

Optimal Series Truncation of the Rigid-Sphere Head-Related Transfer Function for Accurate Binaural Perception

RAHULRAM SRIDHAR,* *AES Student Member*, AND EDGAR Y. CHOUERI, *AES Associate Member*
(rahulram@princeton.edu) (choueiri@princeton.edu)

3D Audio and Applied Acoustics Laboratory, Princeton University, Princeton, New Jersey 08544, USA

A formula is derived for computing the order at which the infinite series representation of the rigid-sphere head-related transfer function (RS-HRTF) must be truncated to minimize the time required to compute the HRTF to a sufficiently high accuracy based on binaural perception metrics. Quick and accurate computation of this HRTF may be useful for implementing spatial audio in computationally limited and portable devices. Using a brute-force approach, the lowest truncation order, N_{\min} , that yields the RS-HRTF that differs from the benchmark (i.e., the RS-HRTF computed with the highest possible accuracy) by less than just-noticeable difference thresholds in interaural time and level differences is approximately computed for a wide range of source distances. By fitting power and rational functions to these computed values, a formula that approximates N_{\min} as a function of frequency and source distance is derived. It is shown that truncation order varies significantly with source distance and that the proposed formula, unlike a previous one, accurately captures this variation. Consequently, using the proposed formula instead of the previous one results in a more accurate RS-HRTF that is also computed 48% faster on average.

0 INTRODUCTION

The head-related transfer function (HRTF) is a mathematical or numerical description of the filtering of free-field acoustic waves exclusively by an individual's morphology (primarily the head, pinnae, and torso) before they reach the individual's eardrums. It is used, for example, to render three-dimensional audio (also known as spatial audio) in virtual reality applications and listening tests for psychoacoustic research.

There are numerous ways of acquiring an individual's HRTF (for comprehensive summaries, see the work byENZNER et al. [1] and XIE [2], for example). One approach [3–5] involves approximating the individual's head to be a rigid (i.e., sound-hard) sphere and computing the HRTF numerically by truncating its infinite series representation. Aside from being easy to compute, this approach is attractive because the *rigid-sphere HRTF* (RS-HRTF) is well-suited to individualization using anthropometric measurements

[6, 7]. It is a good approximation to human HRTFs for frequencies below approximately 2 kHz [4, 8, 9] and is used to correct the frequency response of acoustically measured HRTFs at these frequencies [8, 9], where the signal-to-noise ratio may be poor.

Another application of the RS-HRTF is to extend acoustically-measured far-field HRTFs to the near-field [10, 11] as a starting point for developing computationally efficient models of human HRTFs [12, 13] and to compute interaural level differences (ILDs) [14–16] and interaural time differences (ITDs) [17]. For example, in the so-called "structural HRTF model" proposed by BROWN and DUDA [12], ILD is incorporated using a "head-shadow" filter, which is a first-order, infinite-impulse response (IIR) filter approximation of the far-field RS-HRTF. Additionally, ITD is incorporated by combining the ITD computed from the phase response of this filter with that obtained from the Woodworth formula [18]. Recent work [13, 19] has explored strategies for improving the accuracy and efficiency with which the RS-HRTF may be computed, and it has been demonstrated that a structural HRTF model that uses an RS-HRTF is not only well-suited for real-time applications [12] but offers perceptual benefits over generic

*Author to whom correspondence should be addressed. Electronic mail: rahulram@princeton.edu

HRTFs when used for binaural rendering of higher-order ambisonics [20], for example.

From the preceding discussion we see that quick and accurate computation of the RS-HRTF may be beneficial for implementing spatial audio in power-limited portable devices such as electronic travel aids [21], computationally limited devices such as hearing aids [22], and applications where, for example, an individualized RS-HRTF needs to be quickly computed for a wide range of source directions and distances for real-time dynamic binaural synthesis [20, 23]. Although one approach for improving the speed at which the RS-HRTF is computed is to use filter approximations as described earlier, an alternative is to minimize the order, N , at which the infinite series representation is truncated such that the resulting RS-HRTF is still sufficiently accurate for such applications. Since the RS-HRTF is computed directly from the series representation, this latter approach, unlike the approximate filter models, for example, enables the RS-HRTF to be computed exactly for a given N .

0.1 Previous Work

Cooper and Bauck [3] describe a method for computing the RS-HRTF for sources infinitely far away. In their method, the infinite series representation is truncated when the fractional change in the HRTF for two successive terms first equals or falls below a user-specified threshold, T_c . That is, the truncation order, N , is iteratively increased until convergence is achieved. When $T_c = 0$, the HRTF is computed with the highest possible accuracy (i.e., limited only by finite-precision rounding errors), with larger values of T_c producing HRTFs that are less accurate but also reducing computation time [4]. Cooper and Bauck arbitrarily choose $T_c = 0.001$ in their published Fortran program and do not relate the accuracy of the resulting HRTF relative to those computed with the highest possible accuracy in terms of meaningful quantities such as errors in ITDs and ILDs. As we subsequently show, their program is not optimized to quickly compute the RS-HRTF to a sufficiently high accuracy due to the iterative nature of the computation and the arbitrary choice of T_c .

Using the infinite series representation provided by Rabinowitz et al. [24] for computing the RS-HRTF for sources at any finite distance from the head, Duda and Martens [5] present a comprehensive study of the characteristics of this HRTF, focusing on its source distance dependence. However, to compute the HRTF, Duda and Martens use essentially the same unoptimized method as Cooper and Bauck, modified to use the series solution by Rabinowitz et al. [24].

Instead of an iterative approach, Gumerov and Duraiswami [25], Jarrett et al. [26], and Schymura et al. [27] provide formulas for N as a function of angular wavenumber and sphere radius. These formulas, which typically serve as rules-of-thumb, eliminate the need for iteration and, as we show subsequently, result in lower computation times. However, since they are all derived for far-field sources, they closely resemble the well-known rule-of-thumb proposed by Ward and Abhayapala [28] for representing, using

a finite number of spherical harmonics, a sound field generated within a spherical region of a given radius by an incident plane wave. As we show subsequently, there is a significant dependence of N on source distance for near-field point sources that these existing rules-of-thumb do not take into account, resulting in an inaccurate RS-HRTF for most near-field source distances in addition to computation times that are still longer than necessary.

0.2 Objectives and Approach

From the discussion in Sec. 0.1 we recall that there are two existing approaches for computing the RS-HRTF from its series representation: 1) an *iterative* approach where accuracy may be specified in terms of a parameter, T_c , and 2) an approach where N is determined *a priori* using a formula (typically, a rule-of-thumb) resulting in an HRTF that has a fixed accuracy. In both cases the accuracy of the HRTF is not known in terms of perceptually meaningful quantities such as errors in ITD and ILD.

Our objective here is to derive a formula that approximates the lowest truncation order, N_{\min} , for any given frequency and source distance such that ITDs and ILDs corresponding to the RS-HRTF computed using this formula differ from their respective benchmark values by less than well-known just-noticeable difference (JND) thresholds. Since the formula approximates the *lowest* order, the computation time is close to the minimum required for computing the RS-HRTF that meets the aforementioned accuracy criterion.

After reviewing RS-HRTF theory (Sec. 1) and existing approaches for computing the RS-HRTF (Sec. 2), we first perform, in Sec. 3.1, a brute-force search for N_{\min} using a slightly modified version of the iterative approach developed by Cooper and Bauck [3]. Once we have approximately computed N_{\min} for a wide range of frequencies and source distances, we fit power and rational functions to the data to derive a formula for N_{\min} (Sec. 3.2). Finally, in Sec. 4, we evaluate the proposed formula in terms of computation time and accuracy of the resulting RS-HRTF and compare it to the existing rule-of-thumb formula proposed by Gumerov and Duraiswami [25].

