behavior of the ear, head, and body of the listener for various locations at which acoustical sources may be. If the head is centered at a given point P and the sound source located at elevation θ, azimuth φ, and distance r in a head-centered spherical coordinate system, then the HRTF $H(r, \theta, \varphi, f)$ is the ratio of the Fourier transform of the signal at the ear $F_e(f)$ to the Fourier transform of the signal that would have been received at the point P in free-field $F_P(f)$, where f is the signal frequency. Thus the HRTF is a function of four variables and must be measured or computed as a function of these. Roughly speaking, the size of the head is important above 1 kHz, the general characteristics of the torso are important below 3 kHz, and the detailed structure of the head and pinnae become significant above 3 kHz.

Typically, the HRTF is measured. Both direct and reciprocal HRTF measurements have been reported in the literature. In the former method a source (a speaker) is moved to various locations and the response measured at a microphone placed at the entrance to the blocked ear canal (see e.g., [13], [2]). In the latter, an outwardly directed source (a microspeaker) is placed at the entrance to the blocked ear canal, and the HRTF is determined using Helmholtz' reciprocity principle by measuring the received signal at microphones placed at locations of interest [34].

Since the HRTF results from a scattering process, many computational acousticians have attempted to compute the HRTF by solving the wave equation subject to boundary conditions on the surface of the head [22, 23, 28, 32]. Another area, where there has been interest is in the simulation of the HRTF of various animals, e.g., the dolphin [4] and the bat [9]. The numerical HRTF is usually obtained by solving the Fourier transform of the wave equation (the Helmholtz equation) at wavenumbers corresponding to frequencies of interest, though direct simulation of the wave equation in the time domain has also been attempted [33]. When the solution is obtained using the direct experimental configuration, the simulations must be repeated for each source position, similar to the experimental setting. The solution can also be done in the reciprocal setting.

Numerical simulation is attractive since it offers the possibility of extracting HRTFs without subjecting the user to measurement (beyond those needed to create the discretization of their body). Further, numerical simulation, if it were easy and accurate, offers the promise of allowing one to relate features in the HRTF with anatomical structure of the head, pinna and body. By manipulating the mesh and observing the resulting computed HRTF it might be possible to explore

the sensitivity of the HRTF to particular features. However the match between experiment and simulation over the entire auditory range have not yet been presented, and such sensitivity analysis still remains to be done. Moreover, computations are relatively time consuming (taking days of computer time), and simulations cannot yet be used for parametric studies.

As can be seen from Table I, the progress on simulations has been gradual, and the ability to completely synthesize HRTFs from meshes and simulations is not yet available. This paper presents a first such computation. Progress has been hindered by several factors. First, the simulations are relatively time consuming. Since the wavelengths that must be resolved are relatively small, and approximately 6 to 10 elements are needed per wavelength, large meshes result. When the boundary element methods are solved directly, the solution time scales as the cube of the discretization size and this can lead to hours and days of simulation time, making numerical simulation unattractive in comparison with direct experimentation. Further, in addition to the inordinate time needed, the amount of memory needed to achieve a simulation at higher frequencies means that these simulations can only be done on supercomputers, and this has meant that most studies do not perform the simulations for the high frequencies. Second, the comparison between experimentally measured HRTFs and numerically computed ones has not been demonstrated satisfactorily at all frequencies of interest. The reasons for this could lie in inaccuracies in the simulation procedure or in the quality of the mesh discretization.

For the computations to be useful, the HRTF information must be extracted and presented in a manner similar to the experiments. Further, the results of the computations and experiments must be compared to see that they agree, and any discrepancies explained. Table I summarizes the results of several papers. Due to the inability of previous authors to simulate the entire frequency range, such comparisons were attempted only for restricted ranges, and the results showed less than satisfactory agreement between the experimental and the computed data [14].

In this paper we use the preconditioned multipole accelerated boundary element method [19], a new formulation of the HRTF problem, and good meshes of the head and torso. We then compare the computation results to experiments. In section II, the theoretical background for HRTF computations is discussed. In particular, we derive the equations necessary to use the reciprocal set-up for computations. In section III, we introduce the spherical representation of the

Ref.	Method	Dir/Rec	$f_{\max}\,(kHz)$	Mesh	Time	Exp.
[23]	BEM	Recip.	5.4	22000	28 h/freq	Y
[22]	BEM	Recip.	10.3	30570	9.15 h /freq	Y
[33]	FDTD-PML	Direct	8	$[130]^3$	0.66 h/ direction	N
[28]	BEM	***	20	28000	480 h	N
[14]	BEM	Direct	15	21693	*52.5 h*	Y
This paper	FMBEM	Recip.	20	445728	70 h	Y

TABLE I: The table shows the results from a few papers, and compares the numerical method used, whether a direct or reciprocal formulation was employed, the maximum frequency employed, the size of the mesh, the time taken, and if the computation results are compared with experiments. Data in italics refers to projections rather than actual computations. Ref. [28] uses a different numerical method. None of the papers referred to include the torso in the computations.

HRTF corresponding to insonification from plane-waves from various directions. It is shown how one may then derive the HRTF corresponding to a source at a finite range from this representation. In section IV we briefly describe the numerical methods used in this paper, namely the preconditioned FMM BEM, and its use to compute the spherical harmonic spectrum. Section V presents numerical results for computations with an analytical solution (a sphere), the Neumann KU-100 "Fritz" head, and the Knowles Electronics Manikin for Acoustic Research (KEMAR) head and head-and-torso models. Section VI concludes the paper.

II. PROBLEM FORMULATION

A. Direct computations

To compute the HRTF for a given frequency $f = \omega/(2\pi)$ the boundary value problem for the Helmholtz equation in three dimensions must be solved to determine the scattered field for different plane wave incident angles or source locations generating the incident field. Let the origin of the reference frame be placed at some point inside the head, which is referred to as the "head center" and let $\mathbf{r}$ denote the radius-vector of any spatial point in this frame. Along with the Cartesian

coordinates (x, y, z) the spherical coordinates (r, θ, φ) defined as

$$\begin{aligned}&\mathbf{r} = (x, y, z) = r\left(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta\right), \\ &0 \leqslant \varphi < 2\pi, \quad 0 \leqslant \theta \leqslant \pi, \quad r = |\mathbf{r}|,\end{aligned} \tag{1}$$

are also used. Let S denote the surface of the scatterer and V the infinite external domain.

When the incident field is generated by a source, the sphere S' or radius R (range) centered at the origin, which includes surface S and the source location, must also be considered. This sphere is characterized by the vector $\mathbf{r}'$, or alternately by the radius R and a unit vector $\mathbf{s}$,

$$\mathbf{r}' = R\mathbf{s}, \quad \mathbf{s} = \left(\sin\theta'\cos\varphi', \sin\theta'\sin\varphi', \cos\theta'\right), \quad |\mathbf{s}| = 1. \tag{2}$$

Thus, the incident complex pressure field generated by a unit intensity monopole source placed on S' is

$$\Phi^{in}(\mathbf{r}; \mathbf{s}) = \Phi^{in}_R(\mathbf{r}; \mathbf{r}') = G\left(\left|\mathbf{r} - \mathbf{r}'\right|\right) = G\left(|\mathbf{r} - R\mathbf{s}|\right), \tag{3}$$

where

$$G(r) = \frac{e^{ikr}}{4\pi r}, \quad k = \frac{\omega}{c}, \tag{4}$$

is the free-field Green's function for the Helmholtz equation, k is the wavenumber, and c is the speed of sound in air.

The other case considered is when the field is a plane wave incident from direction $\mathbf{s}$:

$$\Phi^{in}(\mathbf{r}; \mathbf{s}) = e^{-ik\mathbf{s}\cdot\mathbf{r}}, \quad |\mathbf{s}| = 1. \tag{5}$$

In fact this field can be considered as a limiting case of the filed generated by the monopole at large R, since

$$\Phi^{in}(\mathbf{r}; \mathbf{s}) = e^{-ik\mathbf{s}\cdot\mathbf{r}} = \lim_{R\to\infty}\left[4\pi R e^{-ikR} G(\mathbf{r} - R\mathbf{s})\right] = \lim_{R\to\infty}\left[4\pi R e^{-ikR}\Phi^{in}_R(\mathbf{r}; R\mathbf{s})\right]. \tag{6}$$

The total acoustic field can be represented as

$$\Phi = \Phi^{in} + \Phi^{scat}, \tag{7}$$

where the scattered field Φ^{scat} is a regular function in V which can be found by solution of the Helmholtz equation

$$\nabla^2\Phi^{scat} + k^2\Phi^{scat} = 0, \quad \mathbf{r} \in V, \tag{8}$$

with the Sommerfeld radiation condition

$$\lim_{r\to\infty}\left[r\left(\frac{\partial\Phi^{scat}}{\partial r}-ik\Phi^{scat}\right)\right]=0, \tag{9}$$

and boundary conditions on S, which for simplicity and as is customary in HRTF calculations, is assumed to be a sound-hard surface (the impedance boundary conditions can also be imposed):

$$\left.\frac{\partial\Phi^{scat}}{\partial n}\right|_S=-\left.\frac{\partial\Phi^{in}}{\partial n}\right|_S,\quad\left(\left.\frac{\partial\Phi}{\partial n}\right|_S=0\right). \tag{10}$$

This problem is closed and its solution is unique. However, to compute the HRTF we need to compute Φ only at one point on the surface S to obtain the HRTF. This is the point corresponding to the receiver (microphone placed at the entrance to the ear canal) location. Further, if a symmetric head model is used, the HRTF for the right ear is computed for all directions, the HRTF for the left ear can be simply obtained from this solution. Let the location of the receiver be denoted as $\mathbf{r}_*$.

So all that is needed from the solution of the above boundary value problem is to determine the function

$$H(\mathbf{s})=\begin{cases}\Phi(\mathbf{r}_*;R\mathbf{s})\\ \Phi(\mathbf{r}_*;\mathbf{s})\end{cases},\quad\mathbf{r}_*\in S, \tag{11}$$

where the upper and lower formulae correspond to the monopole and plane wave incident fields.

The HRTF is the complex valued function $H(\mathbf{s})$ that is normalized by the value of the incident field at the head center, which by the choice of the reference frame is the origin. The unitary plane wave satisfies $e^{-ik\mathbf{s}\cdot\mathbf{r}}=1$ at $\mathbf{r}=\mathbf{0}$, so $H(\mathbf{s})$ is the complex HRTF in this case.

For the incident field generated by a source we have

$$HRTF(\mathbf{s};R)=\frac{H(\mathbf{s};R)}{G(|R\mathbf{s}|)}=\frac{H(\mathbf{s};R)}{G(R)}=4\pi Re^{-ikR}H(\mathbf{s};R)=4\pi Re^{-ikR}\Phi(\mathbf{r}_*;R\mathbf{s}). \tag{12}$$

Due to the linearity of the scattering problem and Eq. (6),

$$HRTF(\mathbf{s};\infty)=HRTF^{(pw)}(\mathbf{s};\infty)=H^{(pw)}(\mathbf{s}), \tag{13}$$

Here and below the superscript (pw) is used to show that the respective function or constant is related to the plane wave incidence case. So the plane wave HRTF is nothing but a limit at $R\to\infty$ of the HRTF for the case of finite range R.

B. Modified boundary value problem for the reciprocity principle

Even though the direct solution can provide the HRTF it is obvious that HRTF computations will be excessively long, as a separate solution is needed for each direction. Further, the only part of the solution retained is the acoustic pressure at one receiver point, with the rest discarded! To avoid this, the reciprocity principle [26] can be used (see [22, 23]). This principle states that the positions of the source and receiver in a given domain can be exchanged, which does not affect the value of the potential at the receiver location. In our case the receiver is placed on the domain boundary. However, the singularity of the source can cause numerical and conceptual problems unless it is explicitly treated. Such a treatment is presented here.

To derive appropriate forms of the boundary value problem for this situation, we will place the source at the point

$$\mathbf{r}_{*\epsilon} = \mathbf{r}_* + \epsilon \mathbf{n}, \tag{14}$$

where $\mathbf{n}$ is the normal to the boundary directed inside V and ϵ is assumed small enough, so that due to the continuity of the pressure, the values obtained for $\Phi(\mathbf{r}_*; \mathbf{s})$ and $\Phi(\mathbf{r}_{*\epsilon}; \mathbf{s})$ are close. We will then consider the limiting case $\epsilon \to 0$ which will provide the result for the receiver placed exactly on the boundary independent on ϵ.

1. *Incident field generated by a source*

Let $\phi(\mathbf{r}; \mathbf{r}_{*\epsilon})$ be a solution of the following problem

$$\begin{aligned} &\phi = \phi^{in} + \phi^{scat}, \quad \phi^{in} = G\left(|\mathbf{r} - \mathbf{r}_{*\epsilon}|\right), \quad \nabla^2 \phi^{scat} + k^2 \phi^{scat} = 0, \quad \mathbf{r} \in V, \\ &\lim_{r\to\infty} \left[r \left(\frac{\partial \phi^{scat}}{\partial r} - ik\phi^{scat} \right) \right] = 0, \quad \left. \frac{\partial \phi^{scat}}{\partial n} \right|_S = - \left. \frac{\partial \phi^{in}}{\partial n} \right|_S, \quad \left(\left. \frac{\partial \phi}{\partial n} \right|_S = 0 \right). \end{aligned} \tag{15}$$

Consider the case, when the incident field in the direct problem formulation is generated by a source located at $\mathbf{r}'$, the solution to which we denote as $\Phi(\mathbf{r}; \mathbf{r}')$. Due to the reciprocity principle, we have

$$\phi\left(\mathbf{r}'; \mathbf{r}_{*\epsilon}\right) = \Phi(\mathbf{r}_{*\epsilon}; \mathbf{r}'). \tag{16}$$

Therefore, function $H(\mathbf{s})$ (11) can be computed in this case as

$$H(\mathbf{s}) = \lim_{\epsilon\to 0} \Phi(\mathbf{r}_{*\epsilon}; \mathbf{r}') = \lim_{\epsilon\to 0} \phi\left(\mathbf{r}'; \mathbf{r}_{*\epsilon}\right) = \lim_{\epsilon\to 0} \phi\left(R\mathbf{s}; \mathbf{r}_{*\epsilon}\right). \tag{17}$$

Using the decomposition of the total field into the incident and scattered fields (15) this can be rewritten as

$$H(\mathbf{s}) = G(|\mathbf{r}_* + R\mathbf{s}|) + \lim_{\epsilon \to 0} \phi^{scat}(R\mathbf{s}; \mathbf{r}_{*\epsilon}). \tag{18}$$

Since $\mathbf{r}' \notin S$, and, in fact, this point is far from the boundary, the function $\phi^{scat}(\mathbf{r}'; \mathbf{r}_{*\epsilon})$ is regular and well behaved. The limit in (18) is finite and exists. However, in numerical solution of the problem the boundary value $\phi^{scat}(\mathbf{r}_*; \mathbf{r}_{*\epsilon}) \sim \epsilon^{-1}$, which causes substantial numerical errors for small ϵ. (In fact, previous authors have attempted to deal with the locally high gradients by modifying the mesh to account for this, see e.g., [22].)

A straightforward use of Green's identity for determination $\phi^{scat}(\mathbf{r}'; \mathbf{r}_{*\epsilon})$ is also problematic. To treat the singular behavior of the scattered field, we will use the fact that for a sound hard smooth boundary a source located at $\mathbf{r}_{*\epsilon}$ generates an *image*, which is the source of the same intensity located at $\mathbf{r}^-_{*\epsilon} = \mathbf{r}_* - \epsilon\mathbf{n}$. We will represent ϕ^{scat} as

$$\phi^{scat} = G\left(\left|\mathbf{r} - \mathbf{r}^-_{*\epsilon}\right|\right) + \psi, \tag{19}$$

where ψ is a regular radiating function satisfying the Helmholtz equation in V, which also satisfies the following boundary condition

$$\left.\frac{\partial \psi}{\partial n}\right|_S = \left[-\frac{\partial}{\partial n} G\left(\left|\mathbf{r} - \mathbf{r}^-_{*\epsilon}\right|\right) - \frac{\partial \phi^{in}}{\partial n}\right]\Bigg|_S = -\left.\frac{\partial}{\partial n}\left[G\left(\left|\mathbf{r} - \mathbf{r}^-_{*\epsilon}\right|\right) + G(|\mathbf{r} - \mathbf{r}_{*\epsilon}|)\right]\right|_S. \tag{20}$$

This formula allows us to take the limit $\epsilon \to 0$, which shows that

$$\left.\frac{\partial \psi}{\partial n}\right|_S = \begin{cases} 0, & \mathbf{r} = \mathbf{r}_*, \\ -2\frac{\partial}{\partial n} G(|\mathbf{r} - \mathbf{r}_*|), & \mathbf{r} \neq \mathbf{r}_* \end{cases}. \tag{21}$$

As soon as ψ can be computed even at $\epsilon = 0$ the required function $H(\mathbf{s})$ can be found from Eqs (18) and (19) as

$$H(\mathbf{s}) = 2G(|\mathbf{r}_* - R\mathbf{s}|) + \psi(R\mathbf{s}; \mathbf{r}_*). \tag{22}$$

The physical meaning of this procedure is that the far field of a source located near a flat sound hard boundary is close to the field of a single source of double intensity without any boundary.

2. Use of Green's identity

Since ψ is a regular function in V, Green's identity can be applied to determine ψ at any field point including $\mathbf{r}'$ via boundary values of this function, which can be computed using numerical

solvers, such as the BEM:

$$-\psi\left(\mathbf{r}'\right) = L\left[q\right] - M\left[\psi\right], \quad q = \left.\frac{\partial\psi}{\partial n}\right|_S, \tag{23}$$

where L and M denote the following boundary operators:

$$L\left[q\right] = \int_S q\left(\mathbf{r}\right) G\left(\left|\mathbf{r}-\mathbf{r}'\right|\right) dS(\mathbf{r}), \quad M\left[\psi\right] = \int_S \psi\left(\mathbf{r}\right) \frac{\partial G\left(\left|\mathbf{r}-\mathbf{r}'\right|\right)}{\partial n\left(\mathbf{r}\right)} dS(\mathbf{r}), \tag{24}$$

where the normal orientation is consistent with Eq. (14).

