JASA ARTICLE

Spherical harmonic decomposition of a sound field based on observations along the equator of a rigid spherical scatterer

Jens Ahrens,[1,a)] Hannes Helmholz,[1] David Lou Alon,[2] and Sebastià V. Amengual Garí[2]
[1]*Chalmers University of Technology, 412 96 Gothenburg, Sweden*
[2]*Facebook Reality Labs Research, Facebook, 1 Hacker Way, Menlo Park, California 94025, USA*

ABSTRACT:
We present a method for computing a spherical harmonic representation of a sound field based on observations of the sound pressure along the equator of a rigid spherical scatterer. Our proposed solution assumes that the captured sound field is height invariant so that it can be represented by a two-dimensional (2D) plane wave decomposition (PWD). The 2D PWD is embedded in a three-dimensional representation of the sound field, which allows for perfectly undoing the effect of the spherical scattering object. If the assumption of height invariance is fulfilled, then the proposed solution is at least as accurate as a conventional spherical microphone array of the same spherical harmonic order, which requires a multiple of the number of sensors. Our targeted application is binaural rendering of the captured sound field. We demonstrate by analyzing the binaural output signals that violations of the assumptions that the solution is based on—particularly height invariance and consequently also horizontal propagation—lead to errors of moderate magnitude. © *2021 Acoustical Society of America*. https://doi.org/10.1121/10.0005754

(Received 19 February 2021; revised 31 May 2021; accepted 7 July 2021; published online 4 August 2021)
[Editor: Wonkyu Moon] Pages: 805–815

I. INTRODUCTION

Spherical microphone arrays (SMAs) have been shown to be a convenient solution for analyzing and capturing spatial sound fields whereby they exhibit properties that are independent of the angle of sound incidence (Abhayapala and Ward, 2002; Meyer and Elko, 2002; Rafaely, 2004; Zotkin *et al.*, 2010). Arrays that comprise a rigid spherical scatterer on which the microphones are placed are most convenient. Spherical harmonics (SHs) have been shown to constitute a flexible representation of the captured sound field that facilitates applications like beamforming (Zotkin *et al.*, 2010) and binaural rendering (Rafaely and Avni, 2010). Binaural rendering computes the signals that would arise at the ears of a listener when exposed to the sound field captured by the array. This will be the targeted application in our work. It requires the listener's head-related transfer functions (HRTFs) and can be performed either directly in the SH domain (Rafaely and Avni, 2010) or by sampling of the plane wave decomposition (PWD) of the sound field (Li and Duraiswami, 2006) by means of beamforming.

SMAs have the downside that the direction-independent spatial resolution that they exhibit requires a significant number of microphones to be distributed over the surface of the spherical scattering object. When it comes to capturing or reproducing sound scenes for virtual reality, there may be scenarios in which it may be sufficient to restrict the considerations to the horizontal plane (Ahrens, 2012; Galindo *et al.*, 2020) because this is the most common real-world scenario, and the human auditory system is optimized for it (Blauert, 1996). Circular microphone arrays with appropriate scatterers therefore exhibit considerable potential in this context as they may achieve the same spatial resolution with an order of magnitude fewer microphones.

Cylindrical scatterers with a circular microphone array along the circumference and partly also with additional microphones were investigated (Betlehem and Poletti, 2019; Galindo *et al.*, 2020; Kaiser *et al.*, 2012; Parthy *et al.*, 2011; Teutsch and Kellermann, 2006; Trevino *et al.*, 2014; Zotkin *et al.*, 2010). The evaluation was performed mostly with respect to beamforming performance and sound source localization capabilities. The evaluation in Betlehem and Poletti (2019) focused on the accuracy of the extracted sound field representation. Kaiser *et al.* (2012), Meyer (2001), and Tiana-Roig *et al.* (2011) investigated beamforming based on a circular array along the equator of a spherical scatterer.

SH decomposition based on a circular array on a planar scatterer was presented in Zaunschirm and Zotter (2014) and based on concentric circular arrays without a scatterer in Chen *et al.* (2015). A method for reproduction of a sound field captured by a circular microphone array on the equator of a spherical scatterer by a concentric circular loudspeaker array that was partially formulated in SHs was presented in Koyama *et al.* (2016).

We aim at a SH representation of the captured sound field for being able to apply tracking of the listener's head orientation with respect to arbitrary rotations during binaural rendering. In our work, we assume a circular microphone array along the equator of a spherical scattering object similarly to Kaiser *et al.* (2012), Koyama *et al.* (2016), Meyer (2001), Tiana-Roig *et al.* (2011), and Weller *et al.* (2011).

[a)]Electronic mail: jens.ahrens@chalmers.se, ORCID: 0000-0002-1706-3564.