1 THEORY

With the center of the rigid sphere positioned at the origin, we represent the locations of the left and right “ears” in a left-handed Cartesian coordinate system using the position vectors \mathbf{r}_{EL} and \mathbf{r}_{ER} , respectively. Similarly we represent the direction of a sound source by the position vector, $\hat{\mathbf{r}}_{\text{S}}$. Using $\hat{\mathbf{r}}_{\text{S}}$ and \mathbf{r}_{EX} (where X can be L or R), we may compute the so-called *angle of incidence*, Θ_{X} , between \mathbf{r}_{EX} and $\hat{\mathbf{r}}_{\text{S}}$ as

$$\Theta_{\text{X}} = \arctan \left(\frac{\|\hat{\mathbf{r}}_{\text{S}} \times \mathbf{r}_{\text{EX}}\|}{\hat{\mathbf{r}}_{\text{S}} \cdot \mathbf{r}_{\text{EX}}} \right), \quad (1)$$

where $\|\cdot\|$ denotes the Euclidean norm, \times and \cdot denote cross and dot products, respectively, and \arctan is the “four-

quadrant” inverse tangent.¹ The range of Θ_x is $[0, 180^\circ]$. When distinguishing between left and right ears is not required, we use Θ to represent the angle of incidence.

Let μ denote a non-dimensional frequency such that

$$\mu = 2\pi f \frac{r_{RS}}{c} = kr_{RS}, \quad (2)$$

where f is frequency, c is the speed of sound, k is the angular wavenumber, and r_{RS} is the radius of the sphere. For a simple-harmonic point source (with flow of the form $S_\omega e^{i\omega t}$) at a finite distance, r_S , from the origin, the RS-HRTF, H , is given by [5]

$$H(\rho, \mu, \Theta) = -\frac{\rho}{\mu} e^{i\mu\rho} \sum_{m=0}^{\infty} B_m^*(\rho, \mu) P_m(\cos \Theta), \quad (3)$$

where P_m is the m^{th} -degree Legendre polynomial and $\rho = r_S/r_{RS} > 1$ is the non-dimensional distance to the source, and where

$$B_m(\rho, \mu) = (2m + 1) \frac{h_m(\mu\rho)}{h'_m(\mu)}, \quad (4)$$

with B_m^* denoting its complex conjugate, h_m the m^{th} -order spherical-Hankel function of the first kind, and h'_m its first derivative with respect to the argument. For a source infinitely far away, H is given by [3, 5]

$$H(\mu, \Theta) = \frac{1}{\mu^2} \sum_{m=0}^{\infty} A_m^*(\mu) P_m(\cos \Theta), \quad (5)$$

where

$$A_m(\mu) = \frac{(-i)^{m-1} (2m + 1)}{h'_m(\mu)}. \quad (6)$$

Note that taking the complex conjugate of B_m and A_m in Eqs. (3) and (5), respectively, amounts to using spherical-Hankel functions of the second kind (representing waves propagating outwards from the source for the chosen time convention) in these equations. We choose to write our equations in terms of spherical-Hankel functions of the first kind to maintain consistency with the publication by Duda and Martens [5], since we refer to it frequently in the present work.

We also note that Eqs. (3) and (5) are not applicable for $\mu = 0$ and show (with derivations provided in our paper on a formula for low-frequency ILD [16]), from Eq. (3), that

$$\lim_{\mu \rightarrow 0} H(\rho, \mu, \Theta) = \frac{2\rho}{g(\rho, \Theta)} - \rho \times \begin{cases} \ln \left(\frac{g(\rho, \Theta) + 1 - \rho \cos \Theta}{\rho(1 - \cos \Theta)} \right), & \Theta \neq 0, \\ \ln \left(\frac{\rho}{\rho - 1} \right), & \Theta = 0, \end{cases} \quad (7)$$

where $g(\rho, \Theta) = \sqrt{\rho^2 - 2\rho \cos \Theta + 1}$ and $\ln(\cdot)$ denotes taking the natural logarithm of the argument. It may be easily verified from Eq. (5) that

$$\lim_{\mu \rightarrow 0} H(\mu, \Theta) = 1, \quad (8)$$

which is a well-known result.

Equation (7) may be used for computing the value of H as $\mu \rightarrow 0$ exactly for any source location. As described by Sridhar and Choueiri [16], one of the benefits of having such a simple equation is to efficiently compute low-frequency ILD, since the latter is a dominant cue for near-field distance perception in the free-field [14, 29, 30, 31]. In terms of filter design for practical audio applications, knowing the exact value of H as $\mu \rightarrow 0$ may be less critical. Additionally, in the case of the RS-HRTF, using the value of H computed for the smallest non-zero value of μ (depending on the chosen sampling rate) is often a good approximation for H as $\mu \rightarrow 0$. Nevertheless, since Eq. (7) is simple to implement, there is no compelling reason to avoid its use when computing H .

2 COMPUTATION

For $\mu = 0$, we compute H exactly using the closed-form expressions shown in Eqs. (7) and (8). However, for $\mu \neq 0$, we compute H by truncating the infinite series shown in Eq. (3) (for finite ρ) and Eq. (5) (for $\rho \rightarrow \infty$) to a finite order, N . Truncation of the series may be achieved by determining N iteratively or using a *rule-of-thumb* formula. Before we discuss each approach note that we perform all computations without parallelization using MATLAB running on a laptop computer with a 2.3 GHz Intel Core i7 processor. Additionally, all numbers used in the calculation of H are stored in the double-precision, floating-point format [32]. The values of the Legendre polynomials and the spherical-Hankel functions and its derivatives are all computed using the recursion relations shown in APPENDIX A of the paper by Duda and Martens [5]. Functions that come pre-packaged with MATLAB (e.g., native implementations of the aforementioned special functions) are not used to allow fair and transparent comparisons between the different approaches investigated in the present work. For more details, the interested reader may download and review our software [33].

2.1 Iterative Approach

In this approach, proposed by Cooper and Bauck [3], H is computed by increasing N iteratively from 0 until convergence of the infinite series is achieved. Specifically, let $S_M^A(\mu, \Theta)$ given by

$$\begin{aligned} S_M^A(\mu, \Theta) &= \sum_{m=0}^{M-1} A_m^*(\mu) P_m(\cos \Theta) \\ &= \sum_{m=0}^{M-1} \Psi_m^A(\mu, \Theta), \quad \mu \neq 0, \end{aligned} \quad (9)$$

¹This arctan function is commonly represented in many computer programming languages (such as C) and software (like MATLAB) as `atan2`.

denote the M^{th} partial sum of the series in Eq. (5), with $M = 1, 2, \dots$, $\Psi^A = A * P$, and where we ignore the $1/\mu^2$ term. Also, let $d_1(x, y) = |x - y|/|x|$, for any two complex numbers x and y , define a convergence metric. The value of H for $\rho \rightarrow \infty$ and for given values of μ and Θ is then the value of S_M^A obtained by iteratively increasing M up to the largest value that satisfies the conditions

$$d_1(S_{M+2}^A, S_{M+1}^A) > T_c, \text{ and } d_1(S_{M+1}^A, S_M^A) > T_c. \quad (10)$$

The value of H for finite values of ρ is the value of S_M^B obtained similarly, where

$$\begin{aligned} S_M^B(\rho, \mu, \Theta) &= \sum_{m=0}^{M-1} B_m^*(\rho, \mu) P_m(\cos \Theta) \\ &= \sum_{m=0}^{M-1} \Psi_m^B(\rho, \mu, \Theta), \quad \mu \neq 0, \rho > 1, \end{aligned} \quad (11)$$

and $\Psi^B = B * P$.