Use of Green's identity allows us to link the function $H\left(\mathbf{s}\right)$ (22) with the boundary values of ψ:

$$H\left(\mathbf{s}\right) = 2G\left(\left|\mathbf{r}_* - R\mathbf{s}\right|\right) + \int_S \left[\psi\left(\mathbf{r};\mathbf{r}_*\right) \frac{\partial G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)}{\partial n\left(\mathbf{r}\right)} - q\left(\mathbf{r};\mathbf{r}_*\right) G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)\right] dS(\mathbf{r}). \tag{25}$$

3. Plane wave incident field

Consider now the case, when the incident field in the direct problem is a plane wave (5). In this case too $H\left(\mathbf{s}\right)$ can be found from the solution of the same problem (source on the boundary). Indeed, consider the solution of the previous problem, where instead of a unit intensity source we have a source of intensity $4\pi R e^{-ikR}$. Due to the linearity of the problem function $H\left(\mathbf{s}\right)$ in this case can be computed according to (25) as

$$H\left(\mathbf{s}\right) = 4\pi R e^{-ikR}\left[2G\left(\left|\mathbf{r}_* - R\mathbf{s}\right|\right) + \int_S \left[\psi\left(\mathbf{r};\mathbf{r}_*\right) \frac{\partial G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)}{\partial n\left(\mathbf{r}\right)} - q\left(\mathbf{r};\mathbf{r}_*\right) G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)\right] dS(\mathbf{r})\right]. \tag{26}$$

Consider now limit as $R \to \infty$. Eq. (6) shows that in this case the incident field for the direct problem will be $e^{ik\mathbf{s}\cdot\mathbf{r}}$. The same formula applied to Eq. (26) results in

$$H^{(pw)}\left(\mathbf{s}\right) = 2e^{-ik\mathbf{s}\cdot\mathbf{r}_*} - \int_S e^{-ik\mathbf{s}\cdot\mathbf{r}}\left[ik\mathbf{n}\left(\mathbf{r}\right)\cdot\mathbf{s}\psi\left(\mathbf{r};\mathbf{r}_*\right) + q\left(\mathbf{r};\mathbf{r}_*\right)\right] dS(\mathbf{r}), \tag{27}$$

since

$$\lim_{R\to\infty}\left[4\pi R e^{-ikR}\frac{\partial G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)}{\partial n}\right] = \mathbf{n}\cdot\nabla_{\mathbf{r}} \lim_{R\to\infty}\left[4\pi R e^{-ikR} G\left(\left|\mathbf{r}-R\mathbf{s}\right|\right)\right] = \mathbf{n}\cdot\nabla_{\mathbf{r}} e^{ik\mathbf{s}\cdot\mathbf{r}} = -ik\mathbf{n}\cdot\mathbf{s}e^{ik\mathbf{s}\cdot\mathbf{r}}. \tag{28}$$

III. SPHERICAL REPRESENTATIONS

Since $H\left(\mathbf{s}\right)$ is a regular function defined on the unit sphere, and the spherical harmonics form a basis on this surface, it is natural to represent $H\left(\mathbf{s}\right)$ as a sum of spherical harmonics:

$$H\left(\mathbf{s}\right) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} H_n^m Y_n^m\left(\mathbf{s}\right), \tag{29}$$

where the orthonormal spherical harmonics for $\mathbf{s} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ are defined as

$$Y_n^m(\mathbf{s}) = (-1)^m \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\theta) e^{im\varphi}, \tag{30}$$
$$n = 0, 1, 2, ..., \quad m = -n, ..., n,$$

and $P_n^{|m|}(\mu)$ are the associated Legendre functions consistent with that in [1], or Rodrigues' formulae

$$P_n^m(\mu) = (-1)^m \left(1-\mu^2\right)^{m/2} \frac{d^m}{d\mu^m} P_n(\mu), \quad n \geqslant 0, \quad m \geqslant 0, \tag{31}$$
$$P_n(\mu) = \frac{1}{2^n n!}\frac{d^n}{d\mu^n}\left(\mu^2-1\right)^n, \quad n \geqslant 0,$$

where $P_n(\mu)$ are the Legendre polynomials.

In this basis the coefficient of expansion (29) can be found as

$$H_n^m = \int_{S_u} H(\mathbf{s}) Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}), \tag{32}$$

where S_u is the unit sphere surface, due to orthonormality

$$\int_{S_u} Y_{n'}^{m'}(\mathbf{s}) Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}) = \delta_{mm'}\delta_{nn'}, \tag{33}$$

where δ is the Kronecker symbol.

A. Plane wave incident field

Substituting Eq. (27) into Eq. (32), we obtain

$$H_n^{(pw)m} = \int_{S_u} \left[2e^{-ik\mathbf{s}\cdot\mathbf{r}_*} - \int_S e^{ik\mathbf{s}\cdot\mathbf{r}} \left[ik\mathbf{n}(\mathbf{r})\cdot\mathbf{s}\psi(\mathbf{r};\mathbf{r}_*) + q(\mathbf{r};\mathbf{r}_*)\right] dS(\mathbf{r})\right] Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}). \tag{34}$$

Note then that the Gegenbauer formula for plane wave expansion and addition theorem for spherical harmonics (e.g. see [15]) provides:

$$e^{-ik\mathbf{s}\cdot\mathbf{r}} = 4\pi \sum_{n'=0}^{\infty}\sum_{m'=-n'}^{n'} i^{-n'} Y_{n'}^{m'}(\mathbf{s}) R_{n'}^{-m'}(\mathbf{r}), \quad R_{n'}^{m'}(\mathbf{r}) = j_{n'}(kr) Y_{n'}^{m'}\left(\frac{\mathbf{r}}{r}\right), \tag{35}$$

$$-ik\mathbf{n}\cdot\mathbf{s}e^{-ik\mathbf{s}\cdot\mathbf{r}} = \mathbf{n}\cdot\nabla_{\mathbf{r}} e^{-ik\mathbf{s}\cdot\mathbf{r}} = 4\pi \sum_{n'=0}^{\infty}\sum_{m'=-n'}^{n'} i^{-n'} Y_{n'}^{m'}(\mathbf{s}) \left[\mathbf{n}\cdot\nabla R_{n'}^{-m'}(\mathbf{r})\right].$$

where $j_n(kr)$ are the spherical Bessel functions of the first kind and $R_n^m(\mathbf{r})$ are the regular spherical wave functions. Substituting these expansions into Eq. (34) and using the orthonormality property (33), we obtain

$$H_n^{(pw)m} = 4\pi i^{-n}\left\{2R_n^{-m}(\mathbf{r}_*) + \int_S \left[\psi(\mathbf{r};\mathbf{r}_*)\,\mathbf{n}(\mathbf{r})\cdot\nabla R_n^{-m}(\mathbf{r}) - q(\mathbf{r};\mathbf{r}_*)\,R_n^{-m}(\mathbf{r})\right] dS(\mathbf{r})\right\}. \quad (36)$$

B. Arbitrary incident field

Before providing a formula for the spherical representation of $H(\mathbf{s})$ for a source, we note that the plane wave solution allows us to compute of function $H(\mathbf{s})$ for an arbitrary incident field regular inside a sphere which includes the scatterer. Indeed, any such field can be represented as

$$\Phi^{in}(\mathbf{r}) = \int_{S_u} F(\mathbf{s}')e^{-ik\mathbf{s}'\cdot\mathbf{r}}dS(\mathbf{s}'), \quad (37)$$

which is a superposition of plane waves $e^{-ik\mathbf{s}'\cdot\mathbf{r}}$ of intensity $F(\mathbf{s}')dS(\mathbf{s}')$. The function H is a result of the application of some linear operator to the incident field, and therefore we can write

$$H = \int_{S_u} F(\mathbf{s}')H^{(pw)}(\mathbf{s}')\,dS(\mathbf{s}'), \quad (38)$$

Since $H^{(pw)}(\mathbf{s}')$ is also a function that can be decomposed as (29), we can rewrite the above equation as

$$H = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} H_n^{(pw)m}\int_{S_u} F(\mathbf{s}')Y_n^m(\mathbf{s}')\,dS(\mathbf{s}'), \quad (39)$$

where $H_n^{(pw)m}$ are determined by Eq. (36). We note now that the integral in Eq. (39) represents coefficients of expansion of surface function $F(\mathbf{s}')$ over the spherical harmonics basis similarly to Eqs (29) and (32) :

$$\int_{S_u} F(\mathbf{s}')Y_n^m(\mathbf{s}')\,dS(\mathbf{s}') = F_n^{-m}, \quad F(\mathbf{s}') = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} F_n^m Y_n^m(\mathbf{s}'). \quad (40)$$

In this case the incident field (37) can be represented as an expansion over functions $R_n^m(\mathbf{r})$, since substituting expansion of the plane wave (35) into Eq. (37) we obtain:

$$\begin{aligned}\Phi^{in}(\mathbf{r}) &= 4\pi\sum_{n=0}^{\infty}\sum_{m=-n}^{n} i^{-n}R_n^{-m}(\mathbf{r})\int_{S_u} F(\mathbf{s}')Y_n^m(\mathbf{s}')\,dS(\mathbf{s}') \quad (41)\\ &= 4\pi\sum_{n=0}^{\infty}\sum_{m=-n}^{n} i^{-n}F_n^{-m}R_n^{-m}(\mathbf{r}) = 4\pi\sum_{n=0}^{\infty}\sum_{m=-n}^{n} i^{-n}F_n^m R_n^m(\mathbf{r}).\end{aligned}$$

Denoting $C_n^m = 4\pi i^{-n} F_n^m$, we can see that

$$\Phi^{in}(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_n^m R_n^m(\mathbf{r}), \quad H = \frac{1}{4\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n C_n^{-m} H_n^{(pw)m}. \tag{42}$$

which shows how from a given spherical representation of the incident field we can obtain H.

C. Incident field generated by a source

Despite a procedure similar to that for the plane wave can be applied to Eq. (25) to obtain coefficients H_n^m for the case of the incident field generated by a source, we can directly apply Eq. (42) to this case, which allows us to link $H(\mathbf{s})$ for both incident fields of interest and, moreover, obtain useful formulae for arbitrary range R.

Indeed the incident field generated by a source (3) inside the sphere S' can be represented in the form (e.g. [15])

$$\begin{aligned} \Phi^{in}(\mathbf{r};\mathbf{s}) = G(|\mathbf{r} - R\mathbf{s}|) &= ik \sum_{n=0}^{\infty} \sum_{m=-n}^{n} S_n^{-m}(R\mathbf{s}) R_n^m(\mathbf{r}) \\ &= ik \sum_{n=0}^{\infty} \sum_{m=-n}^{n} h_n(kR) Y_n^{-m}(\mathbf{s}) R_n^m(\mathbf{r}), \quad S_n^m(\mathbf{r}) = h_n(kr) Y_n^m\left(\frac{\mathbf{r}}{r}\right), \end{aligned} \tag{43}$$

where $h_n(kr)$ are the spherical Hankel functions of the first kind and $S_n^m(\mathbf{r})$ are the regular spherical wave functions. Comparing this with Eq. (42), we can see that for the incident field of monopole source

$$H(\mathbf{s}) = \frac{ik}{4\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n h_n(kR) H_n^{(pw)m} Y_n^m(\mathbf{s}). \tag{44}$$

So if we represent $H(\mathbf{s})$ in the form (29), then coefficients of expansion H_n^m will be simply related to $H_n^{(pw)m}$ as

$$H_n^m = \frac{ik}{4\pi} i^n h_n(kR) H_n^{(pw)m}. \tag{45}$$

This result can be easily checked since we have

$$\begin{aligned} H_n^{(pw)m} &= \lim_{R\to\infty} \left[4\pi R e^{-ikR} H_n^m \right] = \lim_{R\to\infty} \left[4\pi R e^{-ikR} \frac{ik}{4\pi} i^n h_n(kR) H_n^{(pw)m} \right] \\ &= H_n^{(pw)m} \lim_{R\to\infty} \left[ikR e^{-ikR} i^n i^{-n} \frac{e^{ikR}}{ikR} \left(1 + O\left(\frac{1}{R}\right) \right) \right] = H_n^{(pw)m}. \end{aligned} \tag{46}$$

Also, the above formulae allow us to link the functions $H(\mathbf{s})$ and $H^{(pw)}(\mathbf{s})$ via integral transform avoiding having to perform conversion to spherical representations. Indeed, we have from Eqs (44) and (32):

$$\begin{aligned} H(\mathbf{s}) &= \frac{ik}{4\pi}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} i^n h_n(kR)\, Y_n^m(\mathbf{s}) \int_{S_u} H^{(pw)}(\mathbf{s}')\, Y_n^{-m}(\mathbf{s}')\, dS(\mathbf{s}') \\ &= \frac{ik}{4\pi}\sum_{n=0}^{\infty} i^n h_n(kR) \int_{S_u} H^{(pw)}(\mathbf{s}') \sum_{m=-n}^{n} Y_n^m(\mathbf{s})\, Y_n^{-m}(\mathbf{s}')\, dS(\mathbf{s}') \\ &= \frac{ik}{(4\pi)^2}\sum_{n=0}^{\infty} i^n (2n+1)\, h_n(kR) \int_{S_u} H^{(pw)}(\mathbf{s}')\, P_n(\mathbf{s}\cdot\mathbf{s}') dS(\mathbf{s}') \\ &= \frac{ik}{(4\pi)^2}\lim_{p\to\infty}\int_{S_u} \Lambda_s^{(p)}(kR, \mathbf{s}\cdot\mathbf{s}')\, H^{(pw)}(\mathbf{s}')\, dS(\mathbf{s}'). \end{aligned} \tag{47}$$

Here

$$\Lambda_s^{(p)}(\rho,\mu) = \sum_{n=0}^{p-1} i^n (2n+1)\, h_n(\rho)\, P_n(\mu), \tag{48}$$

is a singular translation kernel of bandwidth p (see [29]), and we used the addition theorem for spherical harmonics

$$P_n(\mathbf{s}_1\cdot\mathbf{s}_2) = \frac{4\pi}{2n+1}\sum_{m=-n}^{n} Y_n^m(\mathbf{s}_1)\, Y_n^{-m}(\mathbf{s}_2). \tag{49}$$

Note that the series (48) diverges as $p\to\infty$, so that the ordering of the infinite sum and integral in (47) cannot be exchanged. However, for a bandlimited kernel, the sums are finite and such an exchange is valid, and the limit of the integral of a bandlimited kernel exists and is finite. Finally, we note that if computations are performed only for a fixed range R (e.g. to compare with experimental data obtained for a given R), then it may be simpler to compute integral (25) directly. We also note that the link (45) is universal and can be used to reduce the size of experimentally measured data set. Indeed, we can express range independent function $H^{(pw)}(\mathbf{s})$ via $H(\mathbf{s})$ in a similar way as derived in (47):

$$H^{(pw)}(\mathbf{s}) = \frac{1}{ik}\int_{S_u} M_s(kR, \mathbf{s}\cdot\mathbf{s}')\, H(\mathbf{s}')\, dS(\mathbf{s}'), \tag{50}$$

with the kernel

$$M_s(\rho,\mu) = \sum_{n=0}^{\infty}\frac{2n+1}{i^n h_n(\rho)} P_n(\mu). \tag{51}$$

So if $H(\mathbf{s})$ is measured or computed then $H^{(pw)}(\mathbf{s})$ can be found, and then $H(\mathbf{s})$ for some different range can be computed using Eq. (47).

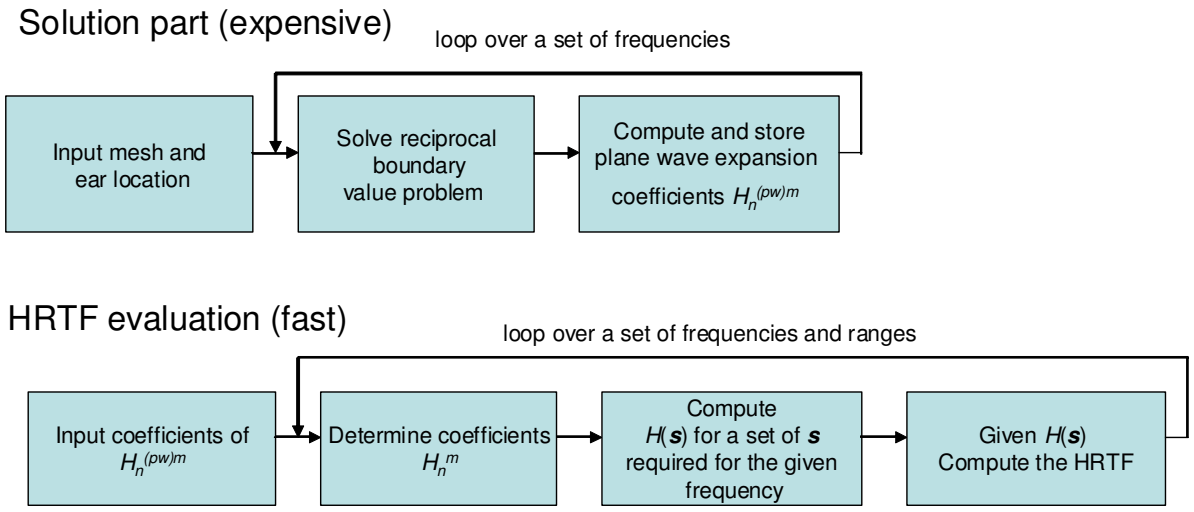

FIG. 1: The flowchart illustrates the two stages in the HRTF computation procedure proposed in this paper. The solution part includes the BEM solver for the reciprocal boundary value problem and the spherical harmonic decomposition of the *range independent* HRTF. In the evaluation part the spherical harmonic spectra are used to interpolate the HRTF corresponding to any direction and range.

IV. COMPUTATIONAL METHODS

The FMM accelerated BEM formulation for the Neumann problem for the Helmholtz equation in 3D [19] is at the heart of the computational technique used to compute the HRTF. The process of obtaining function $H(\mathbf{s})$ and related amplitude or phase of the HRTF is represented by the flow chart in Fig. 1.

As soon as the geometrical model is specified (mesh and ear location) the problem for the scattered field ψ can be solved subject to the boundary conditions (21). The last step of the solution part is determination of the expansion coefficients $H_n^{(pw)m}$ of the function $H^{(pw)}(\mathbf{s})$, which is done using Eq. (36). We note that the coefficients $H_n^{(pw)m}$ (whose number depends on the frequency) provide the theoretically most compact representation of $H^{(pw)}(\mathbf{s})$ within a given bandwidth. Further, the use of the spherical representation provides a fast and accurate interpolation of $H^{(pw)}(\mathbf{s})$ within a given bandwidth for any other direction (thus, this also provides low-pass spherical filtering of the HRTF).

The evaluation part can then be performed relatively quickly for an arbitrary range R. The case $R \to \infty$ is, of course, just $H_n^{(pw)m}$ itself. Computation of the coefficients H_n^m for fields generated

by monopole sources is fast, as it can be done using Eq. (45). After that the function $H(\mathbf{s})$ can be determined for any given direction using Eq. (29) and the HRTF or any other related functions can be computed. We note that the evaluation part can be extended to include, for example, computation of the head related impulse response (HRIR), which is the Fourier transform of the HRTF. Also, the procedure can include computations of the HRTF for incident fields generated by sources other than monopole (e.g. to include beampatterns of speakers). For non-monopole sources the incident field should be decomposed over the spherical harmonics and the resulting solution can be easily found using the computed $H_n^{(pw)m}$ and Eq. (42).

A. BEM accelerated by the FMM

The boundary element method used in the present study is described in details in our recent paper [19], where a review of related work can also be found. This software uses the Burton-Miller formulation [7]. The boundary integral equation solved is

$$-\frac{1}{2}\left[\psi\left(\mathbf{r}'\right)+\lambda q(\mathbf{r}')\right]=\left(L+\lambda L'\right)[q]-\left(M+\lambda M'\right)[\phi], \quad \mathbf{r}'\in S, \tag{52}$$

where $\mathbf{r}'$ is a point on the boundary, λ is a complex valued regularization parameter, the operators L and M are provided by Eq. (24), while the other operators are

$$L'[q]=\int_S q(\mathbf{r})\frac{\partial G\left(|\mathbf{r}-\mathbf{r}'|\right)}{\partial n(\mathbf{r})}dS(\mathbf{r}), \quad M'[\psi]=\frac{\partial}{\partial n(\mathbf{r})}\int_S \psi(\mathbf{r})\frac{\partial G\left(|\mathbf{r}-\mathbf{r}'|\right)}{\partial n(\mathbf{r})}dS(\mathbf{r}). \tag{53}$$

The boundary is discretized by a mesh with N_{vert} vertices and N_{el} flat triangular elements. The discretized equation is collocated either at the vertices or at the centers of elements using a constant panel approximation. The boundary integrals over the elements are computed using standard technique in the neighborhood of singularities of the integrands and center point approximations for far field integrals. The singular integrals are computed using the elementary solution fitting technique described in [19]. This results in $N\times N$ linear system of equations, where N is N_{vert} or N_{el}, depending on the collocation method used. The point $\mathbf{r}_*$ is placed at the collocation node, so Eq. (21) provides non-singular values for the normal derivative of ψ at all collocation nodes.

The system is solved iteratively using preconditioned flexible general minimum residual method (fGMRES), which consists of an external loop where at each step one full system matrix-vector multiplication is performed and the internal (preconditioning) loop. In the preconditioning loop an unpreconditioned GMRES is used to solve an approximate linear system. Each iteration in the

internal loop requires one approximate system matrix-vector multiplication. The fast multipole method (FMM) for the 3D Helmholtz equation is used to accelerate the evaluation of the large dense matrix-vector products in the external and in the internal loops. We also use the low accuracy FMM-based preconditioning, where the GMRES in the internal loop of a limited length is executed to converge to relatively large residual value and using low accuracy FMM for the far field. The use of this preconditioner substantially reduces the number of external iterations and, therefore, substantially reduces the dimensionality of the Krylov subspace needed for the solution (which reduces the memory requirements and allows us to handle large problems on conventional desktop workstations).