We will refer to it as an *equatorial microphone array* (EMA). Spherical scattering objects are compact and easy to manufacture, and the mathematical treatment can be performed analytically, which is not possible with other simple geometries like finite-length cylinders.

In Algazi *et al.* (2004), a method for binaural rendering of equatorial array signals without sound field decomposition is presented that maps microphone signals directly to the ears of the listener. The scattering object thereby produces interaural cues that cannot be individualized. Head tracking is possible only along the azimuth.

A least-squares solution to the problem at hand was presented in Weller *et al.* (2011). It uses a two-dimensional (2D) formulation of the problem and was evaluated only for plane waves. The applied amount of regularization was crucial for achieving spectrally balanced loudspeaker-based reproduction.

We presented an analytical solution to the problem in Ahrens *et al.* (2021c). It also comprised a 2D representation of the captured sound field in which the three-dimensional (3D) scattering off the spherical scatterer was removed whereby the sound sources that composed the captured sound field were assumed to be distant. It was not discussed in detail in the original publication but reflected by the accompanying audio examples (Ahrens *et al.*, 2021b) that a violation of the assumption of large distance of the sound sources (and of indirect sound sources like reflective surfaces) produced an intolerable low-frequency boost that can be on the order of several tens of dB. As an example, this boost occurred below 100 Hz for an eighth-order EMA with a radius of 0.0875 m for sources closer than 4 m. The boost became more pronounced and covered a larger frequency range if the sources were even closer. This was a fundamental limitation that prevented deploying this solution in the targeted application scenario. The solution from Weller *et al.* (2011) may be expected to suffer from similar limitations as it uses a similar 2D representation.

Other conversions between 2D and 3D representations have been presented (Ahrens and Spors, 2008, 2012; Thomas *et al.*, 2014), but these cannot be exploited in the present context. The solution that we present here is based on a formulation of the problem that uses a 3D representation of the effect of the microphone array's scattering body while assuming that the captured sound field is 2D. A similar formulation was presented in Koyama *et al.* (2016), but the SH coefficients of the captured sound field were not computed explicitly. The formulation does not make any assumptions on the distances of the sound sources that produce the captured sound field.

II. APPROACH

A. Conventional spherical microphone arrays

We briefly revisit conventional SMAs in this section as this facilitates the presentation of the EMA solution. SMAs with pressure microphones distributed over the surface of a rigid spherical scatterer have been shown to be most favorable if a SH representation of the captured sound field is sought (Rafaely, 2005). We limit our considerations to this scenario. We use a spherical coordinate system defined through azimuth α, colatitude β, and radial distance r from the coordinate origin. All quantities that are observed on the surface of the spherical scatterer are denoted with the superscript $^{\text{surf}}$ in this article.

In general, the sound pressure $S^{\text{int}}(\beta, \alpha, r, \omega)$ in an interior domain, i.e., the sound pressure inside a spherical domain centered at the coordinate origin that is free of sound sources and free of reflecting boundaries, is given by (Gumerov and Duraiswami, 2005)

$$S^{\text{int}}(\beta, \alpha, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \breve{S}_n^m(\omega)\, j_n\left(\omega \frac{r}{c}\right) Y_n^m(\beta, \alpha). \tag{1}$$

$\breve{S}_n^m(\omega)$ are the SH coefficients of the incident sound field, $j_n(\cdot)$ is the nth-order spherical Bessel function, $\omega = 2\pi f$ is the angular frequency, f is the frequency in Hz, and c is the speed of sound in m/s. $Y_n^m(\beta, \alpha)$ are the surface SHs, of which several different definitions exist. We choose to use the complex definition [Gumerov and Duraiswami (2005), Eq. (2.1.59)]

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

in which $P_n^{|m|}(\cdot)$ are the associated Legendre functions.

The presence of a rigid spherical scatterer of radius R that is centered at the coordinate origin affects the sound field given by (1). It is then represented in the domain $r \geq R$ by (Gumerov and Duraiswami, 2005)

$$S(\beta, \alpha, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \breve{S}_n^m(\omega)\, b_n\left(\omega \frac{r}{c}, R\right) Y_n^m(\beta, \alpha), \tag{3}$$

whereby $b_n(\cdot)$ is given by [Gumerov and Duraiswami (2005), Eq. (4.2.10)]

$$b_n\left(\omega \frac{r}{c}, R\right) = j_n\left(\omega \frac{r}{c}\right) - \frac{j_n'\left(\omega \frac{R}{c}\right)}{h_n'^{(2)}\left(\omega \frac{R}{c}\right)} h_n^{(2)}\left(\omega \frac{r}{c}\right). \tag{4}$$

$h_n^{(2)}(\cdot)$ denotes the nth-order spherical Hankel function of second kind, and the prime denotes differentiation with respect to the argument. We choose the exponent of the temporal Fourier transform such that $h_n^{(2)}(\cdot)$ represents outgoing waves. Note that the coefficients $\breve{S}_n^m(\omega)$ are identical in (1) and (3).