Duda and Martens [5] use this approach but with a slightly modified convergence metric, d_2 , where $d_2(x, y) = |x/y|$. In this case H is determined as the value of $S_M^{A/B}$ obtained by iterating over M up to the largest value that satisfies the conditions

$$d_2(\Psi_{M+1}^{A/B}, S_{M+2}^{A/B}) > T_c, \text{ and } d_2(\Psi_M^{A/B}, S_{M+1}^{A/B}) > T_c. \quad (12)$$

The superscript A/B indicates that the partial sums in either Eq. (9) or Eq. (11) may be used, depending on whether H is computed for $\rho \rightarrow \infty$ or finite ρ , respectively. Note that in Eqs. (10) and (12) the metrics d_1 and d_2 , respectively, are evaluated for two successive iterations to check for convergence of the series in Eqs. (9) and (11), respectively. This is because convergence is non-monotonic due, in part, to the oscillatory nature of the Legendre polynomials as order increases for a given value (except 1) of the argument.

Although the inequalities in Eqs. (10) and (12) are mathematically identical, we find that they are not equivalent from a numerical computation perspective due to rounding errors that are intrinsic to finite-precision calculations. For instance we find that using Eq. (12) instead of Eq. (10) with $T_c = 0$ results in up to 291 additional terms being used to compute H to the same accuracy for 36 uniformly spaced sources on the horizontal plane (i.e., the plane consisting of sources with zero elevation) and for μ ranging from 0 to approximately 40.

To explain this discrepancy we note that although the additional terms are not all zero each term of the series in Eq. (3) or Eq. (5) when added to the preceding one does not change the sum. Since the computation of H necessarily involves such rounding errors due to the summing of terms in these series solutions, it is better to use d_1 instead of d_2 in practice, especially when computing the RS-HRTF to the highest possible accuracy. It is worth noting however that this discrepancy exists only when T_c is less than the machine epsilon and increases in magnitude with decreasing T_c .

As specified in Sec. 0.1, the RS-HRTF computed using $T_c = 0$ has the highest possible accuracy and as such may be considered a benchmark. We hereafter refer to this HRTF

as the *benchmark RS-HRTF* and quantities such as ITD and ILD that are computed from it as *benchmark ITDs* and *ILDs*, respectively.

2.2 Rule-of-Thumb Approach

In contrast to the iterative approach, H is computed by truncating the infinite series to order N , which is computed *a priori*, typically using a rule-of-thumb formula [25-27]. Gumerov and Duraiswami propose the formula [25]

$$N = [ekr_{RS}] = [e\mu], \quad (13)$$

where e is the base of the natural logarithm and $[\cdot]$ represents rounding to the nearest integer. Similarly, Abhayapala et al. propose [34]

$$N = \left\lceil \frac{\pi}{2} e\mu \right\rceil, \quad (14)$$

for truncating the series expansion of a plane wave into spherical harmonics, which is then used by Schymura et al. [27] to truncate the series representation of the RS-HRTF.

Jarrett et al. instead propose [26]

$$N = \lceil 1.1k_{\max}r_{RS} \rceil = \lceil 1.1\mu_{\max} \rceil, \quad (15)$$

where k_{\max} is the largest wavenumber of interest (corresponding to μ_{\max} , the largest non-dimensional frequency of interest) and $\lceil \cdot \rceil$ represents rounding to the next largest integer.

In Eqs. (13) and (14) we see that N is computed as a function of μ , whereas in Eq. (15), N is computed independent of μ . We find that, when computing H for a wide range of frequencies, computing N independent of μ is not computationally efficient because the summations in Eqs. (3) and (5) converge rapidly with decreasing μ such that the value of N corresponding to μ_{\max} is often much larger than necessary for smaller values of μ . For simplicity we shall only use Eq. (13) for the rule-of-thumb approach in the rest of this paper.

2.3 Comparison of the Two Approaches

In the iterative approach computation time and accuracy of H depend on the user-defined threshold, T_c , which must be real and non-negative, with values closer to 0 corresponding to higher accuracy and longer computation times in general. In the rule-of-thumb approach the accuracy of H cannot typically be varied unless the formula for computing N includes a parameter for doing so, which the existing ones do not have.

To compare computation times for the two approaches, we first time (as described in APPENDIX A.1) the iterative approach for 5 uniformly spaced values of T_c ranging from 10^{-2} to 10^{-10} . In doing so we implicitly compute, for each T_c , truncation order, N , as a function of ρ , Θ , and μ . As a first comparison we use these values of N (instead of Eq. (13), for example) to time the rule-of-thumb approach.² Denoting the computation times for the iterative and rule-of-thumb

²Technically this would not qualify as a “rule-of-thumb” approach as N is not determined using a rule-of-thumb, but we shall ignore this distinction as one having to do with semantics.

approaches as τ_I and τ_{RT} , respectively, we then compute the relative speed increase (RSI) of the latter approach over the former as

$$\text{RSI} = \frac{\tau_I - \tau_{RT}}{\tau_{RT}} \times 100\%, \quad (16)$$

where RSI is a function of ρ , Θ , μ , and, in this case, T_c .

By averaging the RSI over all four parameters, we find that the rule-of-thumb approach is, on average, 43% faster than the iterative one. Based on our implementation of the two approaches, this increase in speed is *only* because the convergence criteria defined in Eq. (10) need not be evaluated at each iteration. Additionally we find that this average RSI is within approximately 10% of most RSI values computed for the different values of ρ , Θ , μ , and T_c , indicating that the RSI does not change significantly with these parameters (even though the wall-clock times for the various calculations can vary substantially). The largest differences exist between computations for finite values of ρ and those for $\rho \rightarrow \infty$, which is likely due to the different sets of equations (e.g., Eq. (3) versus Eq. (5)) used in their computation. Similar differences also exist for some of the smaller values of μ and T_c , especially when $\rho \rightarrow \infty$.

We also perform an alternative comparison where we compute the RSI of the rule-of-thumb approach with N determined using Eq. (13), over the iterative approach with $T_c = 10^{-6}$ and $\rho \rightarrow \infty$, since Gumerov and Duraiswami [25] observe that “excellent agreement” is achieved between the two approaches for these values. In this case we find that the rule-of-thumb approach is, on average, 41% faster than the iterative one. The similarity between the two RSI values computed above indicates that values of N computed using Eq. (13) are likely very similar to those obtained implicitly when using the iterative approach with $T_c = 10^{-6}$ and $\rho \rightarrow \infty$, which is in agreement with the aforementioned observation by Gumerov and Duraiswami.

Although each approach could potentially benefit from additional optimization, such as memoization³ and vectorization, it is clear that using the rule-of-thumb approach to compute H to a desired accuracy has an intrinsic benefit over the iterative one, *provided the appropriate N can be specified a priori*.