The FMM used in the BEM solver has the following scaling for the CPU time: for problems with large kD, where D is the maximum size of the computational domain, and k the wave number, this algorithm has theoretical time complexity $O((kD)^3)$ for $kD \gg 10^3$ and $O((kD)^{2+\alpha})$ with a relatively small α for $kD \lesssim 10^3$. These theoretical scalings were validated experimentally on solution of some benchmark problems for $0.0001 < kD < 500$ in Ref. [19]. Also at high frequencies the algorithm has low memory complexity $O((kD)^2 \log(kD))$. At low kD the complexity depends on kD weakly and is determined rather by the size of the system, N, (complexity $O(N \log N)$), not by kD. It also requires $O(N \log N)$ memory. This is in contrast to the scaling of the conventional BEM solvers which have a time complexity of $O\left((kD)^6\right)$ and a memory complexity of $O\left((kD)^4\right)$. It is because of this that the present HRTF calculation for the entire frequency range, and for a discretization that includes the torso too, can be accomplished on conventional desktop workstations.

B. Spherical harmonic analysis

1. Truncation error and error bound

First, let us determine the number of spherical harmonics needed to represent the function $H(\mathbf{s})$ for any given frequency or wavenumber k. This can be deduced from the Gegenbauer plane wave decomposition (35), where the infinite sum over n is truncated to a maximum value $n = p-1$, which provides a p^2 term approximation via an expansion over the spherical functions R_n^m, $n = 0, ..., p-1$;

$m = -n, ..., n$. Using the addition theorem for spherical harmonics, we can rewrite (35) as

$$e^{-ik\mathbf{s}\cdot\mathbf{r}} = 4\pi \sum_{n=0}^{\infty} i^{-n} j_n(kr) \sum_{m=-n}^{n} Y_n^m(\mathbf{s})\, Y_n^{-m}\left(\frac{\mathbf{r}}{r}\right) = \sum_{n=0}^{\infty}(2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s}\cdot\frac{\mathbf{r}}{r}\right) \tag{54}$$
$$= \sum_{n=0}^{p-1}(2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s}\cdot\frac{\mathbf{r}}{r}\right) + \epsilon_p(kr),$$

where $\epsilon_p(kr)$ is the truncation error and P_n are the Legendre polynomials, $|P_n| \leqslant 1$. This error for $n \geqslant p > kr$, for which $j_n(kr)$ are strictly positive functions, can be bounded as

$$|\epsilon_p(kr)| = \left|\sum_{n=p}^{\infty}(2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s}\cdot\frac{\mathbf{r}}{r}\right)\right| \leqslant \sum_{n=p}^{\infty}(2n+1) j_n(kr) < \sum_{n=p}^{\infty}(2n+1) j_n(ka), \tag{55}$$

where a is the maximum possible value of r. Since for computation of H_n^m parameter r is the distance from the scatterer center to a point located on the scatterer surface (see (34)), a is the minimal possible value of range R for which spherical decomposition can be used, so $a = R_{\min}$.

The error bounds shown in the literature for the sum (55) are related to decompositions of plane wave over the wave functions (e.g. [8]). A more strict bound valid for any ka for this sum is found in [16], and is

$$|\epsilon_p| \leqslant \sum_{n=p}^{\infty} \frac{(ka)^n}{(2n-1)!!} < \left(\frac{ka}{2}\right) \sum_{n=p-1}^{\infty} \frac{(ka)^n}{2^n n!} = \frac{e^{ka/2}}{(p-1)!}\left(\frac{ka}{2}\right)^p, \quad p \geqslant 4. \tag{56}$$

This substantially overestimates the p required to obtain a given accuracy $|\epsilon_p|$ for large ka, for which the asymptotic bound [16]

$$|\epsilon_p| \lesssim (2p+1)\, j_p(ka) \lesssim \frac{p^{2/3}}{(ka)^{1/2}} \exp\left(-\frac{1}{3}\left(2\eta_p^{(a)}\right)^{3/2}\right), \quad \eta_p^{(a)} = \frac{p - ka + 1/2}{(ka)^{1/3}}, \quad p > ka \gg 1, \tag{57}$$

is more appropriate. This formula results in the following estimate of p for prescribed error $\epsilon = |\epsilon_p|$ [16]

$$p \gtrsim ka + \frac{1}{2}\left(3\ln\frac{1}{\epsilon} + \frac{1}{2}\ln(ka)\right)^{2/3} (ka)^{1/3}, \quad ka \gg 1. \tag{58}$$

We also note a more simplistic approximation of $p(ka)$ for any ka used in the literature (e.g. [6])

$$p = ka + 4(ka)^{1/3} + 3. \tag{59}$$

The error of this approximation growth with ka (as it misses a slow growth $\sim \ln(ka)$ of the coefficient near $(ka)^{1/3}$, which actually can be estimated using Eq. (57) (which is a more accurate estimate than (58)) with p from Eq. (59).

We also note that for computation of the HRTF for finite range we have more restrictive conditions on p which increase it, due to the requirement of the accuracy of source field expansion (43). However, as it is shown in [15] in the derivation of error bounds for the multipole-to-local translation operator for range $R > 2a$ the order of magnitude of this error will be the same as for the expansion of the plane wave. So if we restrict the ranges considered with conditions $R > 2R_{\min}$, which are typical for experimental measurements of the HRTF, then the above formulae for error bounds can be used. Otherwise more restrictive formulae, which include parameter R/a should be used, which also can be found in [15],[16].

2. Computation of expansion coefficients

As soon as the required bandwidth for a given error bound, frequency, and domain size for the spherical harmonic representation are known, the p^2 coefficients $H_n^{(pw)m}$ can be computed using the direct formula (36). Note that computation of functions $R_n^m(\mathbf{r})$ and their directional derivatives $\mathbf{n}(\mathbf{r}) \cdot \nabla R_n^m(\mathbf{r})$ is done via a standard recursive procedure used in the FMM for the Helmholtz equation, which is inexpensive with a computational complexity of $O(Np^2)$.

We note also that in Eq. (36) $R_n^{-m}(\mathbf{r})$ are regular functions and have the same spatial frequency as the surface functions $\psi(\mathbf{r};\mathbf{r}_*)$, $\mathbf{n}(\mathbf{r})$, and $q(\mathbf{r};\mathbf{r}_*)$ expanded in the integral. Since the discretizations are relatively dense with at least 5-10 elements per wave length and a constant panel approximation is used, it is consistent to use just a single value of functions $R_n^m(\mathbf{r})$ at the collocation point. This is also consistent with the accuracy of the far field integrals computation in the FMM accelerated BEM [19]. Therefore, the discrete analog of Eq. (36) is

$$H_n^{(pw)m} = 4\pi i^{-n} \left\{ 2R_n^{-m}(\mathbf{r}_*) + \sum_{j=1}^{N} w_j \left[\psi_j \mathbf{n}_j \cdot \nabla R_n^{-m}(\mathbf{r}_j) - q_j R_n^{-m}(\mathbf{r}_j) \right] \right\} \tag{60}$$

where $n = 0, ..., p-1, \quad m = -n, ..., n$, and w_j is the weight of surface quadrature, and $\psi_j, q_j, \mathbf{n}_j$ are the values of the respective functions at the collocation points $\mathbf{r}_j$.

Note that the weights w_j should be selected consistent with the BEM implementation used. For example, in [19] the surface integrals are expanded over the bases of monopoles and dipoles, while the weights for singular elements were computed via a method of fitting of test functions (in fact, such weights take into account correction over the neighborhood of singular integrals). A similar procedure can be applied for accurate computation of the surface integral (36) for the singularity

near the ear location. In fact, this procedure is simpler, since only one test solution $e^{ik\mathbf{s}_*\cdot\mathbf{r}}$, where $\mathbf{s}_*$ is a unit vector orthogonal to the normal at the ear location can be used. In this case the normal derivative of this function is zero at the ear location. Applying Green's identity for this solution, which is regular inside the head, we can express the weight for the vicinity of the singularity via all other weights and obtain the final formula for computation in the form

$$H_n^{(pw)m} = 4\pi i^{-n}\left\{R_n^{-m}\left(\mathbf{r}_*\right) + \sum_{j=1}^{N} w_j \left[\begin{array}{c} \psi_j \mathbf{n}_j\cdot\nabla R_n^{-m}\left(\mathbf{r}_j\right) + 2ik\left(\mathbf{n}_j\cdot\mathbf{s}_*\right)e^{ik\mathbf{s}_*\cdot\left(\mathbf{r}_j-\mathbf{r}_*\right)}G\left(\left|\mathbf{r}_j-\mathbf{r}_*\right|\right)R_n^{-m}\left(\mathbf{r}_*\right) \\ -q_j\left(R_n^{-m}\left(\mathbf{r}_j\right) - e^{ik\mathbf{s}_*\cdot\left(\mathbf{r}_j-\mathbf{r}_*\right)}R_n^{-m}\left(\mathbf{r}_*\right)\right) \end{array}\right]\right\}, \tag{61}$$

where the nearly singular terms close to the ear location ($j = j_*$, $\mathbf{r}_{j_*} = \mathbf{r}_*$) become small, since for $j = j_*$ we have $\mathbf{n}_{j*}\cdot\mathbf{s}_* = 0$ and $R_n^{-m}\left(\mathbf{r}_{j*}\right) - e^{ik\mathbf{s}_*\cdot\left(\mathbf{r}_{j*}-\mathbf{r}_*\right)}R_n^{-m}\left(\mathbf{r}_*\right) = 0$ (the corresponding terms in the sum for $j = j_*$ should be neglected). We checked numerically that while Eqs (60) and (61) provide close results, the accuracy of computation of the zero mode $H_0^{(pw)0}$ is substantially higher (e.g. 0.01% vs 4% at low frequencies), when using corrected Eq. (61), while for the other modes the accuracy is about the same for both formulae.

Note that computation of the spherical spectrum is a relatively fast procedure for moderate truncation numbers p and can be performed straightforward. However for large high frequency cases computations using Eqs (60) and (61) may take time comparable with the solution of the BEM. Indeed, to compute p^2 coefficients the complexity will be $O(p^2N)$. This cost can be greatly reduced to $O(p_{\min,FMM}^2 N)$ where $p_{\min,FMM}$ is the truncation number used in the FMM for the smallest box in the tree, if the FMM is employed (p can be an order of magnitude larger than $p_{\min,FMM}$ and savings can be of the order of 100 times). Indeed, Eq. (26) shows that the problem of computation of the HRTF after solution of the BEM reduces to summation of monopoles and dipoles located at the surface collocation nodes. To obtain the spherical spectrum only the multipole expansion about the center of the head needs to be generated. This can be done with the FMM, which after discretization brings Eq. (26) to the form

$$H\left(\mathbf{s}\right) = 4\pi R e^{-ikR}\sum_{n=0}^{p-1}\sum_{m=-n}^{n} A_n^m S_n^m\left(R\mathbf{s}\right), \tag{62}$$

see also the expansion (43)), where A_n^m are the multipole expansion coefficients obtained via the FMM. Taking a limit as $R \to \infty$ as before, we can find that the plane wave spherical spectra

coefficients are simply related to A_n^m

$$H_n^{(pw)m} = \frac{4\pi}{k} i^{-n-1} A_n^m. \tag{63}$$

According to Eq. (61) A_n^m are the multipole expansion coefficients about $\mathbf{r} = \mathbf{0}$ of the field of N dipoles with momentum $\mathbf{n}_j$ and strength $D_j = w_j \psi_j$ located at $\mathbf{r} = \mathbf{r}_j$, $j = 1, ..., N$, $N-1$ monopoles of strength $M_j = -w_j q_j$ located at $\mathbf{r} = \mathbf{r}_j$, except of $\mathbf{r} = \mathbf{r}_*$, and a monopole at $\mathbf{r} = \mathbf{r}_*$, the strength of which is

$$M_* = 1 + \sum_{j \neq j_*} w_j \left[2ik \left(\mathbf{n}_j \cdot \mathbf{s}_* \right) e^{ik\mathbf{s}_* \cdot (\mathbf{r}_j - \mathbf{r}_*)} G\left(|\mathbf{r}_j - \mathbf{r}_*| \right) + q_j e^{ik\mathbf{s}_* \cdot (\mathbf{r}_j - \mathbf{r}_*)} \right], \tag{64}$$

and can be computed using direct summation at $O(N)$ cost. To be consistent with the FMM described in [19] we build multipole $S-$expansions for any source box up to level 2, and then apply the $(S|S)-$ translation procedure to the center $\mathbf{r} = \mathbf{0}$ followed by consolidation of all expansions, which results in the full set of coefficients $\{A_n^m\}$.

V. NUMERICAL TESTS

We first conducted some tests on simple geometrical models, for which the HRTF can be obtained either analytically or semi-analytically without using the BEM, reciprocity principle, and decompositions/synthesis over spherical basis functions and compare it with the present method which uses all these. These tests were used to validate the method and figure out the settings appropriate for computations with more complex geometries. After that we performed several tests on a complex head geometry in the presence and absence of the torso and compared this with the results from simpler geometries. Also, we tested the BEM solutions against analytical results for arbitrary meshes, such as by computing the field generated by a source placed inside the body. These tests provided us with the confidence that the method can be used for modeling more complex situations, and we then compared the results of the computations with the HRTF measured in experiments.

A. Comparison with single sphere analytical solution

The analytical expression for the scattering from a single sound hard sphere of radius a can be found elsewhere [20], and [10] provide analysis and discussion of the HRTF for the case of a sphere

placed in the incident field generated by a monopole source:

$$H(\mathbf{s}) = -\frac{1}{4\pi ka^2}\sum_{n=0}^{\infty}(2n+1)\frac{h_n(kR)}{h_n'(ka)}P_n(\mathbf{s}\cdot\mathbf{s}_*) \tag{65}$$
$$= -\frac{1}{ka^2}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{h_n(kR)}{h_n'(ka)}Y_n^{-m}(\mathbf{s}_*)Y_n^m(\mathbf{s}), \quad \mathbf{s}_* = \frac{\mathbf{r}_*}{a},$$

where h_n' is the derivative of h_n and $\mathbf{s}_*$ is the direction to the ear location. Comparing this with Eq. (29), we can see that the expansion coefficients for $H(\mathbf{s})$ are

$$H_n^m = -\frac{1}{ka^2}\frac{h_n(kR)}{h_n'(ka)}Y_n^{-m}(\mathbf{s}_*), \quad n = 0,1,..., \quad m = -n,...,n. \tag{66}$$

According to Eq. (45), we can determine the spherical spectrum for the plane wave incident field

$$H_n^{(pw)m} = \frac{4\pi}{ik}i^{-n}\frac{H_n^m}{h_n(kR)} = -\frac{4\pi}{i(ka)^2 h_n'(ka)}i^{-n}Y_n^{-m}(\mathbf{s}_*), \tag{67}$$

which also provides an analytical expression for $H^{(pw)}(\mathbf{s})$ according to Eq. (29):

$$H^{(pw)}(\mathbf{s}) = -\frac{4\pi}{i(ka)^2}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{i^{-n}}{h_n'(ka)}Y_n^{-m}(\mathbf{s}_*)Y_n^m(\mathbf{s}) = \frac{i}{(ka)^2}\sum_{n=0}^{\infty}(2n+1)\frac{i^{-n}}{h_n'(ka)}P_n(\mathbf{s}\cdot\mathbf{s}_*). \tag{68}$$

In [19] we performed an extensive study of the convergence and accuracy of solution for scattering from a sphere to test the BEM we use in range of $ka = 0.001 - 50$. Note that for the head radius $a = 8.25$ cm and frequency $f = 20$ kHz, we have $ka = 30.225$ ($c = 343$ m/s). So the range for which the BEM was tested covers the range of human audible frequencies.

Fig. 2 compares the computed and analytical plane wave HRTF $H^{(pw)}(\mathbf{s})$ for $ka = 30$ plotted as functions of the angle of incidence $\theta_{in} = \arccos(\mathbf{s}\cdot\mathbf{s}_*)$ ($\mathbf{s}_* = 3^{-1/2}(1,-1,1)$). Figs 3 - 4 compare the computed and analytical spherical power spectra (some modes). Eq. (59) provides in this case truncation number $p = 46$, while equation (57) bounds the error of truncation for this p as 4%, which is more or less consistent with the overall accuracy of the standard BEM computations and reasonable for comparisons with experiments. The numerical results were obtained using vertex collocation BEM with 48602 vertices and 97200 elements discretizing the sphere surface which provided not less than 6 elements (in size) per wavelength (parameter $kd_{\max} = 0.94$, where $d_{\max}$ is the maximum size of the element). The fGMRES solution converged in 10 outer loop iterations to residual 10^{-4}, and for the full matrix-vector product the prescribed accuracy of the FMM was also 10^{-4}. For the inner (preconditioning) loop the FMM with accuracy $5\cdot 10^{-2}$ was used to converge to

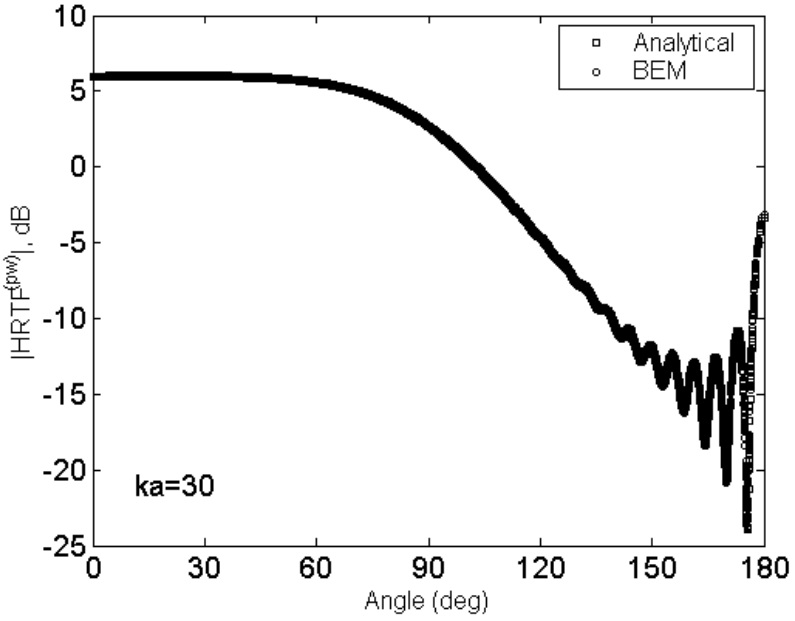

FIG. 2: Amplitude HRTF for plane wave incident field for a sphere at $ka = 30$, computed using analytical series (68) (the squares) and the BEM/spherical harmonic representation. The graph plots the HRTF computed for 48602 directions $\mathbf{s}$. The computed results correlate well and can be plotted as a function of a single variable - the angle of incidence $\theta_{in} = \arccos(\mathbf{s} \cdot \mathbf{s}_*)$.

0.5 residual error, which normally took 3-5 iterations, and never exceeded 10 iterations. The actual errors in the spherical harmonic spectrum (of size p^2 =2116) were ϵ_∞ =1.6% and ϵ_2 =0.57%, while the errors in the HRTF were ϵ_∞ =1.1% and ϵ_2 =0.59% (HRTF was computed for 48602 directions corresponding to each vertex). Here ϵ_∞ and ϵ_2 are the relative errors in the L_∞ and L_2 norms, defined as

$$\epsilon_\infty = \frac{\left\|H^{comp} - H^{exact}\right\|_\infty}{\left\|H^{exact}\right\|_\infty}, \quad \epsilon_2 = \frac{\left\|H^{comp} - H^{exact}\right\|_2}{\left\|H^{exact}\right\|_2}, \tag{69}$$

$$\|H\|_\infty = \max_{j=1,\ldots K} |H_j|, \quad \|H\|_2 = \left(\frac{1}{K}\sum_{j=1}^{K} |H_j|^2\right)^{1/2},$$

where the superscripts mark computed and analytical solution of any function H sampled at K points. We also computed the HRTF for some finite ranges R from $H_n^{(pw)m}$ using the above theory for range dependent HRTF and found that the errors are of the same order as for the plane wave HRTF (e.g. for $R/a = 5$, we have ϵ_∞=1.3% and ϵ_2=0.6%).