Observing $S(\beta, \alpha, r, \omega)$ on the surface of the scatterer where $r = R$ applies allows for simplifying (4) to [Gumerov and Duraiswami (2005), Eq. (4.2.13)]

https://doi.org/10.1121/10.0005754

$$b_n\left(\omega\frac{R}{c},R\right) = -\frac{\mathrm{i}}{\left(\omega\frac{R}{c}\right)^2}\frac{1}{h_n'^{(2)}(\omega R/c)}. \tag{5}$$

$\breve{S}_n^m(\omega)$ are the SH coefficients of the incoming sound field, the computation of which is the ultimate goal of this endeavor. $\breve{S}_n^m(\omega)$ can be obtained from integrating the sound pressure $S^{\mathrm{surf}}(\beta,\alpha,R,\omega)$ observed on the surface of the scatterer weighted with the complex conjugate * of the SH basis functions as (Williams, 1999)

$$\breve{S}_n^m(\omega) = \frac{1}{b_n\left(\omega\frac{R}{c},R\right)}\oint_O S^{\mathrm{surf}}(\beta,\alpha,R,\omega)\; Y_n^m(\beta,\alpha)^*\;\mathrm{d}\Omega. \tag{6}$$

O denotes the spherical surface, and $\mathrm{d}\Omega$ is an infinitesimal spherical surface element. $\breve{S}_n^m(\omega)$ represents the captured sound field with the effect of the scatterer removed.

Conventional SMAs comprise microphones distributed over the entire surface of the scattering object so that the integral in (6) can be approximated discretely by means of quadrature. A consequence of the quadrature is that the SH coefficients can be retrieved only up to a given maximum SH order N such that $n, |m| \le N$.

The spatial domain in which the normalized squared error of an order-limited sound field representation is 4% (–14 dB) or smaller is inside the spherical region around the coordinate origin with radius r such that the argument $(\omega r/c)$ of the spherical Bessel function $j_n(\cdot)$ is smaller than the highest order N contained in the SH expansion [Gumerov and Duraiswami (2005), p. 427]. Consequently, the minimum order limit N_1 at which the normalized squared error is smaller than that is given by (Ward and Abhayapala, 2001)

$$N_1 = \left\lceil \frac{\omega}{c} r_{\max} \right\rceil, \tag{7}$$

whereby $\lceil\cdot\rceil$ denotes the ceiling function. Evaluating (7) for a human head of radius $r_{\mathrm{head}} = 0.0875\,\mathrm{m}$ up to a frequency of $f = 20\,\mathrm{kHz}$ yields a minimum SH order of $N_1 = 33$.

An order limit N_2 that produces negligible truncation error is given by (Zotkin *et al.*, 2010)

$$N_2 \ge \frac{\mathrm{e}\,\frac{\omega}{c} r_{\max} - 1}{2} + 1, \tag{8}$$

which yields $N_2 \ge 44$ for the human head.

B. Binaural rendering

Once the SH coefficients $\breve{S}_n^m(\omega)$ of the captured sound field are known, the sound field can be rendered binaurally. This means that the head of a listener whose HRTFs are known can be virtually placed into the sound field at the position of the SMA. The signal $D(\omega)$ that occurs at a given ear is computed as (Rafaely and Avni, 2010)

$$D(\omega) = \sum_{n=0}^{N}\sum_{m=-n}^{n} a_m\;\breve{S}_n^{-m}(\omega)\;\breve{H}_n^m(\omega), \tag{9}$$

where $\breve{H}_n^m(\omega)$ are the SH coefficients of the HRTFs of that ear. $a_m = (-1)^m$ for the present definition (2) of the SH basis functions, or $a_m = 1$ if the SH definition does not contain the factor $(-1)^m$ [e.g., in Rafaely and Avni (2010)]. A matrix formulation of (9) was presented in Zaunschirm *et al.* (2018).

$\breve{S}_n^m(\omega)$ and $\breve{H}_n^m(\omega)$ can be rotated relative to each other to enable head tracking in the rendering.

Although physically accurate binaural rendering requires an order of $N = 33$ or higher, practical arrays exhibit an order of typically less than 10. HRTFs are usually available at much higher orders so that the array is the limiting factor. The perceptual implication of the errors, which occur exclusively at high frequencies, is small to significant depending on the maximum order N that is used. The perceptual impairment becomes very small for $N \ge 8$ if all sound sources are located in the horizontal plane (Ahrens and Andersson, 2019; Bernschütz, 2016; Lübeck *et al.*, 2020; Zaunschirm *et al.*, 2018). The frequency above which artifacts occur can be determined from (7).