We can show that N varies with *both* μ and ρ by plotting, in Fig. 1, N versus μ for a few different values of ρ . The variation of N with μ and ρ shown in the figure is valid for any Θ since we find that N does not vary significantly with Θ . The solid curves correspond to N computed using the iterative approach whereas the dashed curve shows N computed using Eq. (13). In general N increases with increasing μ and decreasing ρ . However for $\rho = 1.5, 2$, and 3 , we see that N is unchanged below $\mu \approx 1.3, 3$, and 5 , respectively, even though we expect N to decrease with decreasing μ . This is a consequence of T_c being specified independently of μ , which, as we later demonstrate (see Sec. 3.1.2), may be accounted for by specifying accuracy using a different

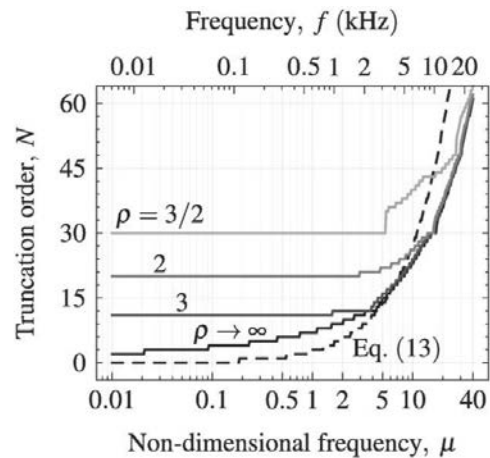


Fig. 1. Solid curves: N versus μ for $\rho = 1.5, 2, 3$ and $\rho \rightarrow \infty$, $T_c = 10^{-6}$, and $\Theta = 170^\circ$. Dashed curve: N computed using Eq. (13). The top axis shows frequency, f , in kHz computed for $r_{RS} = 0.09$ m.

parameter, resulting in the expected variation of N with μ , even for small values of ρ .

Comparing the dashed curve to the solid one corresponding to $\rho \rightarrow \infty$, we find that Eq. (13) is fairly accurate in the neighborhood of $\mu = 7$ but underestimates N for $\mu \ll 7$ and overestimates it for $\mu > 7$. However a more critical issue is that Eq. (13) does not account for the variation of N with ρ , which, as we later demonstrate, has consequences related to the computation time and accuracy of the RS-HRTF.

3 OPTIMIZATION

Although the RS-HRTF is computed more quickly using the rule-of-thumb approach rather than the iterative one, existing formulas for computing N *a priori* suffer from two drawbacks: 1) they do not take into account the variation of N with ρ , and 2) they may be used to compute the RS-HRTF only to a fixed accuracy that may be higher than necessary for most practical applications, resulting in an unnecessarily long computation time. For applications where computational resources are limited, it may be desirable to compute the RS-HRTF to the lowest acceptable accuracy from a perceptual standpoint.

Setting a perceptually motivated accuracy threshold is not trivial because there are many perceptual attributes to consider (for example, Simon et al. [36] list eight such attributes in addition to localization). Even when considering localization alone, spectral cues generated by the pinnae and reflections off the torso, both of which are absent when considering the RS-HRTF only, play important roles in elevation localization and resolving front-back and up-down confusions. However given that ITDs and ILDs are important for localization of lateral sources and in the perception of near-field source distance [37, 38] and that the RS-HRTF is often used to incorporate them in binaural synthesis applications (recall the discussion in Sec. 0), we compute the RS-HRTF that is just accurate enough to ensure that ITD and ILD errors relative to their benchmark values are within

³Memoization is an optimization technique where results of expensive function calls are stored in a cache and returned whenever the function is called with the same inputs [35].

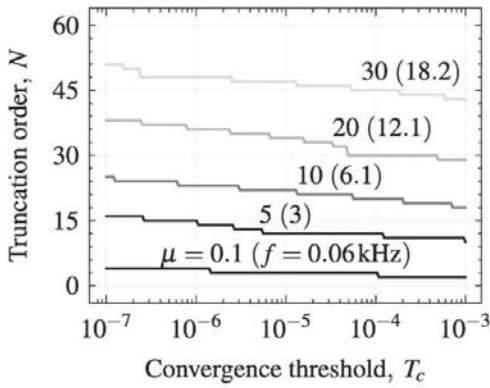


Fig. 2. N versus T_c for $\mu = 0.1, 5, 10, 20,$ and 30 rad, $\rho \rightarrow \infty$, and $\Theta = 170^\circ$. The values in parentheses are frequencies, f , in kHz computed for $r_{RS} = 0.09$ m.

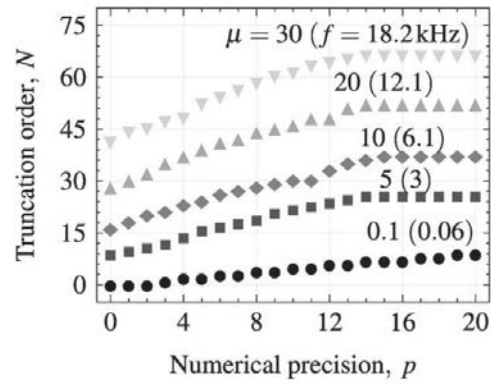


Fig. 3. N versus p for $\mu = 0.1, 5, 10, 20,$ and 30 rad, $\rho \rightarrow \infty$, and $\Theta = 170^\circ$. The values in parentheses are frequencies, f , in kHz computed for $r_{RS} = 0.09$ m.

well-known JND thresholds. Here we derive a formula for N as a function of μ and ρ that minimizes the time required to compute such an RS-HRTF.

To derive the formula we first use a slightly modified version of the iterative approach described in Sec. 2.1 to compute, for a wide range of values for ρ , the RS-HRTF that has approximately the lowest required orders, N_{\min} , while also satisfying the aforementioned accuracy criteria. We then fit power and rational functions to the approximately determined N_{\min} to obtain a formula for N_{\min} as a function of μ and ρ .

3.1 Computation of the Lowest-Order RS-HRTF

It is well known that humans are capable of resolving changes in ITD as low as $10 \mu\text{s}$ for sources close to the median plane [39] and changes in ILD as low as 1 dB for a wide range of baseline ILDs [40] and for frequencies up to approximately 5 kHz [41]. One way to determine N_{\min} is by iteratively increasing the value of T_c from $T_c = 0$ until the largest value is found that results in an RS-HRTF with ITDs and ILDs that differ from their corresponding benchmark values by less than these JND thresholds. However, as we shall now demonstrate, the iterative approach is not suitable for doing so.

3.1.1 Unsuitability of the Iterative Approach

In the iterative approach T_c is only restricted to be real and non-negative. Consequently, unless T_c is changed by an unknown, large enough amount that is a function of μ , ρ , and Θ , the infinite series is truncated to the same order, N . This makes it difficult to iterate over T_c .

Fig. 2 shows a plot of N versus T_c for a few different values of μ and $\rho \rightarrow \infty$. Although N increases with decreasing T_c in general, we see that there is a one-to-many relationship between T_c and N , and T_c cannot be reliably changed to guarantee a change in N . For example, for $\mu = 5$, we see that $N = 12$ for a wide range of T_c but $N = 13$ for a much smaller range. We also see that changing the value of T_c may result in a change in the value of N for some values of μ but not for others unless T_c is varied by a large enough

amount. Although not shown we find a similar behavior for finite values of ρ .