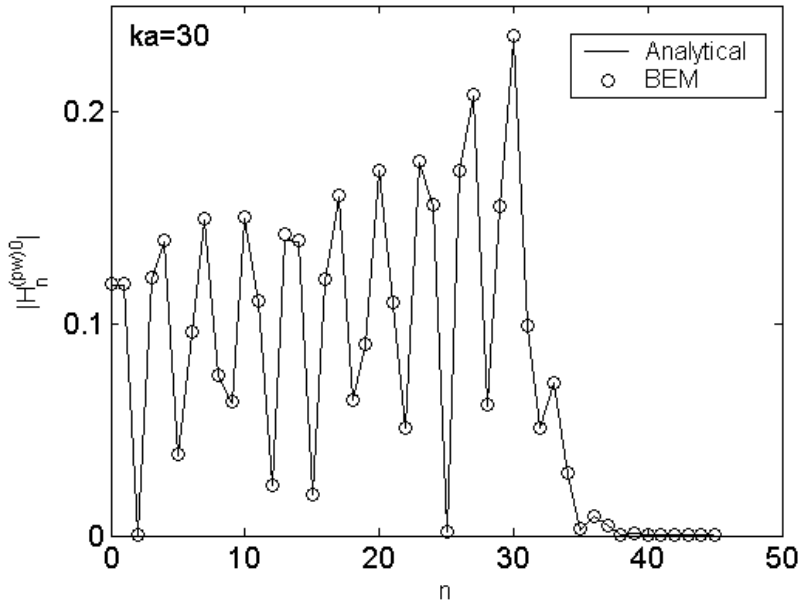

FIG. 3: Amplitudes of complex harmonics $\left|H_n^{(pw)m}(\mathbf{s}_*)\right|$ for $m = 0$ as functions of n found analytically (67) (points are connected by the solid line) and computed using the BEM and Eq. (60) (the circles) for $ka = 30$ and $\mathbf{s}_* = (0.5774, -0.5774, 0.5774)$.

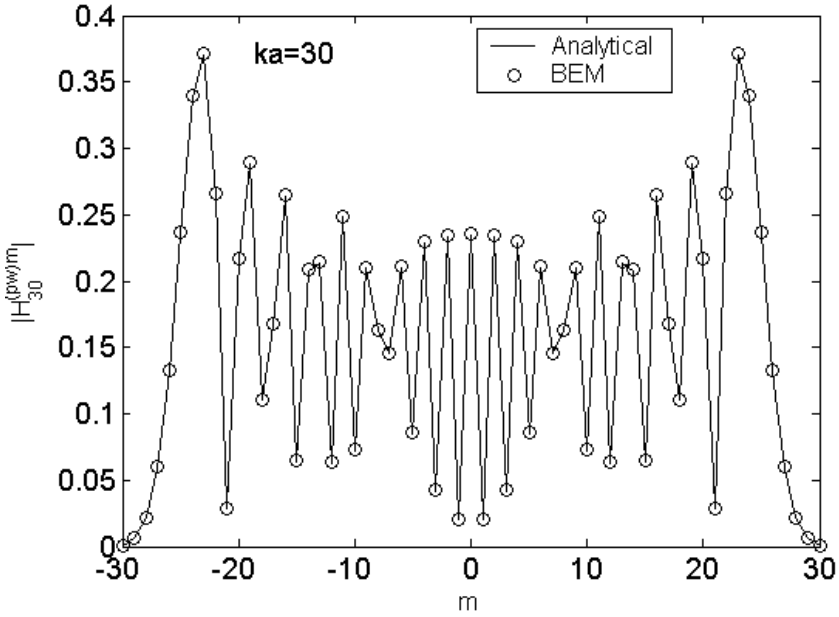

FIG. 4: The same as Fig. 3, but for $n = 30$ and varying m.

The importance of Burton-Miller formulation for accurate solution for a sphere was demonstrated in [19], where we showed that the Green's identity formulation at some values of ka may converge to the wrong solution, corresponding to internal resonances of the sphere. In the present study we used $\lambda = 0.03i/k$ in Eq. (52).

B. Tests for head mesh

1. Mesh generation

Computation of scattering off complex shapes, such as human head, require quality surface meshes. Usual requirements for such a mesh is that the size of a surface element (we used triangular elements) should be much smaller than the acoustic wavelength. The use of the version of the FMM accelerated BEM described in [19] requires high quality meshes, which ideally should have a uniform distribution of the mesh vertices over the surface. Moreover, the valence of each vertex (the number of triangles it belongs to) should be as close to 6 as possible, so that the triangles are as close to equilateral as possible.

On the other hand, the data coming from laser scans of heads and meshed using standard triangulation techniques, such as those based on the Delaunay triangulation, are acceptable for graphics, but have a poor quality for the BEM. Several such meshes were available to the researchers. To make them appropriate for the BEM we developed a remeshing technique which produced meshes of the desired quality. Note that in the literature several techniques for generation of quality meshes are available (e.g. in the CGAL project). A detailed description of our technique is beyond the scope of the present paper, and we only mention that the technique is based on the 3D radial basis function (RBF) interpolation combined with a relaxation technique, which allows the mapping of meshes between topologically equivalent objects. Details of the technique can be found in our technical report [27]. In the case of a human head the original data/mesh was topologically equivalent to a sphere. So, we first generated a high quality mesh on the sphere, which is a relatively easy task and then used the interpolation/relaxation process to provide a mapping between the two meshes. We note that for high frequency computations up to 20 kHz the head mesh contained of order 10^5 elements, and the maximum value of the computational parameter kD was 181. RBF technique for such a large data set was enabled by the use of the FMM for the biharmonic equation and $O(N \log N)$ preconditioning technique as described in [17].

The resulting mesh was checked for quality using several geometrical parameters (minimal and maximum size of the triangles, aspect ratios, etc.) and for several elements an additional subdivision procedure was applied. So it was guaranteed that the surface is discretized with not less than 6 elements per wavelength for the highest frequency computed. Fig. 5 shows the computational mesh for the Neumann KU 100 Dummy Head model (referred to as "Fritz"). This mesh contains 76,335 vertices and 152,666 triangles. We also did computations with two Kemar head meshes which contain 49,585 vertices and 99,166 triangles ("small pinnae" mesh) and 58,800 vertices and 117,596 triangles ("large pinnae" mesh), respectively (see Fig. 15, picture in the center, and pinnae (the right column)). Note that the "small pinnae" Kemar mesh was generated from a laser scan, while the "large pinnae" mesh was obtained in the following way. A mesh of good quality was generated from the CT scan of the pinnae [11]. The pinna areas of the original Kemar head were cut off and the pinnae mesh was positioned at an angle corresponding to measurements in the CIPIC database [2]. After that a mesh "stitching" and relaxation technique was used to merge the meshes. Below, if it is not mentioned specifically, the term "Kemar head mesh" means the original model with the "small pinnae".

2. Performance

Both vertex collocation and panel-center collocation methods were used in the BEM. While both of them produced consistent results, the latter method was 2-3 times more accurate (though a little bit slower, since the number of unknowns is twice as large). This was checked by comparison with analytical solution for a given mesh (the solution was generated by a source placed inside the head, and then the external Neumann problem was solved and the accuracy was checked). Since we used the same meshes for several frequencies, the low frequency cases produced much smaller errors due to substantial oversampling. Even for the highest frequencies the relative errors ϵ_∞ and ϵ_2 computed using Eq. (69) did not exceed a few percent, which we considered acceptable for evaluation of the HRTF and for comparison with the experimental data.

Another issue important for the current method of the HRTF computation is that the solution obtained assumes that the image source of the same intensity approaches a smooth boundary. In the case of center panel collocation method this is correct, since the panels are flat and the normal derivative at the center of the panel corresponding to the ear location tends to zero. In

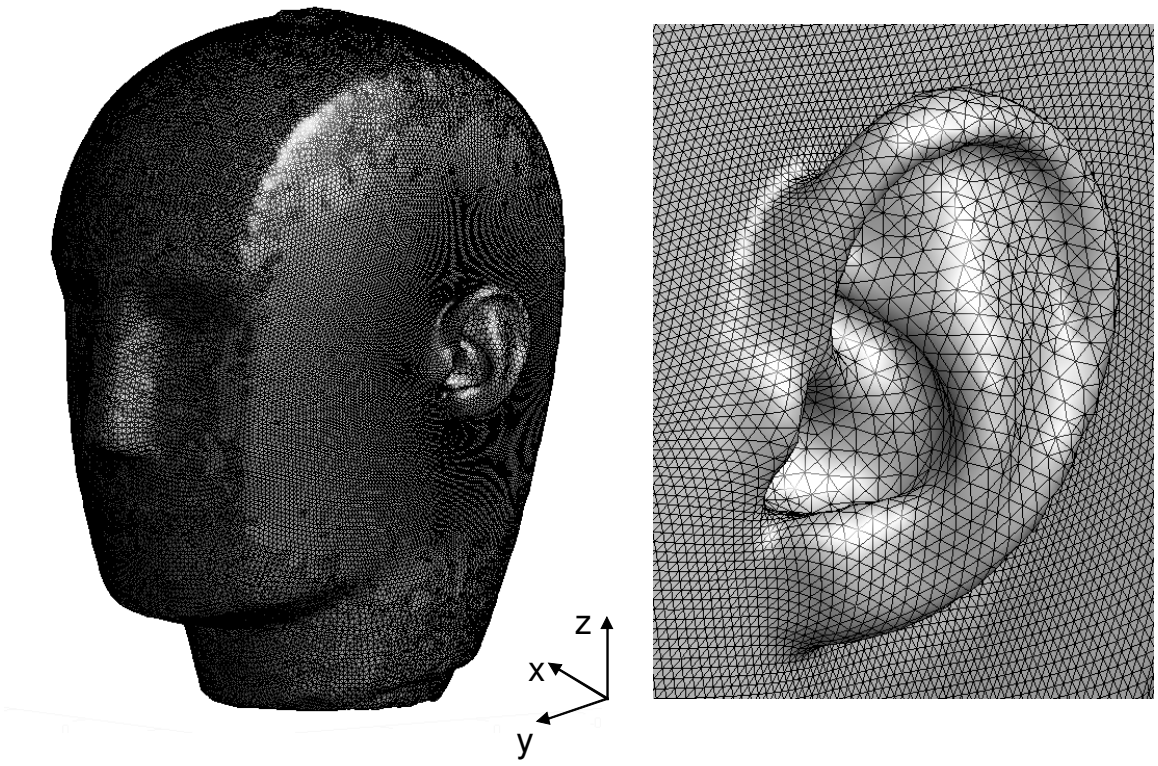

FIG. 5: The mesh used for computation of the HRTF using the FMM accelerated BEM for the Fritz model. The mesh contains 76335 vertices and 152666 triangular faces. The plot on the right zooms in the pinnae region of the plot on the left.

the case of vertex collocation each vertex has some solid angle different from 2π. Unless the vertex corresponding to the ear location subtends a solid angle of 2π (means all panels containing this vertex are located in the same plane) the symmetry of the solution will be broken and some systematic shift in the HRTF can be expected (and which was observed in our computations).

The most expensive computational step (solution of the boundary value problem) was executed on a four core PC (Intel Core 2 Extreme QX6700 2.66 GHz, 8 GB RAM) using parallelization of the FMM with parallelization efficiency about 90%. The settings for the FMM accuracy and preconditioned fGMRES process were the same as for the test case of a spherical head reported above. Computation of the coefficients $H_n^{(pw)m}$ is a much faster procedure than the BEM solution. The table below shows the wall clock time (includes I/O time for each procedure) and memory required for several cases using the center panel collocation method.

Frequency (kHz)	0.172	0.517	1.034	2.067	4.996	9.991	14.987	20.155
BEM solution (s)								
Fritz	131	125	159	334	558	919	1254	1483
Kemar	62	64	86	160	341	417	520	628
Computation of $H_n^{(pw)m}$ (s)								
Fritz	1.2	1.8	2.9	5.6	13	35	64	103
Kemar	0.9	1.3	1.8	3.3	8	21	36	58
Memory (GB)								
Fritz	2.07	2.10	2.14	2.20	2.30	2.37	2.44	2.60
Kemar	3.12	3.16	3.20	3.25	3.30	3.32	3.34	3.43

TABLE II: Performance for the Fritz and Kemar (head) models

It is seen that for the method used enables computation of even the highest frequency case for the time less than half of hour for the Fritz model and about 10 minutes for the Kemar head. In this case for Fritz the fGMRES process required 30 iterations (residual $<10^{-4}$) with about 20 seconds per matrix-vector product and 317 inner loop (preconditioning) iterations with about 2.5 seconds per matrix-vector product, while for Kemar the process converged to residual $<10^{-4}$ for 21 iterations with 7.3 seconds per matrix-vector product and 203 inner loop iterations with 2.1 seconds per matrix-vector product. The table also shows relatively slow increase of the solution time with frequency for a fixed mesh (for high frequencies f the solution time scales approximately as $t \sim f^{0.75}$ in the computational range). Faster times for the Kemar head are due to three factors. First, the mesh for Kemar had approximately 1.5 times less panels than for Fritz. Second, the size of Kemar was smaller (maximum kD for Fritz was 181, while for Kemar 157). Since the algorithm is scaled as $(kD)^{\alpha}, 2 < \alpha < 3$, this gives another factor of 1.3-1.5. Third, due to the difference in meshes and kD's the maximum level of the octree in the FMM was 6 for Fritz and 5 for Kemar, which caused a different balance between the dense and sparse matrix vector product parts in the FMM and also a difference in the consumed memory, which was higher for the Kemar case due to storage of the larger sparse matrix. Finally, we can mention that the total execution time for the run, which includes BEM solution, and computation of the spherical spectrum over 117 frequencies equispaced from 172 to 20,155 Hz took about 32 hours for Fritz and 14 hours for Kemar head.

3. Spherical spectra

Figure 6 presents the power spherical spectra of the plane wave HRTF (modules of coefficients $H_n^{(pw)m}$) for Fritz computed for four frequencies $f = 518, 2071, 10008$, and 20016 Hz, which are stored subsequently in groups according to their index n (within each group the order m is changing from $-n$ to n). This also can be plotted as a function of single index $i = (n+1)^2 + n - m$, $n = 0, ..., p-1$, $m = -n, ..., n$, $i = 1, ..., p^2$. The spectrum truncation number p was selected according to Eq. (59).

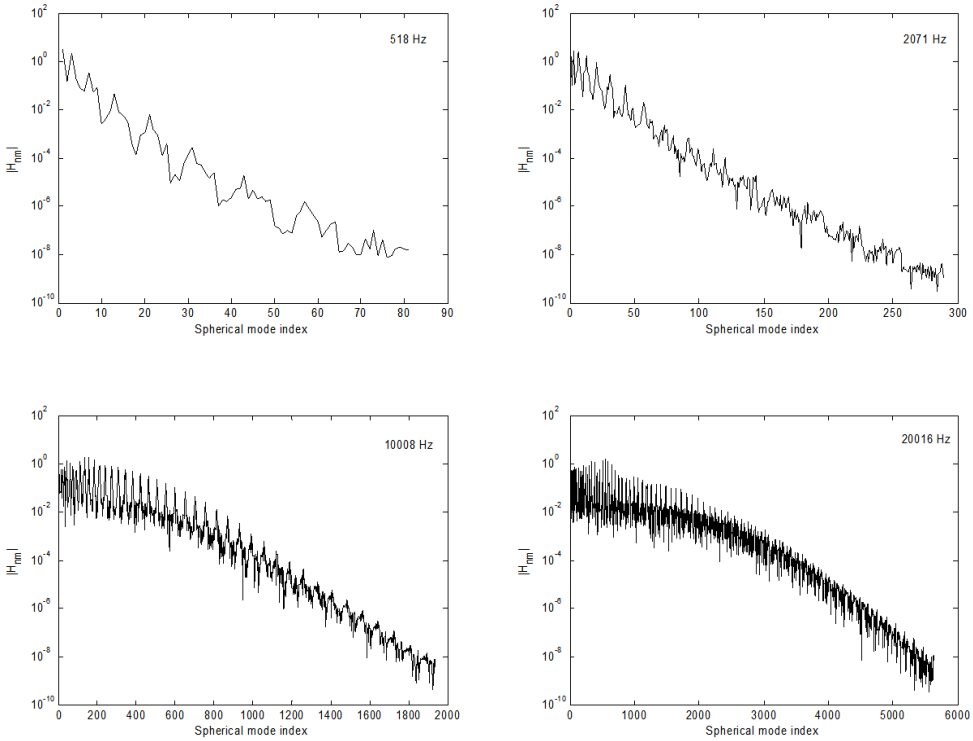

FIG. 6: Power spectra of spherical representation of the HRTF, $\left|H_n^{(pw)m}\right|$ as functions of a single spherical mode index $i = (n+1)^2 + n - m$, $n = 0, ..., p-1$, $m = -n, ..., n$ for different frequencies (Fritz head model).

It is seen that the truncation occurs in the region of spectral (exponential) convergence, and the values of the truncated modes are several orders of magnitude smaller than the values of the modes stored for the HRTF representation. Such representation is valuable for efficient storage and high quality interpolation of the HRTF. Moreover, the spherical harmonics have the property that their degree n is preserved under rotation. This means that each subspace corresponding to a value of n is rotated independently of the other subspaces. Thus truncation in the region of exponential

convergence with respect to n guarantees uniformity of the expansion (29) for any spherical angle. It may also be noted that an efficient rotation transform (see [15]) can be applied to obtain H_n^m for *any rotation* of the head:

$$\{H_n^m\}_{rot} = \mathbf{Rot}\,(\alpha,\beta,\gamma)\,\{H_n^m\}_{stnd}\,, \tag{70}$$

where α, β,and γ are the Euler rotation angles, **Rot** is the rotation transform matrix, which can be decomposed to a direct sum of rotation operators for each subspace n, and subscripts "rot" and "stnd" refer to the sets of coefficients in the rotated and standard (computational) reference frames. Thus, the present representation may be particularly useful in the use of HRTFs in virtual reality applications.

Figure 7 shows computations of the finite range HRTF spectra and their comparisons with the plane wave HRTF for frequency $f = 10008$ Hz. According to Eqs (12), (13), (29), and (45) the coefficients of the HRTF coefficients are related to coefficients H_n^m and $H_n^{(pw)m}$ as

$$(HRTF)_n^m = 4\pi R e^{-ikR} H_n^m = i^{n+1} e^{-ikR} kR h_n\,(kR)\, H_n^{(pw)m}, \tag{71}$$

so $(HRTF)_n^m \to H_n^{(pw)m}$ at $R \to \infty$. For the Fritz model the closest and the farthest point of the surface from the origin are located at distances $a_{\min} = 0.0705$ m and $a_{\max} = 0.1536$ m, respectively. We varied the range R from 0.2 to 1 m to see the effect of the range on the spectrum. There were no visible differences between the power spectra $|(HRTF)_n^m|$ and $\left|H_n^{(pw)m}\right|$ for $R >$ 0.4 m, while for $R = 0.2$ m some small effect was seen in the region of exponential decay of the magnitude. Nonetheless the spectrum is characterized also by the phase angle of the complex coefficients $(HRTF)_n^m$, and to show the effect of the range we plotted the difference in this angle between the finite range and plane wave HRTF spectra:

$$\Delta\phi_n^m = \left(\phi\left((HRTF)_n^m\right) - \phi\left(H_n^{(pw)m}\right)\right)(\mathrm{mod}\,\pi)\,. \tag{72}$$

Figure 7 shows that this difference is small for modes with very low n, while for larger n it is substantial for $R = 0.2$ m and appreciable even for $R = 1$ m (according to Eqs (71) and (72) it does not depend on the order m).

a. Low frequency asymptotics At very low frequencies the wavelengths are larger than or comparable to the head size, one can expect that scattering from an arbitrary head model should be close to that of an object of simple shape, such as a sphere or an ellipsoid. We performed a

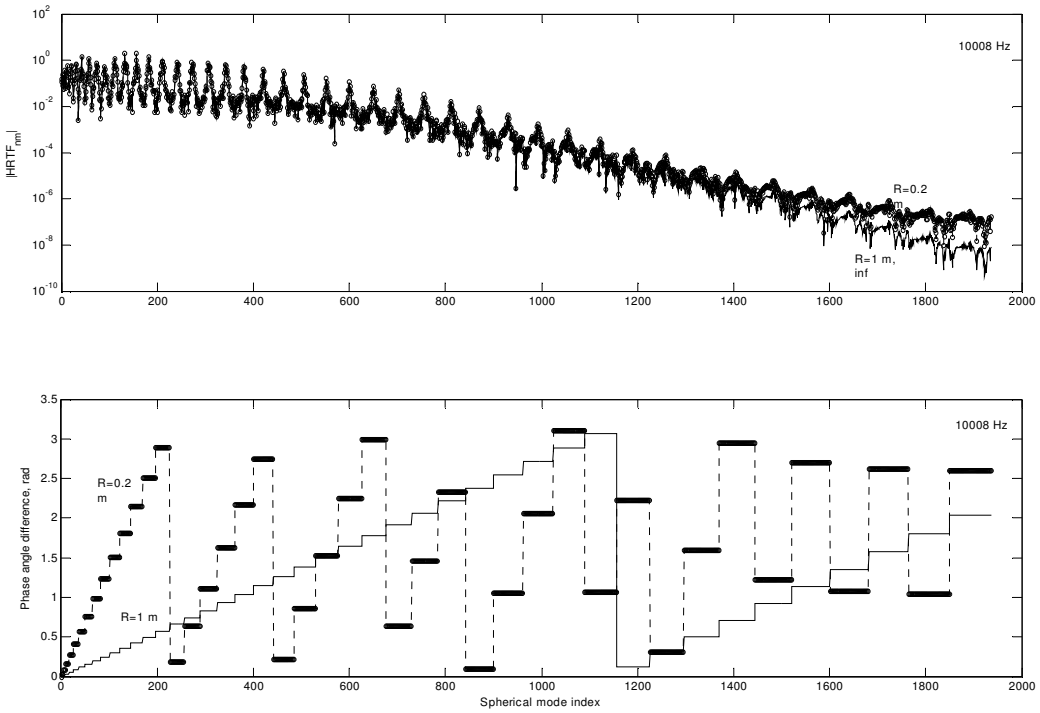

FIG. 7: The power spherical spectra of the HRTFs (on the top) and the difference in phase angles between the finite range and the plane wave HRTF spectra (on the bottom) for the Fritz model at frequency 10008 Hz for ranges $R = 0.2$ m and $R = 1$ m.

set of low frequency runs for the Fritz model in the range 20-1000 Hz with step 20 Hz to compare them with the analytical expressions for scattering off a sphere. The coefficients $H_n^{(pw)m}$ for the sphere are given by Eq. (67). It shows that $H_0^{(pw)0}$ does not depend on the ear location, since $Y_0^0(\mathbf{s}_*) = (4\pi)^{-1/2} = \text{const.}$ (see definition (30)). Since the spherical Hankel functions can be expressed via elementary functions, we have

$$H_0^{(pw)0} = -\frac{(4\pi)^{1/2}}{i\,(ka)^2\, h_0'\,(ka)} = (4\pi)^{1/2}\,\frac{e^{-ika}}{1-ika}. \tag{73}$$

An expansion of this function in a series in ka for $ka \ll 1$ is

$$H_0^{(pw)0} = (4\pi)^{1/2}\left[1 - \frac{1}{2}\,(ka)^2 - \frac{1}{3}i\,(ka)^3 + O\left((ka)^4\right)\right]. \tag{74}$$

In fact, this formula can be used to determine an effective radius, $a_{eff}^{(low)}$, of the head model at low frequencies. Indeed, the principal term here does not depend on the frequency, and so the limiting

value is $(4\pi)^{1/2}$. The second or the third term can be then used to determine the value. Using the latter quantity we obtain

$$a_{eff}^{(low)} = \frac{1}{k}\left[-\frac{3\,\mathrm{Im}\left\{H_0^{(pw)0}\right\}}{(4\pi)^{1/2}}\right]^{1/3} \qquad (ka_{eff}^{(low)} \ll 1), \tag{75}$$

where $H_0^{(pw)0}$ is a computed coefficient for a given head model.

According to Eq. (67) all other coefficients tend to zero as $ka \to 0$ and depend on the ear location. We can note, however, that if $(ka)_* \ll 1$ is some reference value, then for a sphere

$$\frac{H_n^{(pw)m}(ka)}{H_n^{(pw)m}((ka)_*)} = \frac{(ka)_*^2\, h_n'((ka)_*)}{(ka)^2\, h_n'(ka)}, \tag{76}$$

and this does not depend on the ear location.

Figure 8 shows a comparison between the computations of the spherical mode amplitudes for Fritz and the analytical solution for a sphere of radius (75) ($a_{eff}^{(low)} \approx 0.12$ m). The plot on the left shows that Eq. (73) describes the computed data well for $ka < 0.5$, which corresponds to frequencies below 220 Hz. The limiting analytical value $H_0^{(pw)0}(0) = (4\pi)^{1/2}$ is matched with the computed value up to 4 digits, or 0.01% in the relative error. While the qualitative behavior is the same as for a sphere for larger frequencies (up to 1 kHz) and the imaginary part of $H_0^{(pw)0}$ is well described by Eq. (73) even quantitatively, the real part shows substantial deviation from the spherical solution. Presumably, this is related to ellipsoidal distortion of the head shape (indeed for this model we have $a_{\max}/a_{\min} \approx 2$), since the wavelength in this range is still large compared with the pinna. The plot on the right shows that the normalized amplitudes match the dependence (76) well at the low frequency limit. Indeed, while the dependence of the mode amplitude on the mode order m is clearly seen for $ka > 1$, at low frequencies such a dependence vanishes and a good match with the spherical case is observed.

b. Comparison with sphere In figure 9 a comparison between the computed data $H_0^{(pw)0}$ for the Fritz and Kemar head models and the analytical expression (73) for a sphere is shown. This expression depends only on ka, and so is universal, while for scaling of data for the non-spherical shapes different values of a between two values $a_{\min}$ and $a_{\max}$ were selected. It is seen that qualitatively the solution for the sphere matches the computed data. However, if $a = a_{eff}^{(low)}$ is used for scaling the frequency of oscillations of the real and imaginary parts of $H_0^{(pw)0}$ as functions of ka are substantially lower than that for the sphere. A good match in this frequency is observed

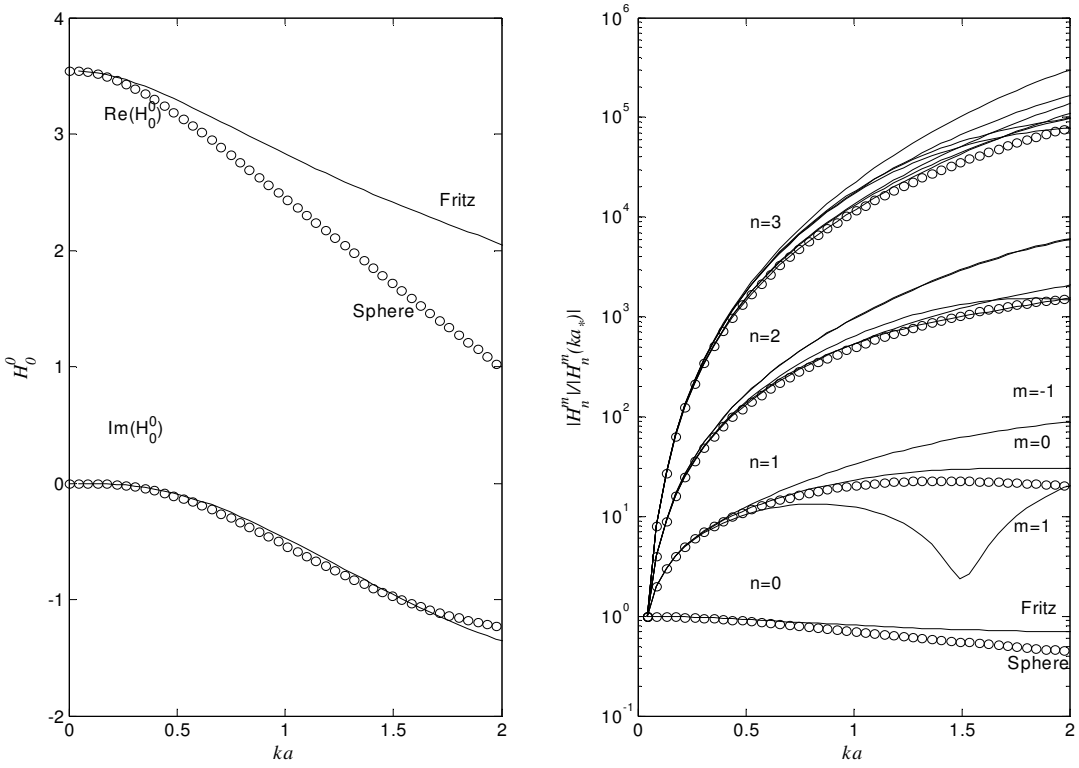

FIG. 8: Spherical mode amplitudes of the plane wave HRTF $H_n^{(pw)m}$ at low frequencies computed for the Fritz model (the solid curves) and for the sphere of equivalent radius a (analytical solution: the circles), determined by Eq. (75).

if the minimal radius of the head, $a = a_{\min}$ is used instead. $H_0^{(pw)0}$ represents the mean value of the plane wave HRTF over all directions (see Eq. (32)). So the frequency dependence of the zero mode qualitatively coincides with that for a sphere. The quantitative difference is explained by shape distortions from the sphere. Indeed, if the shape were represented as an expansion over the spherical harmonics, then each mode would provide a contribution to the mean due to interaction with the conjugate mode (mode of the same degree n, but negative order $-m$) of the acoustic wave. We also can see that the use of $a_{eff}^{(low)}$ as an effective radius is beneficial only for the imaginary part of $H_0^{(pw)0}$ at low frequencies, while the frequency dependence of the real part of this coefficient is modeled well even at low frequencies with $a = a_{\min}$. This figure also shows a difference in scattering from the Fritz and Kemar shapes independently on scaling.

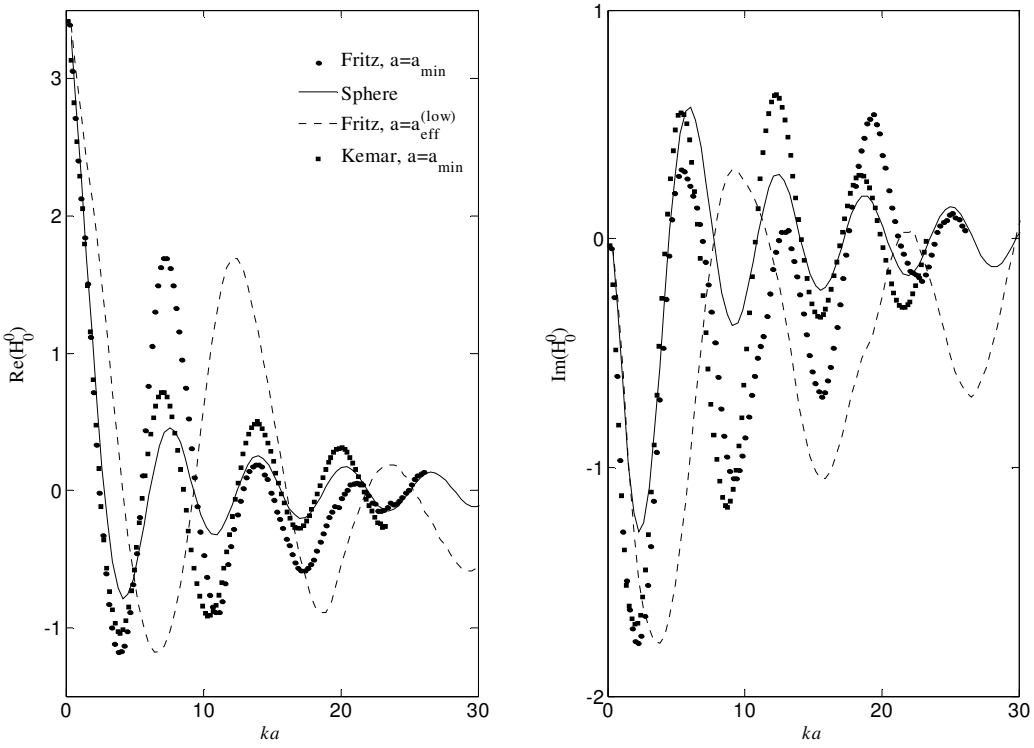

FIG. 9: A comparison of the computations for the Fritz and Kemar heads and spherical analytical model results for the real and imaginary parts of complex amplitude $H_0^{(pw)0}$ on parameter ka. The dependence for the sphere is universal, while different scalings a are used for computed data: for Fritz: $a_{\min} = 0.0705$ m and $a_{eff}^{(low)} = 0.12$ m and for Kemar $a_{\min} = 0.063$ m (see text for definitions of these parameters).

4. HRTF

Figure 10 shows a comparison between the plane wave HRTF for the Fritz model and for a sphere of radius $a_{\min}$. For plotting purposes the HRTF was computed from its spectrum on a 201×201 equiangular spherical grid with respect to angles φ and θ (see Eq. (1), and the reference frame on Fig. 5). The bright side of the imaging sphere corresponds to azimuth angles $-90^\circ \leqslant \varphi \leqslant 90^\circ$, while the dark side corresponds to azimuths $90^\circ \leqslant \varphi \leqslant 270^\circ$. It is seen that at low frequencies, such as at 1 kHz the HRTF for the sphere and for the Fritz head are qualitatively similar. It is also seen that at these frequencies "the bright spot" on the dark side of the imaging sphere is well defined for both the head shapes. At higher frequencies, such as 10 kHz the HRTFs for the Fritz model

and sphere are qualitatively very different. On the ipsilateral side (the side of the ear location) the maximum amplitude of the HRTF for the sphere is at the ear location. At the same time for the Fritz model a notch is observed at this location, where the HRTF amplitude can be 20 dB or so less than its maximum value. The pictures for the dark side are also different. If for the sphere we can see the bright spot surrounded by a circular interference pattern, the Fritz model demonstrates rather irregular filamentary structure of the notches, where the amplitude of the HRTF can be tens of decibels lower than the average (in fact, the bright spot and circular interference pattern are present also in the Fritz model, but it substantially masked by the notch structure).

The frequency dependence of the plane wave HRTF along with its dependence on the spherical angles is presented in Figure 11. Again we can see the qualitative difference between the HRTF for the sphere and for the Fritz and Kemar head models. The dynamic range of the HRTF for Fritz and Kemar is much larger than that for the sphere. The amplitude at the notches can be tens of decibels lower than the average values. It is also seen that while the HRTF for the sphere is clearly symmetric for elevation angles $\theta = 90^\circ \pm \alpha$, the HRTF for the head models is different for the planes defined by elevation angles $\theta = 45^\circ$ and $\theta = 135^\circ$ which cut the head above and below the elevation angle at the plane of ear location, $\theta = 90^\circ$. The figure also shows a difference between the HRTFs for Fritz and Kemar head models and the effect of the pinnae shape, since the difference between the two Kemar head meshes is only due to the different pinnae.

5. Comparison with the computations of other authors

Fortunately, there are computational results for the Fritz head model computed by other authors. Greff and Katz [14] independently computed the HRTF for frequency range from 100 Hz to 15 kHz using conventional BEM on substantially non-uniform meshes containing 21,693 and 12,054 panels, respectively. The results for the "Gain" in dB for two incident angles were scanned from the cited paper and renormalized to obtain true HRTFs. As the HRTF is plotted using a logarithmic scale (dBs), any renormalization results in a shift of graphs along the y-axis. As soon as at low frequencies we have $\left|H^{(pw)}\right| \to 1$, or 0 dB, we shifted the cited results to 0 dB level at the lowest frequency (100 Hz).

Figure 12 shows comparison between the present computations and that of Greff and Katz. It is seen that all the computations agree qualitatively. Quantitatively a good agreement is observed for

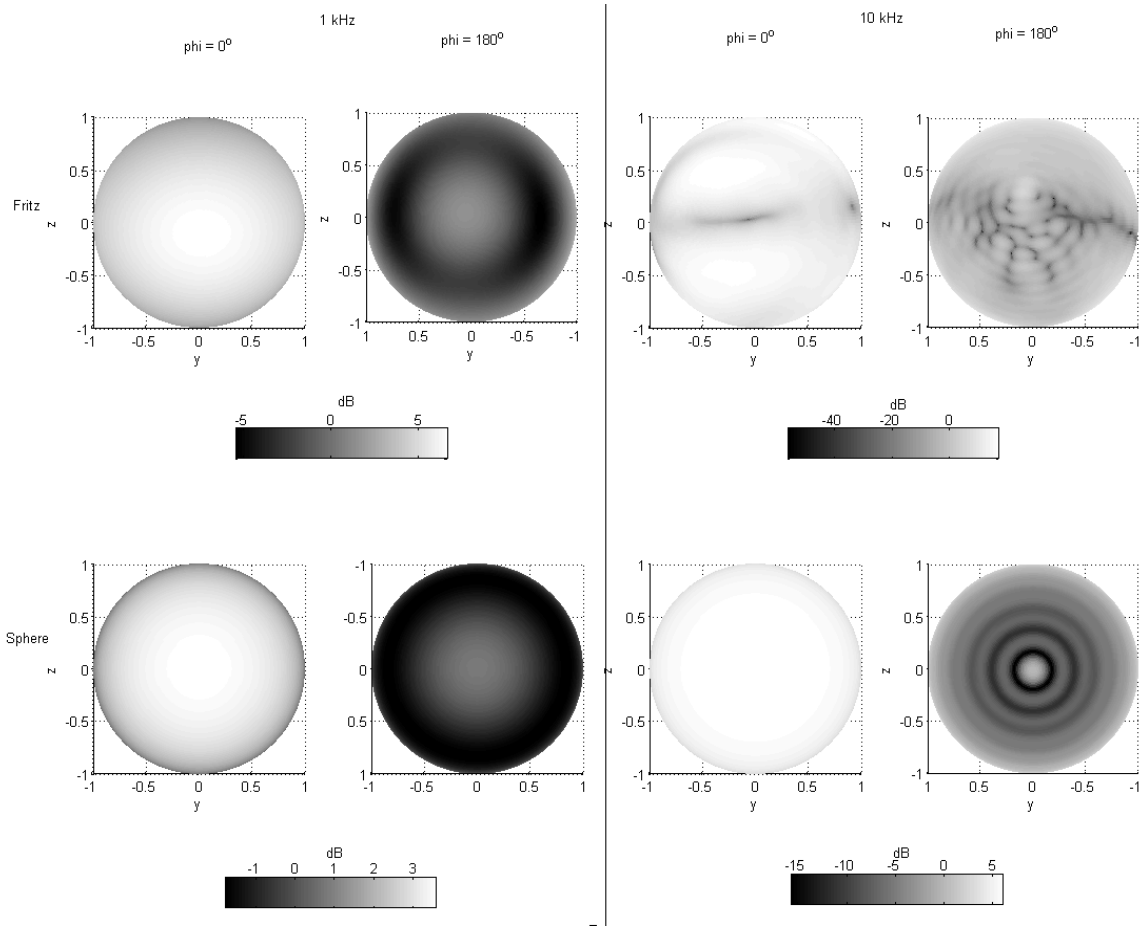

FIG. 10: The plane wave HRTF ($20\log_{10}\left|H^{(pw)}\right|$) computed for the Fritz model (the top row) and analytical for a sphere of radius $a = a_{\min} = 0.0705$ m (the bottom row) at frequencies 1.034 and 9.991 kHz. The HRTF is plotted as a function on a unit sphere of spherical angles θ and φ (see Eq. (1)). The center of the right ear is located at the $\theta = 90°$ and $\varphi = 0°$. The views corresponing to $\varphi = 0°$ are from the side of the right ear, while $\varphi = 180°$ are from the side of the left ear. Axes (x, y, z) are oriented as shown in Fig. 5.