3.1.2 Modified Iterative Approach

Instead of choosing an arbitrary iteration step size for T_c we modify the iterative approach described in Sec. 2.1 to make it more suitable for determining N_{\min} iteratively. Specifically, instead of the convergence metrics defined in Sec. 2.1, we define the metric, $d_3(x, y) = (x - y)_p$, where $(\cdot)_p$ denotes rounding to p decimal digits such that

$$(z)_p := \frac{[\text{Re}(z) \cdot 10^p] + i[\text{Im}(z) \cdot 10^p]}{10^p}, \quad z \in \mathbb{C}, \quad (17)$$

for a given, non-negative integer, p , and where we define $(z)_\infty = z$. We refer to the quantity, p , as *numerical precision* and determine H as the value of $S_M^{A/B}$ obtained by iterating over M up to the largest value that satisfies the conditions

$$d_3 \left(S_{M+M_c}^{A/B}, S_{M+M_c-1}^{A/B} \right) \neq 0, \quad M_c = 1, \dots, 4. \quad (18)$$

Unlike Eqs. (10) and (12), where only two conditions need to be satisfied, we require that four conditions be satisfied because we find that for a few values of μ , ρ , and Θ , checking two conditions is not sufficient to ensure convergence.

Fig. 3 shows a plot of N versus p for the same values of μ shown in Fig. 2. We see that there is generally a one-to-one relationship between p and N for values of p up to a μ -dependent upper limit that also depends on the computer number format. We use the double-precision floating-point format [32], where $p > 15$ and $p \rightarrow \infty$ may be considered equivalent. An *approximate* relationship between p and T_c is $T_c \approx 10^{-p}$ with $p \rightarrow \infty$ corresponding exactly to $T_c = 0$.

3.1.3 Determination of N_{\min}

Since p is integer valued, we may determine N_{\min} by iteratively decreasing p in steps of 1 from $p = 15$ until the smallest value is found that results in an RS-HRTF that satisfies the aforementioned accuracy criteria based on ITD and ILD errors relative to their respective benchmarks. In particular we first compute 10 ms-long rigid-sphere head-

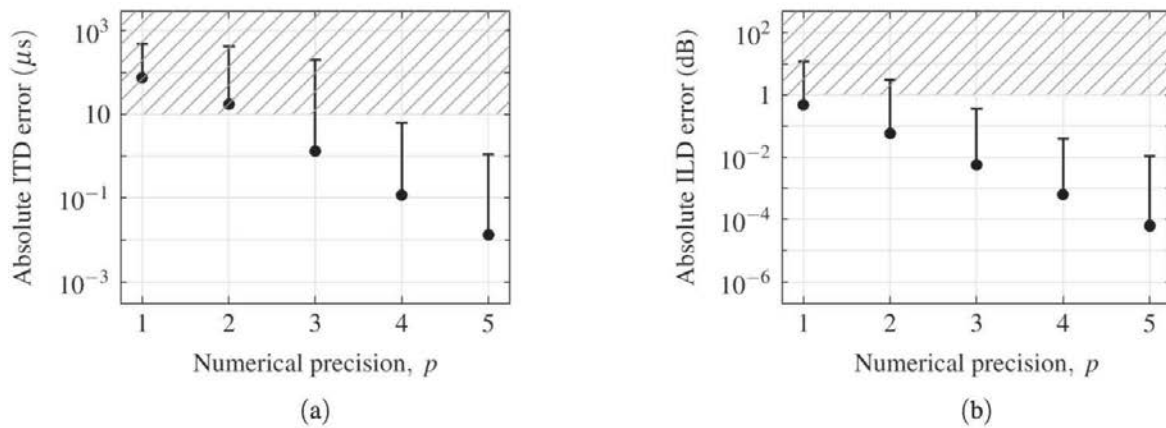


Fig. 4. Absolute error in (a) ITD and (b) ILD, each as a function of p . The circles correspond to the mean while the top of the error bars correspond to the maximum absolute error.

related impulse responses (RS-HRIRs) at a sampling rate of 48 kHz for a sphere with antipodal “ears,” 20 logarithmically spaced values of ρ ranging from 1.5 to 20, and 61 uniformly spaced source directions within 30° on either side of the median plane. From each of these HRIRs we compute ITDs as interaural-phase delays (IPDs) for $0 \leq \mu < 2$. We also compute ILDs as the absolute differences between left and right ear magnitude spectra, in dB, for $0 \leq \mu < 8$ for similar RS-HRIRs corresponding to 61 uniformly spaced source directions within 90° on either side of the median plane.

Fig. 4 shows the mean and maximum absolute errors in ITD (Fig. 4(a)) and ILD (Fig. 4(b)) as a function of p . The means and maxima are taken over all values of μ , ρ , and Θ for which these quantities are computed. The hatched regions correspond to the values of ITD and ILD that are above their respective JND thresholds (shown by dashed lines). From Fig. 4(a) we see that for $p < 3$, the absolute ITD errors are well above the threshold of $10 \mu\text{s}$. When $p = 3$ the errors are in the neighborhood of the threshold value, while for $p > 3$ the errors are below. From Fig. 4(b) we see that for $p = 1$, the absolute ILD errors are well above the threshold of 1 dB. When $p = 2$ the errors are in the neighborhood of the threshold value, while for $p \geq 3$ the errors are below.

From these two figures and the preceding discussion we conclude that when $p \geq 4$ is used to compute the RS-HRTF, values of ITD and ILD computed from the HRTF differ from their corresponding benchmarks by less than their respective JND thresholds. To minimize computation time we choose $p = p_{\min} = 4$. The desired values of N_{\min} approximately correspond to the orders of the RS-HRTF computed using $p = p_{\min}$.

3.2 Derivation of Formula

By analyzing the relationships between N_{\min} and the variables μ , ρ , and Θ we find that N_{\min} varies significantly with μ and ρ but not with Θ . Based on values of N_{\min} computed as described in Sec. 3.1 for $p = p_{\min}$, 201 linearly spaced values of μ between 0 and ~ 33 , 10 logarithmically spaced

values of ρ between 2 and 10, and averaged over 19 uniformly spaced values of Θ between 0 and 90° , we find, using a heuristic approach, that N_{\min} varies with μ and ρ according to the relationship

$$N_{\min} \approx [\alpha + \beta\mu^\gamma], \quad \mu > 0, \quad (19)$$

where α , β , and γ are given by

$$\begin{aligned} \alpha &= \frac{1.41\rho + 3.9}{\rho - 1.36}, \\ \beta &= \frac{2.73\rho - 4.75}{\rho - 1.21}, \\ \gamma &= \frac{0.8\rho - 1.01}{\rho - 1.45}. \end{aligned} \quad (20)$$

As $\rho \rightarrow \infty$ we have $\alpha = 1.41$, $\beta = 2.73$, and $\gamma = 0.8$. Equation (20) is applicable for $\rho \geq 2$. For $1 < \rho < 2$ the value of N_{\min} computed for $\rho = 2$ may be used. For $\mu = 0$ we may compute the RS-HRTF exactly using Eq. (7) (for finite ρ) or Eq. (8) (for $\rho \rightarrow \infty$). The numerical values in each of the rational functions shown in Eq. (20) are computed by generating nonlinear least squares fits to the data using a trust-region algorithm [42]. A summary of the steps involved in computing these values is given in APPENDIX B.1.

Fig. 5 shows the variation of N_{\min} with μ and ρ for $\Theta = 60^\circ$. The values of N_{\min} corresponding to individual data points are computed iteratively using $p = p_{\min}$ as described in Sec. 3.1, whereas those corresponding to the solid lines are computed using Eq. (19). From this figure we see that N_{\min} increases with μ as expected and decreases with increasing ρ , with the latter trend being more prominent for smaller values of μ . We also see that there is a good fit between N_{\min} computed using Eq. (19) and those computed iteratively. Similar behavior is observed for other values of Θ , although the fit is slightly better for some values of Θ than others.