the range of frequencies below 5 kHz, while for larger frequencies some mismatch takes place. There can be a few reasons for that. In our opinion, the most probable reasons are the difference in the location of the measurement point for the HRTF used by different authors, and the substantially coarser mesh used by those authors. The center of the panel, which we selected for the placement of the point source, in fact, had spherical angles $\theta \approx 89°$ and $\varphi \approx 0°$ instead of the ideal $\theta = 90°$ and $\varphi = 0°$. While a 1 degree shift in the direction of the plane wave should have a small effect, the reference panel may have a different location inside the pinnae cavity and different normal compared

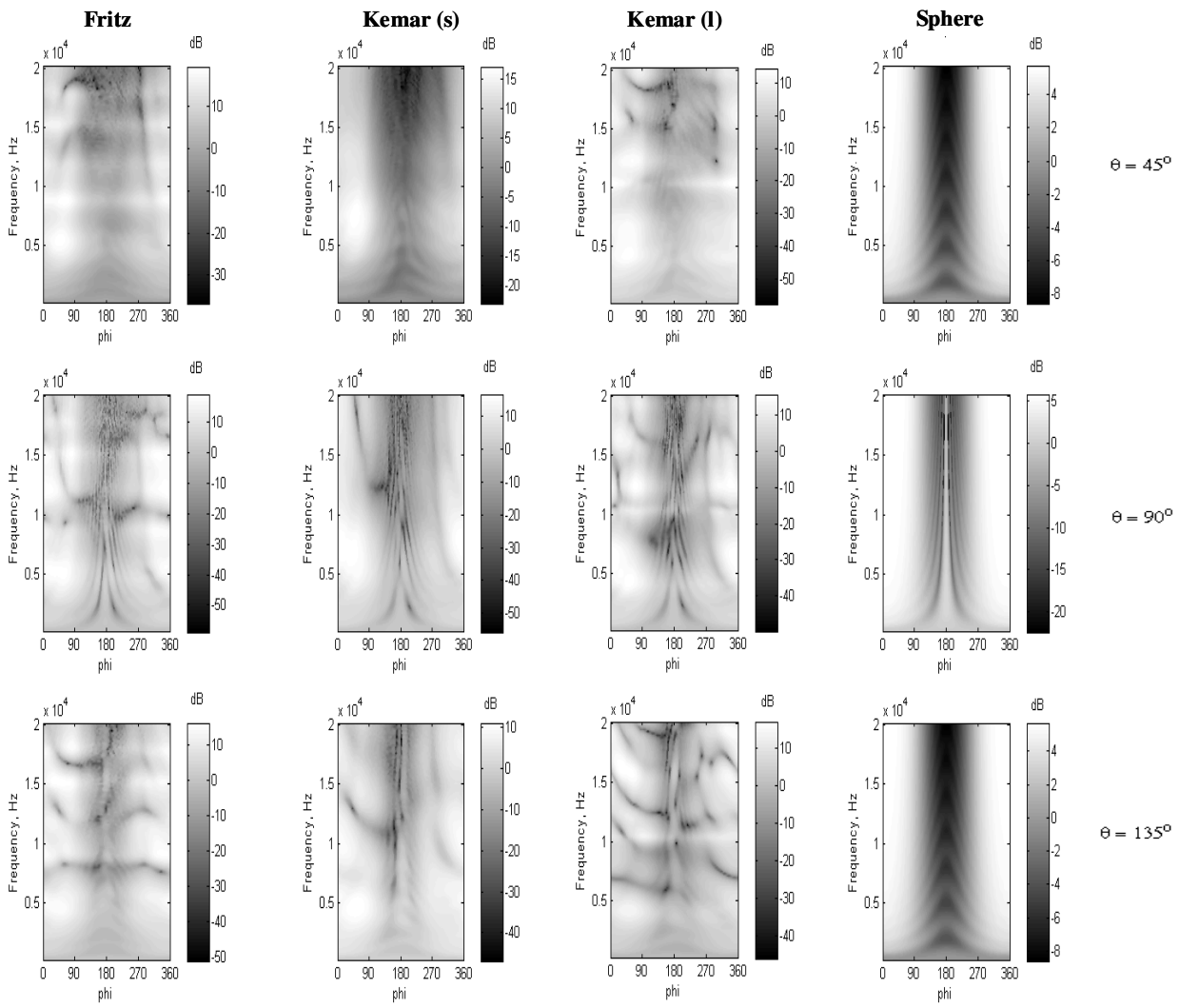

FIG. 11: The plane wave HRTF ($20\log_{10}\left|H^{(pw)}\right|$) for the righ ear ($\theta = 90°$, $\varphi = 0°$) computed for Fritz (the left column) and Kemar heads (the lef center column for the "small pinnae" model and the right center column for the "large pinnae" model), and analytical for a sphere of radius $a = a_{\min} = 0.0705$ m (the right column) at fixed spherical angles θ for the range of audible frequencies and spherical angles φ (see Eq. (1) and Fig. 5 for definitions).

to that used in computations of the referred authors. While this should have a negligible effect at lower frequencies, the location of the notch and higher frequency scattering can be somehow affected. In terms of comparisons with experiments and actual measurements, agreement between the computed data can be considered as good, since there are much larger differences between the

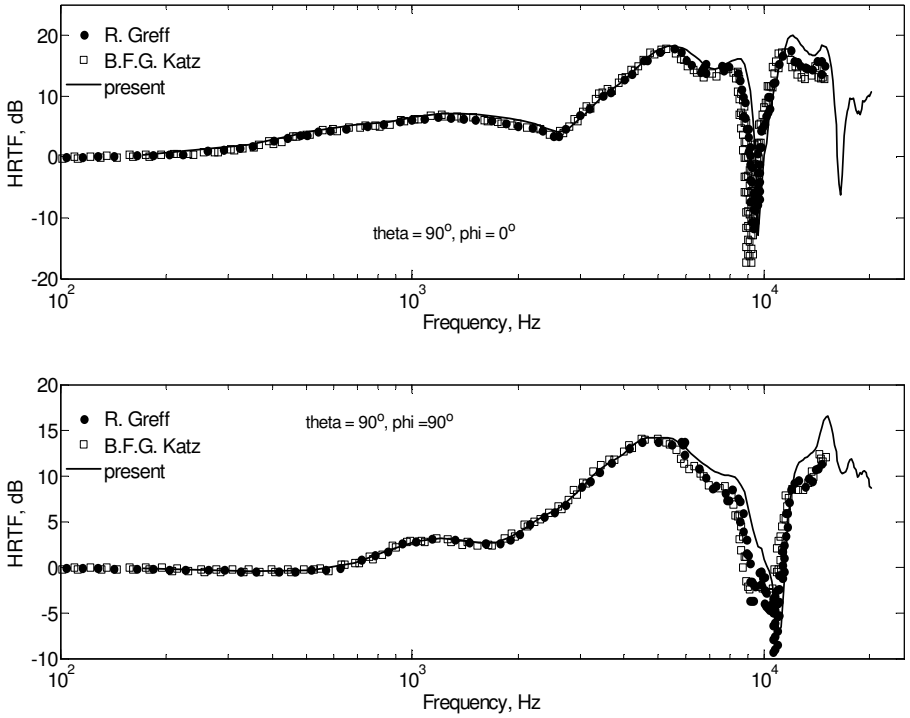

FIG. 12: Comparison of the plane wave HRTFs ($20\log_{10}\left|H^{(pw)}\right|$) for two incident angles computed independently for the Fritz model by Greff and Katz [14] and using the present method. Original normalized data of Greff and Katz are renormalized to match the HRTF asymptotics at zero frequency ($\left|H^{(pw)}\right| \to 1$).

experiments and computations. Modeling of a finite size transducer placed inside the ear by a single measurement point is always an issue (e.g. see the discussion in [14]). While there can be some other reasons for the quantitative discrepancy, e.g. related to the size of the mesh panels, accuracy of the method used, etc., the comparison convince that the HRTF in all cases is computed well enough.

6. Comparison with experiments

Two sets of experimental data for the Fritz model were available to the authors via their own measurements, using both a direct and a reciprocal set-up. These measurements were conducted as a part of the "Club Fritz" activity [24]. Unfortunately, measurements from other authors in the round-robin studies were not available to us. Comparisons with both direct and reciprocal

experiments provided approximately the same quality of agreement, so only data on the direct measurements which were conducted with a denser sampling are reported below for the sake of space.

In these experiments, the impulse response function was measured at both ears and then its spectrum was analyzed to determine the HRTF for 823 directions. The signal generation/processing technique used windowing, so the experimental data for frequencies above 14 kHz are not reliable and frequencies above 16 kHz are completely suppressed. Also, as is the case with most HRTF measurements (see e.g., [3] for a discussion) the experimental spectra were not reliable for low frequencies below 700 Hz. So for comparisons with computations only experimental HRTF for the range of frequencies from 700 Hz to 14 kHz were used.

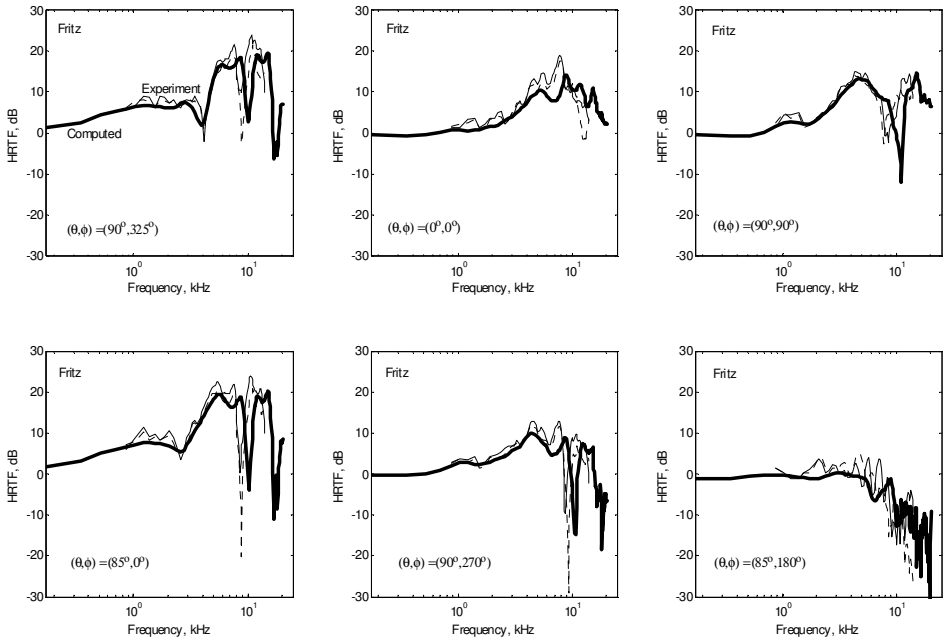

FIG. 13: Comparison of the computed and experimental HRTF for the Fritz model for several angles. Computed data for the right ear are shown by the thick lines. The thin solid lines show experiemental data for the right ear, while the dotted lines show data for the right ear derived from the experimental data for the left ear assuming symmetry of the model.

We compared experiments and computations for all 823 measurement locations, and several typical comparisons are shown in Fig. 13, where all non-normalized experimental data were shifted by 12 dB to provide consistent magnitudes with computations. A qualitative match of experiments

and computations is clearly seen, while a quantitative discrepancy is also observed. Such a discrepancy may arise from several factors, including modeling issues (the microphone in computations is modelled as a point on the pinnae surface) and experimental errors. The latter, for example include some imprecision in the recording of the tilt in the model position, which may result in a few degrees of inconsistency in the directions for which the HRTF was measured and computed. To evaluate such a type of the error, we plotted in Fig. 13 by the dashed lines measurements for the left ear, for which directions were modified to get the "right ear data from the left ear measurements". This was easy to do, since the 823 directions were symmetric with respect to the azimuth $\varphi = 90^\circ$ and the same set could be used for evaluation of the response of both ears (pairs $\varphi = 90^\circ \pm \alpha$ with the same θ were swapped). In the case of perfect symmetry and ideal measurements the actual data for the right ear and derived from the data for the left ear should coincide for this symmetric model. However, Fig. 13 shows discrepancy of these data, which can be used to judge the experimental errors and model misalignment. It is seen then that in some cases inconsistency of the experimental data is of the same order of magnitude as the inconsistency between the experiment and computations, while in some other cases the experimental error is smaller that between the experiments and computations (e.g. 1-2 kHz shift in the HRTF notches), which we refer to the idealizations of the computational model. Fig. 14 shows the data measured and computed at a constant $\theta = 90^\circ$ elevation (coronal plane through the ears) for the frequency range of validity of the experimental data. Experimental data for $\theta = 90^\circ$ measured at 58 azimuth angles φ were used for computations. This figure shows that an overall qualitative match exist between the computations and experiments, while quantitative discrepancy (in terms of notch locations) is clear for frequencies above 7-8 kHz or so.

C. Head and torso model computations

While computations for the head alone provide important information on the HRTF, the effect of torso on the HRTF is substantial. Modeling of this effect using a simple geometric approximations was performed in [3, 31]. In the present study we used a hybrid model consisting of detailed head shape of Kemar and torso modelled as an ellipsoid disjoint from the head. Parameters of the ellipsoid fitting the actual torso can be found in [3].

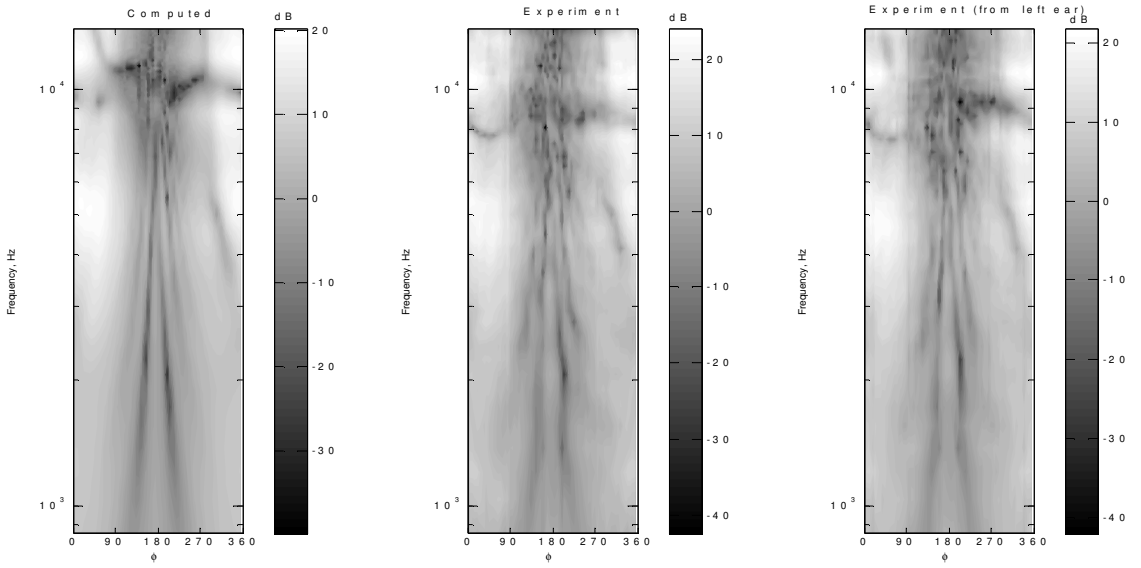

FIG. 14: Computed and measured HRTF (in dB) for the Fritz model (right ear) at angle $\theta = 90°$ for frequency range 800 Hz-14 kHz. The middle plot is obtined from the right ear measurements, while the right plot is obtained from the left ear measurements applying symmetry.

1. Mesh

The finest "small pinnae" mesh used for computations had 213,427 vertices and 426,846 panels (222642/445276 for the finest "large pinnae mesh"). Both head and torso meshes satisfy the quality requirements for the BEM/FMM. Figure 15 illustrates the model and meshes used.

These meshes are suitable for computations of the HRTF in the entire audible range and provide not less than 6 elements per the shortest wavelength computed (frequency 20.155 kHz). The dimensionless size of the computational domain for this frequency was $kD = 410$, which is in the range of the BEM/FMM capabilities for personal computers. To reduce the computational cost we used two more meshes ("mid frequency" and "low frequency" meshes), which had 90,547/181,086 and 59,827/119,646 vertices/panels, respectively (for the "small pinnae"). These meshes had the same head meshing as the high frequency mesh, while we used coarser discretizations for the torso. The meshes were applied for frequency ranges 7.2-14.2 kHz and below 7.2 kHz, respectively. In these ranges the head mesh certainly satisfied 6 elements/per wavelength criteria, while the torso discretizations were a bit rougher, but anyway not less than 5.5 elements/per wavelength.

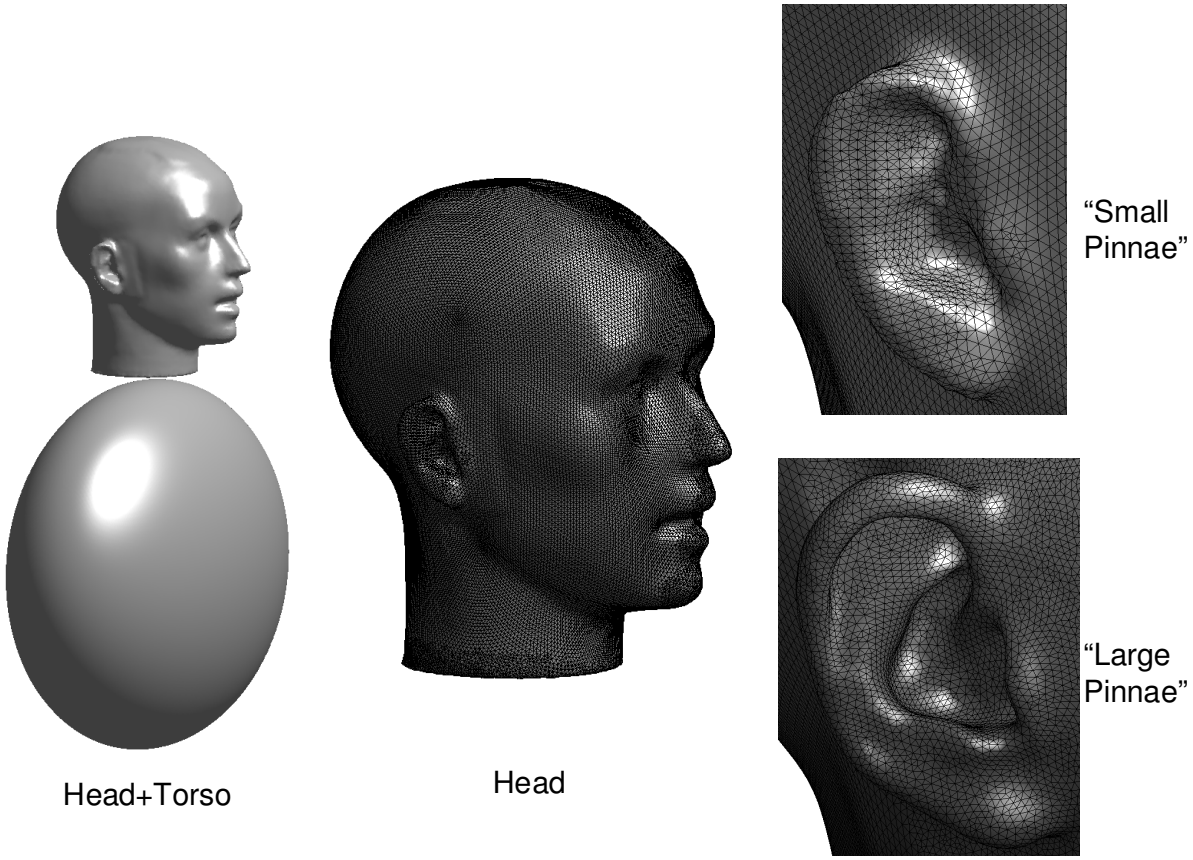

FIG. 15: The mesh used for computation of the HRTF using the FMM accelerated BEM for the Kemar model with ellipsoidal torso. The mesh contains 213427 vertices and 426846 triangular faces ("Small Pinnae") or 222642 vertices and 445276 faces ("Large Pinnae"). The four plots show zooming in the head and the right pinnae regions.