Comparing Fig. 5 to Fig. 1 and referring back to the discussion in Sec. 2.3, we see that, unlike the variation of N with μ shown in Fig. 1, N_{\min} decreases with decreasing μ even for values of ρ as low as $\rho = 2$. This suggests that

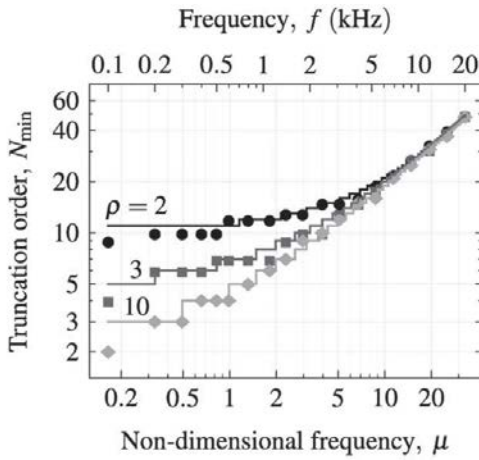


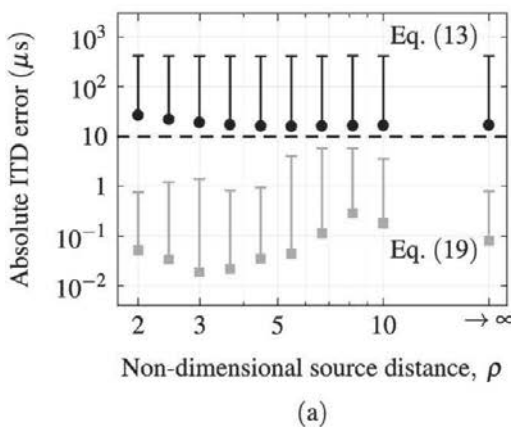
Fig. 5. N_{\min} versus μ for $\Theta = 60^\circ$. The N_{\min} values corresponding to individual data points are computed iteratively whereas those corresponding to the solid lines are computed using Eq. (19). The top axis shows frequency, f , in kHz for $r_{RS} = 0.09$ m.

ρ is more suitable than T_c for specifying accuracy when computing the RS-HRTF iteratively.

4 EVALUATION

To evaluate the accuracy of the proposed formula we compute absolute errors in ITD and ILD estimated from an RS-HRTF computed using the formula (i.e., Eq. (19)) relative to their respective benchmarks. We also compare the formula to the rule-of-thumb (see Eq. (13)) by Gumerov and Duraiswami [25], both in terms of accuracy using the aforementioned error calculations and computation time.

Fig. 6 shows the mean and maximum absolute errors in ITD (Fig. 6(a)) and ILD (Fig. 6(b)) as a function of ρ , where ITD and ILD are computed as described in Sec. 3.1. The means and maxima are taken over all values of μ and Θ for which these quantities are computed (see Sec. 3.1). The dashed lines in the figures correspond to JND thresholds.



From Fig. 6(a) we see that the mean and maximum absolute ITD errors for the proposed formula are consistently lower than those for the rule-of-thumb. We also see that the errors for the rule-of-thumb are above the JND threshold of $10 \mu\text{s}$ for all values of ρ , whereas those for the proposed formula are all below this threshold.

From Fig. 6(b) we see that the mean and maximum absolute ILD errors for the proposed formula are well below the JND threshold of 1 dB, whereas the maximum errors for the rule-of-thumb are above this threshold for most values of ρ , especially in the near-field (i.e., for $\rho \leq 10$, approximately). For $\rho \rightarrow \infty$ the absolute ILD errors for both formulas are comparable. Interestingly we see from Fig. 6(a) that even for $\rho \rightarrow \infty$, the absolute ITD error for the rule-of-thumb is above the JND threshold. Since ITD is computed for $\mu < 2$ only, this indicates that the rule-of-thumb is inaccurate for small μ even for $\rho \rightarrow \infty$, where N is underestimated by no more than 5 (see Fig. 1).

To evaluate the computational efficiency of the two formulas we first measure their computation times as described in APPENDIX A.1. We then use the RSI calculation described in Sec. 2.3 (see Eq. (16)) to compare the computation times with each other and with that obtained when using the iterative approach described in Sec. 3.1 with $p = p_{\min}$. We find that the average RSI when using the proposed formula instead of the rule-of-thumb is 48% and that this value increases to 165% when comparing the proposed formula to the iterative approach.

From these results we see that the proposed formula for N_{\min} as a function of μ and ρ may be used to compute an RS-HRTF that is sufficiently accurate for most practical applications while also providing significant computational savings over existing methods for computing this HRTF.

5 SUMMARY AND CONCLUSIONS

We derived, in Sec. 3.2, a formula (see Eq. (19)) that approximates the truncation order, N_{\min} , that minimizes

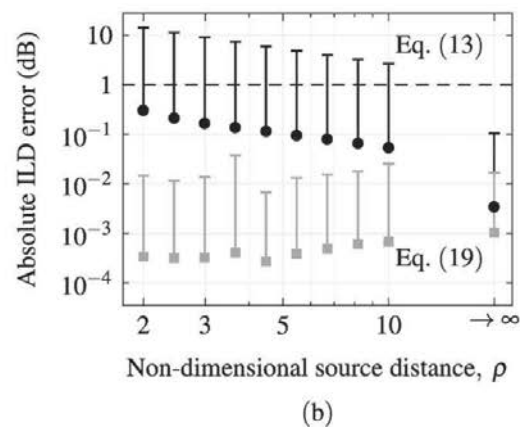


Fig. 6. Absolute error in (a) ITD and (b) ILD, each as a function of ρ . The open circles and squares correspond to the mean while the tops of the error bars correspond to the maximum absolute errors.

the time required to compute the RS-HRTF such that corresponding ITDs and ILDs differ from their respective benchmarks by less than their respective JND thresholds. To evaluate the proposed formula we used it to compute the RS-HRTF for many source distances and directions (see Sec. 4) and showed that mean and maximum absolute ITD and ILD errors relative to their respective benchmarks (defined in Sec. 2.1) remain below JND thresholds. We also showed that the computation is, on average, 48% faster compared to using an existing rule-of-thumb formula (see Eq. (13)). Furthermore we showed that errors in ITD and ILD frequently exceed corresponding JND thresholds when using the rule-of-thumb formula partly because, unlike the proposed formula, it does not account for the dependence of truncation order on source distance.

Many spatial audio applications use the RS-HRTF with the intention of quickly obtaining accurate ITDs and ILDs without significant computational effort. For example the RS-HRTF has been used to extend acoustically measured far-field HRTFs to the near-field [10, 11], where ensuring that low-frequency ILD is correctly computed is critical due to the importance of this cue for near-field distance perception [14]. Since the proposed formula may be used to quickly compute an RS-HRTF that is sufficiently accurate in terms of ITD and ILD, it may be used to implement distance-dependent localization algorithms in devices such as hearing aids where computational resources are limited.

6 ACKNOWLEDGMENT

This work was jointly sponsored by Focal-Jmlab and Tesla, Inc. The authors wish to thank Dr. Joseph G. Tylka for fruitful discussions and reviewing an early draft of the manuscript and the anonymous reviewers for their valuable feedback during the peer-review process. All plots in this paper were generated in Mathematica using the open-source MaTeX software [43] for LaTeX typesetting.

7 REFERENCES

- [1] G. Enzner, Chr. Antweiler, and S. Spors, "Trends in Acquisition of Individual Head-Related Transfer Functions," in J. Blauert (Ed.), *The Technology of Binaural Listening*, Modern Acoustics and Signal Processing, pp. 57–92 (Springer, Berlin, Heidelberg, Berlin/Heidelberg, Germany, 2013), 1st ed. https://doi.org/10.1007/978-3-642-37762-4_3.
- [2] B. Xie, *Head-Related Transfer Function and Virtual Auditory Display*, 2nd ed. (J. Ross Publishing, Fort Lauderdale, FL, 2013). ISBN 978-1-60427-070-9.
- [3] D. H. Cooper and J. L. Bauck, "On Acoustical Specification of Natural Stereo Imaging," presented at the *65th Convention of the Audio Engineering Society* (1980 Feb.), paper 1616. <http://www.aes.org/e-lib/browse.cfm?elib=3759>.
- [4] D. H. Cooper, "Calculator Program for Head-Related Transfer Function," *J. Audio Eng. Soc.*, vol. 30, no. 1/2, pp. 34–38 (1982 Feb.). <http://www.aes.org/e-lib/browse.cfm?elib=3857>.
- [5] R. O. Duda and W. L. Martens, "Range Dependence of the Response of a Spherical Head Model," *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048–3058 (1998). <https://doi.org/10.1121/1.423886>.
- [6] V. R. Algazi, C. Avendano, and R. O. Duda, "Estimation of a Spherical-Head Model From Anthropometry," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 472–479 (2001 Jun.). <http://www.aes.org/e-lib/browse.cfm?elib=10188>.
- [7] R. Sridhar and E. Choueiri, "Capturing the Elevation Dependence of Interaural Time Difference With an Extension of the Spherical-Head Model," presented at the *139th Convention of the Audio Engineering Society* (2015 Oct.), paper 9447. <http://www.aes.org/e-lib/browse.cfm?elib=18003>.
- [8] R. O. Duda, V. R. Algazi, and D. M. Thompson, "The Use of Head-and-Torso Models for Improved Spatial Sound Synthesis," presented at the *113th Convention of the Audio Engineering Society* (2002 Oct.), paper 5712. <http://www.aes.org/e-lib/browse.cfm?elib=11294>.
- [9] D. Y. N. Zotkin, J. Hwang, R. Duraiswaini, and L. S. Davis, "HRTF Personalization Using Anthropometric Measurements," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684)*, pp. 157–160 (New Paltz, NY) (2003 Oct.). <https://doi.org/10.1109/ASPAA.2003.1285855>.
- [10] D. Romblo and B. Cook, "Near-Field Compensation for HRTF Processing," presented at the *125th Convention of the Audio Engineering Society* (2008 Oct.), paper 7611. <http://www.aes.org/e-lib/browse.cfm?elib=14762>.
- [11] A. Kan, C. Jin, and A. van Schaik, "A Psychophysical Evaluation of Near-Field Head-Related Transfer Functions Synthesized Using a Distance Variation Function," *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2233–2242 (2009). <https://doi.org/10.1121/1.3081395>.
- [12] C. P. Brown and R. O. Duda, "A Structural Model for Binaural Sound Synthesis," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 5, pp. 476–488 (1998 Sep.). <https://doi.org/10.1109/89.709673>.
- [13] D. Romblo and H. Bahu, "A Revision and Objective Evaluation of the 1-Pole 1-Zero Spherical Head Shadowing Filter," in *Proceedings of the AES International Conference on Audio for Virtual and Augmented Reality* (2018 Aug.). <http://www.aes.org/e-lib/browse.cfm?elib=19696>.
- [14] D. S. Brungart and W. M. Rabinowitz, "Auditory Localization of Nearby Sources. Head-Related Transfer Functions," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1465–1479 (1999). <https://doi.org/10.1121/1.427180>.
- [15] S. Spagnol, E. Tavazzi, and F. Avanzini, "Distance Rendering and Perception of Nearby Virtual Sound Sources With a Near-Field Filter Model," *Appl. Acoust.*, vol. 115, pp. 61–73 (2017 Jan.). <https://doi.org/10.1016/j.apacoust.2016.08.015>.
- [16] R. Sridhar and E. Choueiri, "A Formula for Low-Frequency Interaural Level Difference," presented at the *149th Convention of the Audio Engineering*

Society (2020 Oct.), paper 10412. <http://www.aes.org/e-lib/browse.cfm?elib=20949>.

[17] G. F. Kuhn, "Model for the Interaural Time Differences in the Azimuthal Plane," *J. Acoust. Soc. Am.*, vol. 62, no. 1, pp. 157–167 (1977). <https://doi.org/10.1121/1.381498>.

[18] R. S. Woodworth and H. Schlosberg, *Experimental Psychology (Rev. Ed.)*, 3rd ed. (Oxford and IBH Publishing, New Delhi, India, 1971). ISBN 9788120417083.

[19] C. G. Buchanan and M. J. Newton, "Dynamic Balanced Model Truncation of the Spherical Transfer Function for Use in Structural HRTF Models," in *Proceedings of the AES International Conference on Audio for Virtual and Augmented Reality* (2018 Aug.). <http://www.aes.org/e-lib/browse.cfm?elib=19673>.

[20] M. Geronazzo, J. Y. Tissieres, and S. Serafin, "A Minimal Personalization of Dynamic Binaural Synthesis With Mixed Structural Modeling and Scattering Delay Networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 411–415 (Barcelona, Spain) (2020 May). <https://doi.org/10.1109/ICASSP40776.2020.9053873>.

[21] T. Gu, S. Spagnol, G. Wersényi, M. Bujacz, O. Bălan, M. Herrera Martínez, A. Moldoveanu, and R. Unnthorsson, "Current Use and Future Perspectives of Spatial Audio Technologies in Electronic Travel Aids," *Wireless Comm. Mobile Comput.*, vol. 2018, 3918284 (2018 Mar.). <https://doi.org/10.1155/2018/3918284>.

[22] G. Courtois, P. Marmaroli, M. Lindberg, Y. Oesch, and W. Balande, "Implementation of a Binaural Localization Algorithm in Hearing Aids: Specifications and Achievable Solutions," presented at the *136th Convention of the Audio Engineering Society* (2014 Apr.), paper 9034. <http://www.aes.org/e-lib/browse.cfm?elib=17181>.

[23] A. Lindau, J. Estrella, and S. Weinzierl, "Individualization of Dynamic Binaural Synthesis by Real Time Manipulation of ITD," presented at the *128th Convention of the Audio Engineering Society* (2010 May), paper 8088. <http://www.aes.org/e-lib/browse.cfm?elib=15385>.

[24] W. M. Rabinowitz, J. Maxwell, Y. Shao, and M. Wei, "Sound Localization Cues for a Magnified Head: Implications From Sound Diffraction About a Rigid Sphere," *Presence: Teleop. Virtual Environ.*, vol. 2, no. 2, pp. 125–129 (1993). <https://doi.org/10.1162/pres.1993.2.2.125>.

[25] N. A. Gumerov and R. Duraiswami, "Computation of Scattering From N Spheres Using Multipole Reexpansion," *J. Acoust. Soc. Am.*, vol. 112, no. 6, pp. 2688–2701 (2002). <https://doi.org/10.1121/1.1517253>.

[26] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor, "Rigid Sphere Room Impulse Response Simulation: Algorithm and Applications," *J. Acoust. Soc. Am.*, vol. 132, no. 3, pp. 1462–1472 (2012). <https://doi.org/10.1121/1.4740497>.