2. Performance

In terms of computations adding the torso substantially increases the computation time, as the size of the computational domain increases. As the performance for the models with "large" and "small" pinnae is approximately the same, only the latter performance data are reported below. The total time for the BEM solution was slightly below 70 hours and required about 8 GB RAM at the peak for the highest frequency. Figure 16 shows the dependence of the BEM solution time on the frequency and the number of iterations in the outer fGMRES loop. If the later quantity is constant and the mesh is the same for all frequencies the time should be well approximated with parabola $y = ax^2 + b$, where the first term is due to frequency dependent translations in the

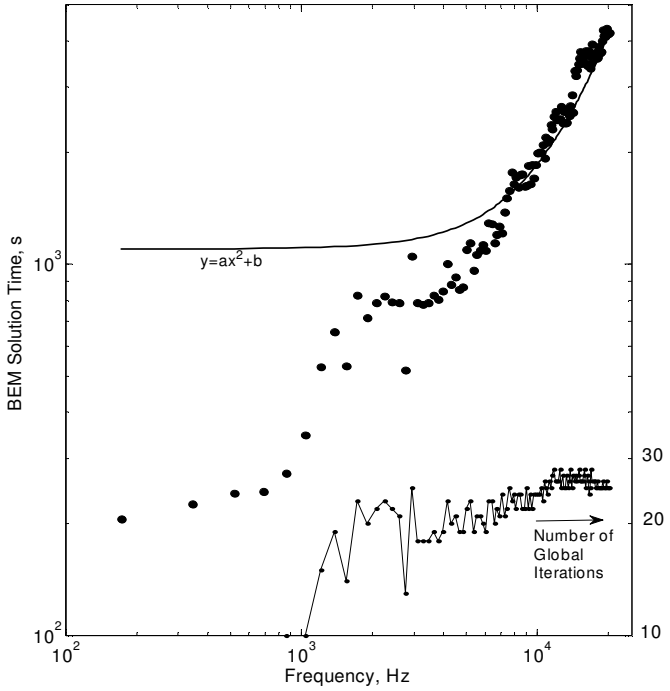

FIG. 16: Wall clock BEM solution time for Kemar+Torso model (the filled disks) and the number of iterations in the outer fGMRES loop. The dashed curve shows fit with function $y = ax^2 + b$ (log-log plot), which is a theoretical approximation of the time assuming constant number of iterations.

FMM, and the second term due to the constant sparse matrix multiplication. The actual time shows deviations because of the changes in the number of fGMRES iterations. The number of iterations grows very slowly with the frequency; for frequencies below 1 kHz the number of global iterations did not exceed 10 while it reached a maximum of 28 for the high frequencies. For the highest frequency the time per matrix-vector multiplication in the outer loop was 57 s, while for the inner preconditioning loop that was 9.8 s (for the lowest frequency these times are 14.6 and 0.6 s, respectively). The times shown in the figure reflect the most expensive part of the numerical solution, and do not include computation of the spectrum and evaluation of the HRTF.

3. Effect of torso

As a result of larger model size, plus the fact that the spectral representations can be applied to the region outside a large sphere concentric with the head, which includes the full torso, the number of harmonics for the head+torso model was substantially larger than that for the head alone. Figure 17 illustrates spectral data for three frequencies computed for the head with torso and head alone ("small pinnae"). In this figure the data with amplitude $\left|H_n^{(pw)m}\right| < 10^{-4}$ are truncated, while the actual truncation using Eq. (59) provides magnitudes down to $\left|H_n^{(pw)m}\right| \sim 10^{-10}$. We found that for all frequencies several low frequency spatial modes of models with and without the torso match almost perfectly. The difference arises at high frequency spatial modes starting from some value dependent on the frequency of the acoustic field. The higher the acoustic frequency the higher is this starting value. So, e.g. at 500 Hz spectral differences are observable for modes with $n \geqslant 3$, while for 5 kHz the difference is visible for $n \gtrsim 10$ and for 15 kHz for $n \gtrsim 25$. The high frequency spatial modes in the presence of torso have a considerable magnitude in a wider range than that for the head alone.

Figure 18 shows the plane wave HRTF corresponding to the spectra from Fig.17. It is seen that the presence of torso results in interference patterns due to reflections from the torso, which spatial frequency is higher at higher acoustic frequencies. Since the speed of sound is constant, the wavenumber and acoustic frequency simply are proportional. According to the spectral analysis at high acoustic frequencies the torso does not affect low frequency spatial modes, while high frequency spatial component shows up as a wave reflected from the torso, which is superposed with the HRTF for the head alone. At low frequencies when the wavelength is comparable with the head size the torso affects low frequency spatial components, which create effects of totally different spatial dependence of the HRTF.

Figure 19 represents the HRTF along the centerlines when the azimuth varies at constant elevation ($\theta = 90°$) and the elevation varies at constant azimuth ($\varphi = 0°$, $\delta = 90° - \theta$ and $\varphi = 180°$, $\delta = 90° + \theta$; angle δ is the same as in [3]). The effect of torso is seen on both pictures, while for the latter case it is especially clear as shows a characteristic pattern for angles $180° < \delta < 270°$. This pattern is similar to that measured for the snowman model and the Kemar model without ears reported in [3] at frequencies below 5 kHz. The present computations show that, in fact, the effect of torso as strong at high frequencies as at low frequencies, since reflections from the torso

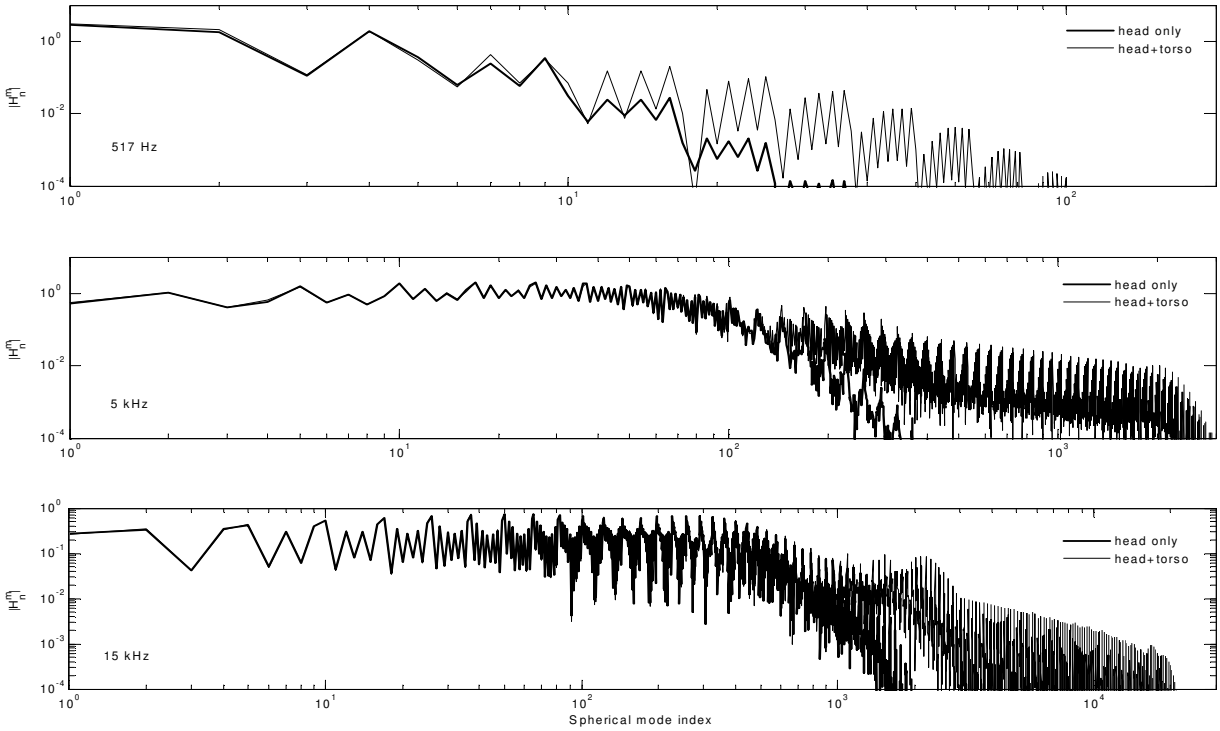

FIG. 17: The power spectra of spherical representation of the plane wave HRTF, $\left|H_n^{(pw)m}\right|$ as functions of a single spherical mode index $i = (n+1)^2 + m$, $n = 0, ..., p-1$, $m = -n, ..., n$ for different frequencies (Kemar head alone and Kemar with torso, "small pinnae").

occur for waves of all frequencies. On the other hand, it is also clear that the presence of the torso affects a little the notch structure, which is due to pinnae shape.

D. Comparisons with experimental data

Several sets of experimental data from different researchers were compared with the computed data for Kemar. The basic conclusion we came to is that computations for the "small pinnae" model do not agree well with the experiments with the "small pinnae" at frequencies above 6 kHz or so, while computations for the "large pinnae" agree good enough with the "large pinnae" experiments up to 20 kHz. The major reason for this we see in the ways how the ear meshes were produced. In the former case that were done via the laser scans, which lose information on the

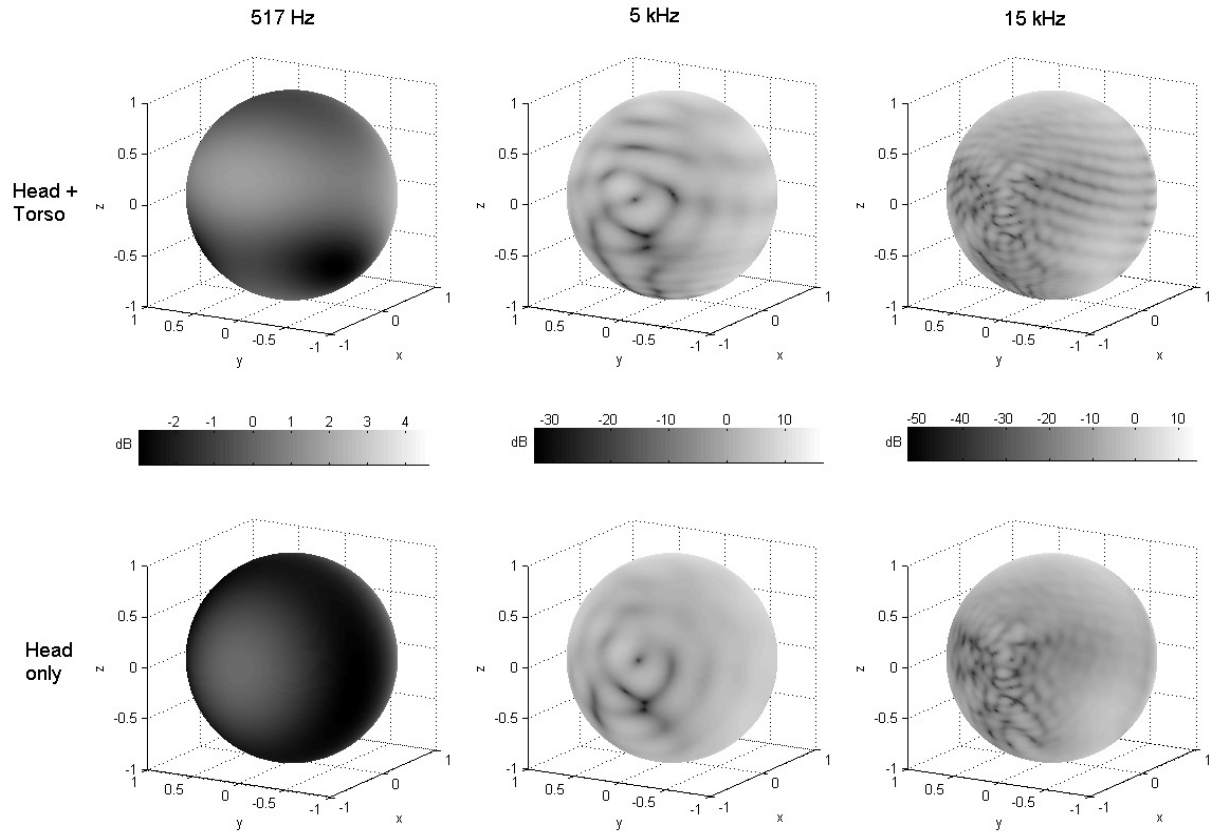

FIG. 18: The plane wave right ear HRTF ($20\log_{10}\left|H^{(pw)}\right|$) computed for the "small pinnae" Kemar with torso (the top row) and without torso (the bottom row) at different frequencies. The HRTF is plotted as a function on a unit sphere of spherical angles θ and φ (see Eq. (1)). The center of the right ear is located at the $\theta = 90^\circ$ and $\varphi = 0^\circ$ ($x = 1$, $y = 0$, $z = 0$). The views are from the side of the left ear. Axes (x, y, z) are oriented as shown in Fig. 5.

depth and so the mesh has low fidelity to the real shape. In the latter case the CT scans were used for mesh generation, which provide much better information on the features of the real shape. The importance for the use of quality scans for meshes and effect on the computation was also discussed in [28].

The comparisons presented below use experimental data from the CIPIC database [2]. All the data there are standardized in terms of measurement technique and data are presented in the hoop coordinates φ_{hoop} (azimuth) and θ_{hoop} (elevation). These angles are different from the spherical polar angles we use, since in the experiments the poles of spherical coordinates were located at

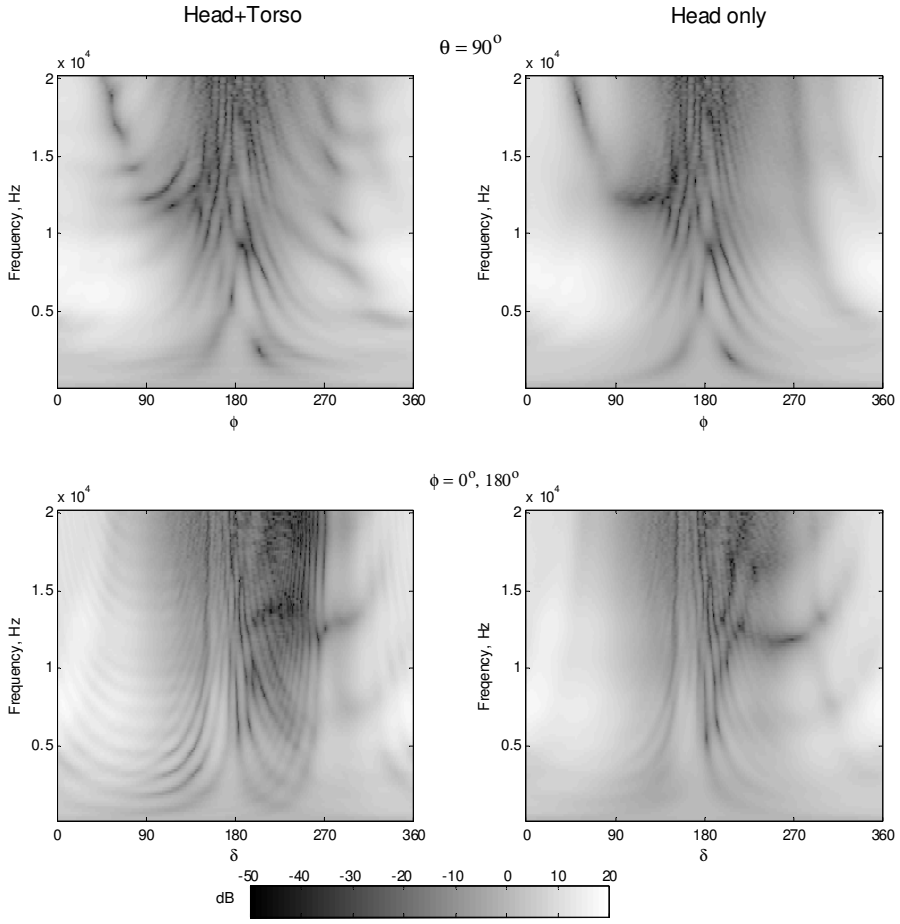

FIG. 19: The plane wave right ear HRTF ($20\log_{10}\left|H^{(pw)}\right|$) computed for the "small pinnae" Kemar with torso (the left column) and without torso (the right row) for fixed elevation ($\theta = 90°$, the top row) and for the centerline going through the rears and the top of the head (the bottom row). In the latter case angle δ is measured from the right ear; $\delta = (90° - \theta) \bmod 360°$ for $\varphi = 0°$ and $\delta = 90° + \theta$ for $\varphi = 180°$ ($0 \leqslant \theta \leqslant 180°$).

the ear positions. Relation between the angles (φ, θ) and $(\varphi_{hoop}, \theta_{hoop})$ follows from Eq. (1) and expression of the Cartesian coordinates via the hoop coordinates

$$\mathbf{r} = (x, y, z) = r\left(\sin\varphi_{hoop}, \cos\varphi_{hoop}\cos\theta_{hoop}, \cos\varphi_{hoop}\sin\theta_{hoop}\right), \tag{77}$$
$$-\pi/2 < \theta_{hoop} \leqslant 3\pi/2, \quad -\pi/2 \leqslant \varphi_{hoop} \leqslant \pi/2, \quad r = |\mathbf{r}|,$$

so the right ear location, for which all comparisons are made, is corresponds to the pole $\varphi_{hoop} = \pi/2$.

Figure 20 shows comparison of the computation (plane wave HRTF) and experiments for the frequency range 0.5-20 kHz for frontal direction ($\varphi_{hoop} = \theta_{hoop} = 0$). It is seen that for frequencies above 6 kHz or so there is a substantial qualitative deviation of the computations from experi-

ments for the "small pinnae" case. On the other hand, qualitatively computations for the "large pinnae" match with the experiments as all maxima and minima of the frequency dependence are reproduced. Moreover, a good quantitative match is observed for frequencies up to 10 kHz and for the frequency and the depth of the strongest notch. It is also interesting that computations for head alone quantitatively are in better agreement with experiments for frequencies above 10 kHz, while computations with torso agree a bit better qualitatively (the experimental data on this figure were normalized for 3 dB shift to match maximum value of the HRTF for the "large pinnae" case). This shows that to reproduce effect of torso quantitatively a better model of the torso than simple fitting ellipsoid should be used.

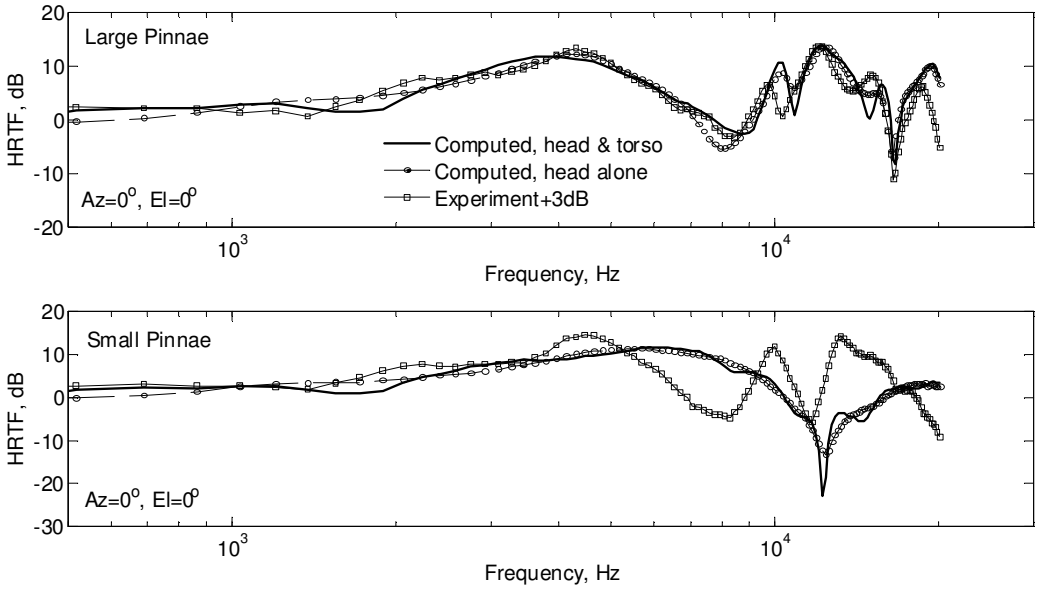

FIG. 20: Comparison of the computed HRTF for head and torso, head alone and CIPIC experiments (right ear) for the frontal direction (azimuth=0°, elevation=0°). The top graph corrsponds to the "large pinnae" model, while the bottom graph shows comparison for the "small pinnae" model. The experimental data are shifted up to 3 dB.

Figure 21 supports these observations for the entire range of azimuths at fixed elevation and for all elevations at fixed azimuths. Its is clear that computations for the "large pinnae" match experiments in terms of the notch structure (dark filamentary patterns), while experiments for the "small pinnae" show more complex notch structure than the computed one. The effect of torso is well seen on the charts plotted for the fixed azimuth. Both experiments and head and torso computations show the ripples due to torso reflections, which are not present in the computations

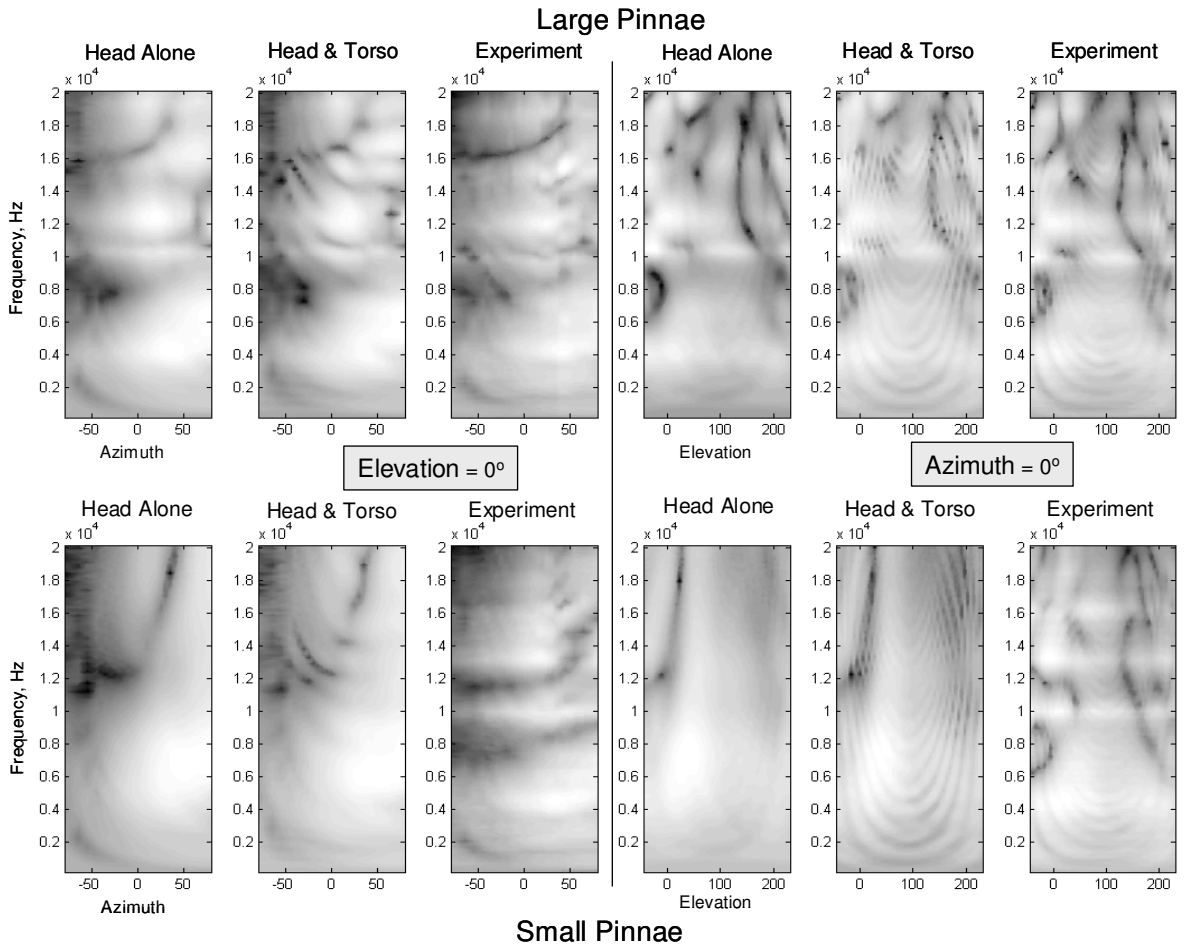

FIG. 21: Comparison of computed (head alone and head and torso) and experimental HRTF data (in dB) for the entire range of audible frequencies for fixed elevation (0°) and all measured azimuths (-80° $\leqslant \varphi_{hoop} \leqslant$ 80°) (the six pictures on the left) and for fixed azimuth (0°) and all measured elevations ($-45° \leqslant \theta_{hoop} \leqslant 230.625°$) (the six pictures on the right). The top row shows data for the "large pinnae" case and the bottom row shows data for the "small pinnae" case. Experimental data are taken for the right ear. The dynamic range is varying from about -60 dB (the darkest spots) to about +10 dB (the brightest areas).

with head alone. The computed and experimental ripple patterns and ripple frequencies agree well, while as we noted above the quantitative agreement for a given point can be not very good, which happens due to some ripple phase mismatch and slight overall distortion of the ripple pattern due to approximate torso shape modeling.

Figure 22 shows comparisons of computed (head and torso) and experimental data for three fixed

frequencies (at about 5, 14, and 19 kHz) and all measured azimuths and elevations, which gives an idea on spatial dependence of the HRTF for a given frequency. Again we can see consistency of the computed and experimental data for the "large pinnae case", while such consistency is observed for the "small pinnae" case only at frequency 5 kHz. At larger frequencies the latter case shows a different notch structure. We did not plot on this graph the computed data for the head alone, which are similar to computations taking into account the presence of the torso in terms of notch locations, but do not have reflection ripples. Such ripples are clearly seen in experiments and the ellipsoid torso approximation reproduces this structure good enough.

VI. CONCLUSIONS

The main purpose of the present study was development and testing of an efficient method for computation of the HRTF over the audible frequency range. We believe that this objective was achieved. Indeed, comparing our results with those of the direct boundary element methods, the present FMM accelerated BEM shows a superior performance and is suitable for computation of the HRTF in the full audible range including not only the head, but the head and torso models. The method appeared to be stable and produces good results for meshes containing hundreds of thousands boundary elements, which provides accurate computation and representation of the HRTF at frequencies as high as 20 kHz.

Representation of the HRTF via its discrete spherical spectrum, suggested in this paper (and for fitting of experimental data in [12]), is rather natural. First, this is a relatively compact representation, and, second it provides spectral accuracy consistent with the accuracy of computations, and third, fast evaluation of the HRTF for any set of directions. In particular, this is important for rotation transforms, range dependence, and corrections of the HRTF due to body motion in the use of the HRTF in virtual reality.

Comparisons with experiments for the head alone and the head and torso, showed that a good agreement can be obtained in the entire audible range. However, to achieve this goal several things should be handled with care. This relates to accurate scans of pinnae (laser scans may provide loss of important features, while the CT scans are much more accurate though difficult to acquire), careful reproduction of experimental environment (e.g. positioning of the head model, modeling of torso (if present), etc.), and quality computational meshes.

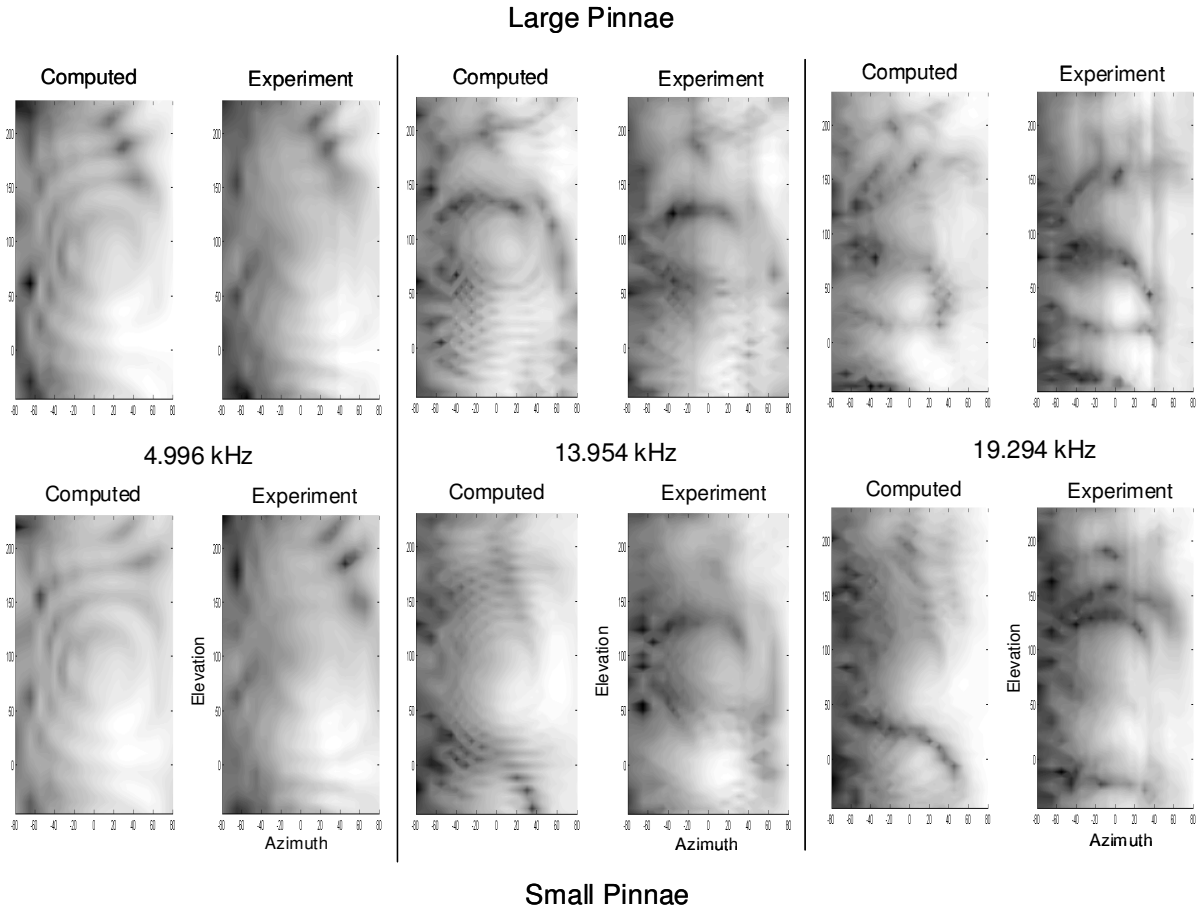

FIG. 22: Comparisons of the computed (head and torso) and CIPIC experimental HRTF data (in dB) for three fixed frequencies (5.2, 14.1, and 19.2 kHz) and all azimuths and elevations measured in the experiments. The upper row showss the "large pinnae" case, while the bottom row shows for the "small pinnae" case.The dynamic range for the higest frequency case is varying from about -60 dB (the darkest spots) to about +10 dB (the brightest areas) .

Of course, the availability of a computational tool for the HRTF is only a first step. These should now be used to explore the dependence of HRTFs on shape, to compute HRTFs for different animals/individuals, and to study the influence of things like clothing and hair. To aid in this effort it would be essential to combine the different research tools such as those developed in this study (the meshing procedures, the fast multipole software, and the software for analysis of the spherical spectra) in to a usable tool. Then this tool should be used for computation of the HRTF for other meshes, modeling real human head/torso interactions, for comparisons with data in available HRTF

databases, and conducting experiments on perceptual experiences with the computed HRTFs. We were able to compute full range HRTF for tens of hours on a PC. However, for parametric studies, further accelerations are needed. These may be provided by supercomputers, but we believe that the method can be further accelerated in the personal computing environments being made available, e.g. using graphic processors. Indeed, some promising results are obtained for the FMM for the Laplace equation on graphical processors [18], while the method should be developed further for the Helmholtz equation. Also we note that research is needed on the use of HRTF spherical spectra in applications.

VII. ACKNOWLEDGMENTS

These computations are a result of a long term project that has been performed over the past eight years at the Perceptual Interfaces and Reality Lab., at the University of Maryland. At various stages our research has been supported by the National Science Foundation, DARPA, the state of Maryland, and Fantalgo, LLC. All these agencies are thanked for their support. We thank Profs. R.O. Duda and V.R. Algazi of the University of California, Davis for several discussions on the HRTF, for providing the CIPIC database, and for information on it. They are also to be thanked for providing the CT-scans of the Kemar large pinna. We would also like to thank Dr. Kexue Liu for creating the pinna mesh, Dr. Elena Grassi for the measurements of the Fritz HRTFs, Dr. Yuvi Kahana for the original Kemar mesh, and Dr. Brian Katz for the original Fritz mesh. We would like to acknowledge the use of software provided by Fantalgo, LLC for the fast multipole boundary element method, and for the fast multipole interpolation used in the remeshing procedure.

---

[1] M. Abramowitz and I.A. Stegun. Handbook of Mathematical Functions. National Bureau of Standards, Washington, D.C., 1964.

[2] V.R. Algazi, R.O. Duda, D.M. Thompson and C. Avendano. "The CIPIC HRTF Database," Proc. 2001 IEEE WASPAA, 21-24, 2001.

[3] V.R. Algazi, R.O. Duda, R. Duraiswami, N.A. Gumerov, and Z. Tang. "Approximating the head-related transfer function using simple geometric models of the head and torso," J. Acoust. Soc. Am., 112:2053-2064, 2002.

[4] J.L. Aroyan. "Three-Dimensional Modeling of Hearing In Delphinus Delphis," J. Acoust. Soc. Am. 110:3305-3318, 2001.

[5] D.W. Batteau. "The role of the pinna in human localization", Proc. Royal Soc. London, 168(ser. B):158-180, 1967.

[6] C.F. Bohren and D.R. Huffman. Absorption and scattering of light by small particles. Wiley, New York. 1983.

[7] A.J. Burton and G.F. Miller. "The application of the integral equation methods to the numerical solution of some exterior boundary-value problems," Proc. Roy. Soc. London, Series A, Math. Phys. Sci., 323:201–210, 1971.

[8] W.C. Chew, J.M. Jin, E. Michelssen, and J. Song (eds.). Fast and Efficient Algorithms in Computational Electromagnetics (Artech House, Boston), 2001.

[9] F. De Mey, J. Reijniers, and H. Peremans. "Simulated head related transfer function of the phyllostomid bat Phyllostomus Discolor," J. Acoust. Soc. Am. 124:2123-2132, 2008.

[10] R.O. Duda and W.L. Martens. "Range Dependence of the Response of A Spherical Head Model," J. Acoust. Soc. Am. 104:3048-3059, 1998.

[11] R.O. Duda. Personal communication, 2002.

[12] 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.

[13] W. Gardner and K. Martin. "HRTF Measurements of a KEMAR," J. Acoust. Soc. Am., 97:3907-3908, 1995.

[14] R. Greff and B.F.G. Katz. "Round Robin Comparison of HRTF Simulation Results: Preliminary Results." Audio Engineering Society Convention Paper, Presented at the 123rd Convention, 2007.

[15] N.A. Gumerov and R. Duraiswami. Fast Multipole Methods for the Helmholtz Equation In Three Dimensions. Elsevier, 2005.

[16] N.A. Gumerov and R. Duraiswami. "A scalar potential formulation and translation theory for the time-harmonic Maxwell equations," J. Comput. Physics, 225:206-236, 2007.

[17] N.A. Gumerov and R. Duraiswami. "Fast radial basis function interpolation via preconditioned Krylov iteration," SIAM J. Sci. Comput., 29:1876-1899, 2007.

[18] N.A. Gumerov and R. Duraiswami. "Fast multipole methods on graphics processors," Journal of Computational Physics, 227:8290-8313, 2008.

[19] N.A. Gumerov and R. Duraiswami. "A Broadband Fast Multipole Accelerated Boundary Element Method For the 3D Helmholtz Equation," J.Acoust. Soc. Am., 125:191-205, 2009.

[20] S. Hanish. A Treatise on Acoustic Radiation (Naval Research Laboratory, Washington, D.C., 1981.

[21] W.M. Hartmann. "How we localize sound", Physics Today, Nov. 1999, 24-29.

[22] Y. Kahana, P.A. Nelson, M. Petyt, S. Choi. "Numerical Modelling of the Transfer Functions of a

Dummy-Head and of the External Ear," AES 16th International Conference on Spatial Sound Reproduction, 1:16, 1999.

[23] B.F.G. Katz. "Boundary Element Method Calculation of Individual Head-Related Transfer Function. I. Rigid Model Calculation," J. Acoust. Soc. Am. 110:2440-2448, 2001.

[24] B.F.G. Katz and D.R. Begault. "Round Robin Comparison of HRTF Measurement Systems: Preliminary Results," 19th International Congress On Acoustics, 2007.

[25] K. Liu, R. Duraiswami, and L.S. Davis. "A Simple and Optimal Energy Surface Reconstruction Algorithm from Volumetric Data," University of Maryland Department of Computer Science Technical Report CS-TR-4441 (Also UMIACS-TR-2003-11), 2003.

[26] P.M. Morse and K.U. Ingard. Theoretical Acoustics, Princeton Univ. Press, New Jersey, 1968.

[27] A.E. O'Donovan, N.A. Gumerov, R. Duraiswami. "Good Meshes For the Fast Multipole Accelerated Boundary Element Method Via Radial Basis Function Interpolation and Relaxation," University of Maryland, Department of Computer Scince Technical Report, CS-TR-4937 and UMIACS-TR-2009-07, 2009.

[28] M. Otani and S. Ise. "Fast Calculation System Specialized For Head-Related Transfer Function Based on Boundary Element Method," J. Acoust. Soc. Am. 119:2589-2598, 2006.

[29] V. Rokhlin. "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," Appl. and Comp. Harmonic Analysis, 1, 82-93, 1993.

[30] Y. Saad. "A Flexible Inner-Outer Preconditioned GMRES Algorithm," SIAM J. Sci. Comput. 14:461-469, 1993.

[31] N. Gumerov, R. Duraiswami, and Z. Tang. "Numerical study of the influence of the torso on the HRTF," Proceedings ICASSP 2002, Orlando, Florida.

[32] T. Walsh, L. Demkowicz, and R. Charles. "Boundary Element Modeling of the External Human Auditory System." J. Acoust. Soc. Am. 115:1033-1043, 2004.

[33] T. Xiao and Q-H. Liu. "Finite Difference Computation of Head-Related Transfer Function for Human Hearing," J. Acoust. Soc. Am. 113:2434-2441, 2003.

[34] D.N. Zotkin, R. Duraiswami, E. Grassi, and N.A. Gumerov. "Fast Head-Related Transfer Function Measurement Via Reciprocity," J. Acoust. Soc. Am. 120:2202-2215, 2006.