[27] C. Schymura, F. Winter, D. Kolossa, and S. Spors, "Binaural Sound Source Localisation and Tracking Using a Dynamic Spherical Head Model," in *Proceedings of the INTERSPEECH 2015* (Dresden, Germany) (2015 Sep.).

https://www.isca-speech.org/archive/interspeech_2015/i15_0165.html.

[28] D. B. Ward and T. D. Abhayapala, "Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 697–707 (2001 Sep.). <https://doi.org/10.1109/89.943347>.

[29] D. S. Brungart, N. I. Durlach, and W. M. Rabinowitz, "Auditory Localization of Nearby Sources. II. Localization of a Broadband Source," *J. Acoust. Soc. Am.*, vol. 106, no. 4, pp. 1956–1968 (1999). <https://doi.org/10.1121/1.427943>.

[30] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, "Auditory Distance Perception in Humans: A Summary of Past and Present Research," *Acta Acust. United Acust.*, vol. 91, no. 3, pp. 409–420 (2005 May/Jun.). <https://www.ingentaconnect.com/content/dav/aaua/2005/00000091/00000003/art00003>.

[31] A. J. Kolarik, B. C. J. Moore, P. Zahorik, S. Cirstea, and S. Pardhan, "Auditory Distance Perception in Humans: A Review of Cues, Development, Neuronal Bases, and Effects of Sensory Loss," *Atten. Percept. Psychophys.*, vol. 78, no. 2, pp. 373–395 (2016). <https://doi.org/10.3758/s13414-015-1015-1>.

[32] IEEE, "IEEE Standard for Floating-Point Arithmetic," *Standard 754-2008* (2008 Aug.). <https://doi.org/10.1109/IEEESTD.2008.4610935>.

[33] R. Sridhar, "Rigid-Sphere Head-Related Transfer Function Computation Software," <https://doi.org/10.5281/zenodo.4568323>. (accessed Apr. 22, 2021).

[34] T. D. Abhayapala, T. S. Pollock, and R. A. Kennedy, "Characterization of 3D Spatial Wireless Channels," in *Proceedings of the IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, vol. 1, pp. 123–127 (Orlando, FL) (2003 Oct.). <https://doi.org/10.1109/VETEFC.2003.1284991>.

[35] D. Michie, "'Memo' Functions and Machine Learning," *Nature*, vol. 218, no. 5136, pp. 19–22 (1968 Apr.). <https://doi.org/10.1038/218019a0>.

[36] L. S. R. Simon, N. Zacharov, and B. F. G. Katz, "Perceptual Attributes for the Comparison of Head-Related Transfer Functions," *J. Acoust. Soc. Am.*, vol. 140, no. 5, pp. 3623–3632 (2016). <https://doi.org/10.1121/1.4966115>.

[37] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization (Revised Edition)* (The MIT Press, Cambridge, MA, 1996). ISBN 9780262268684, <https://doi.org/10.7551/mitpress/6391.001.0001>.

[38] B. Moore, *An Introduction to the Psychology of Hearing*, 6th ed. (Brill, Leiden, The Netherlands, 2013). ISBN 978-90-04-25242-4, <https://brill.com/view/title/24210>.

[39] A. W. Mills, "On the Minimum Audible Angle," *J. Acoust. Soc. Am.*, vol. 30, no. 4, pp. 237–246 (1958). <https://doi.org/10.1121/1.1909553>.

[40] R. M. Hershkowitz and N. I. Durlach, "Interaural Time and Amplitude jnds for a 500-Hz Tone," *J. Acoust. Soc. Am.*, vol. 46, no. 6B, pp. 1464–1467 (1969). <https://doi.org/10.1121/1.1911887>.

[41] W. A. Yost and R. H. Dye Jr., "Discrimination of Interaural Differences of Level as a Function of Frequency,"

J. Acoust. Soc. Am., vol. 83, no. 5, pp. 1846–1851 (1988). <https://doi.org/10.1121/1.396520>.

[42] MathWorks, Inc., “Least-Squares Fitting,” <https://www.mathworks.com/help/curvefit/least-squares-fitting.html> (accessed May 13, 2020).

[43] S. Horvát, “MaTeX,” (2020 Oct), doi:10.5281/zenodo.4116357, <https://doi.org/10.5281/zenodo.4116357>, Please go to the GitHub repository to download an immediately usable version of this software. Accessed: 2021-04-18.

A.1 MEASURING COMPUTATION TIME

To measure the computation time of a given method, we use the method to compute H for $\rho = 2, 3, 5$ and $\rho \rightarrow \infty$, five uniformly spaced values of Θ between 0 and 180° , and 40 uniformly spaced values of μ between 0.82 and 32.97 (corresponding to frequencies of 500 Hz and 20 kHz, respectively, for $r_{RS} = 0.09$ m). The wall-clock time for each calculation (i.e., for a given selection of ρ , Θ , and μ) is determined using MATLAB’s `tic` and `toc` functions by averaging over 1,000 repetitions of each calculation. This results in a $4 \times 5 \times 40$ array of computation times for a given method. Note that before we time each calculation we perform a few untimed calculations to prevent any first-time computation costs from biasing the result.

B.1 COMPUTATION OF COEFFICIENTS IN EQ. (20)

To compute the coefficients in Eq. (20) we begin by using a trust-region algorithm (see MATLAB Curve Fitting Tool-

box [42]) to determine the values of α , β , and γ as functions of ρ from the values of N_{\min} computed as described in Sec. 3.1 with $p = p_{\min}$. The computed values of α , β , and γ are given in Table 1.

Table 1. Values of α , β , and γ (the coefficients in Eq. (19)) computed for 10 logarithmically spaced values of ρ .

ρ	α	β	γ
2	10.3555	0.9367	1.0695
2.3916	7.2092	1.4175	0.9705
2.8599	5.2887	1.8173	0.9098
3.4200	4.2109	2.0757	0.8765
4.0896	3.3774	2.3068	0.8497
4.8904	2.7869	2.4486	0.8347
5.8480	2.6488	2.4593	0.8331
6.9932	2.5140	2.4658	0.8320
8.3625	2.2972	2.5027	0.8277
10	2.2837	2.4533	0.8325

We find that the generated fit produces a coefficient of determination, R^2 , of 0.9989 and a root-mean-square error (RMSE) of 0.41 when averaging over ρ .

Having computed α , β , and γ for 10 unique values of ρ , we then fit first-order rational functions as shown in Eq. (20) to compute these parameters directly for any given ρ . We determine the coefficients of the rational functions similarly and find that the fits produce R^2 values of 0.997, 0.983, and 0.991 and RMSE values of 0.165, 0.08, and 0.008 for α , β , and γ , respectively.

THE AUTHORS



Rahulram Sridhar



Edgar Y. Choueiri

Rahulram Sridhar is a research assistant in the 3D Audio and Applied Acoustics (3D3A) Laboratory at Princeton University, where he is conducting his Ph.D. research on the individualization of head-related transfer functions for 3D sound reproduction. His research interests include spatial audio, audio signal processing, and psychoacoustics.

•

Edgar Choueiri is a professor of applied physics at the Mechanical and Aerospace Engineering department of Princeton University and associated faculty at the Department of Astrophysical Sciences. He heads Princeton’s Electric Propulsion and Plasma Dynamics Lab and the 3D Audio and Applied Acoustics Lab. His research interests are plasma physics, plasma propulsion, acoustics, and 3D audio.