C. The equatorial array

The SMA methodology for computing the SH coefficients of the captured sound field cannot directly be applied with EMAs. The main reason is that the microphone placement does not constitute a quadrature of the spherical surface with respect to which the basis functions $Y_n^m(\beta,\alpha)$ are orthogonal. Discretizing the integral in (6) therefore does not yield the desired coefficients.

All SHs vanish on the equator, i.e., $Y_n^m(\pi/2,\alpha) = 0$, whenever $n+m$ is odd (Gumerov and Duraiswami, 2005). This prevents certain information on the sound field from being retrieved. Also, the EMA cannot differentiate two sound fields that are copies of one another mirrored on the horizontal plane.

What we can do is perform an integration of the sound pressure weighted with a complex exponential along the equator, which is equivalent to a Fourier series expansion or circular harmonic (CH) expansion given by [Rabenstein *et al.* (2014), Sec. 4.32.3.4]

$$S_m^{\circ\,\mathrm{surf}}(\pi/2,R,\omega) = \frac{1}{2\pi}\int_0^{2\pi} S^{\mathrm{surf}}(\pi/2,\alpha,R,\omega)\;\mathrm{e}^{-\mathrm{i}m\alpha}\;\mathrm{d}\alpha. \tag{10}$$

This yields the coefficients $S_m^{\circ\,\mathrm{surf}}(\pi/2,R,\omega)$, which we refer to as CH coefficients of $S^{\mathrm{surf}}(\pi/2,\alpha,R,\omega)$ in the remainder. $S^{\mathrm{surf}}(\pi/2,\alpha,R,\omega)$ can be reconstructed from $S_m^{\circ\,\mathrm{surf}}(\pi/2,R,\omega)$ using

$$S^{\mathrm{surf}}(\pi/2,\alpha,R,\omega) = \sum_{m=-\infty}^{\infty} S_m^{\circ\,\mathrm{surf}}(\pi/2,R,\omega)\;\mathrm{e}^{\mathrm{i}m\alpha}. \tag{11}$$

The integral in (10) can be approximated by a sum over the discrete microphone positions. One consequence of this discretization is that only modes up to a given maximum order $|m| \leq N$ can be obtained so that the reconstruction (11) is order limited similarly to the SMA case (cf. Sec. II A).

To establish the relation between the CH coefficients $S_m^{\circ\,\mathrm{surf}}(R,\omega)$ from (10) and the SH coefficients $\breve{S}_n^m(\omega)$ from (3), we change the order of summations in (3) to reveal the CH coefficients $\overset{\circ}{S}{}_m^{\mathrm{surf}}(R,\omega)$ as

$$S^{\mathrm{surf}}(\pi/2,\alpha,R,\omega) = \sum_{m=-\infty}^{\infty}\sum_{n=|m|}^{\infty}\breve{S}_n^m(\omega)\; b_n\!\left(\omega\frac{R}{c},R\right) Y_n^m(\pi/2,0)\,\mathrm{e}^{\mathrm{i}m\alpha}, \tag{12}$$

whereby the relation $Y_n^m(\pi/2,\alpha) = Y_n^m(\pi/2,0)\,\mathrm{e}^{\mathrm{i}m\alpha}$ is exploited. Comparing (12) with (11) reveals that the coefficients $S_m^{\circ\,\mathrm{surf}}(\pi/2,R,\omega)$, which we will be denoting $S_m^{\circ\,\mathrm{surf}}(R,\omega)$ in the remainder for convenience and that we obtain from the microphone signals via (10), are given by

$$S_m^{\circ\,\mathrm{surf}}(R,\omega) = \sum_{n=|m|}^{\infty}\breve{S}_n^m(\omega)\; b_n\!\left(\omega\frac{R}{c},R\right) Y_n^m(\pi/2,0). \tag{13}$$

In the remainder of this derivation, we assume that the captured sound field is height invariant, i.e., 2D. We will analyze the consequences if the captured sound field does not fulfill this assumption in Sec. III.

For a plane wave with unit amplitude propagating into azimuth angle θ parallel to the horizontal plane, $\breve{S}_n^m(\omega)$ is given by [Gumerov and Duraiswami (2005), Eq. (2.3.6)]

$$\breve{S}_n^m(\omega) = 4\pi\mathrm{i}^{-n}\; Y_n^m(\pi/2,\theta)^*, \tag{14}$$

so that

$$S_m^{\circ\,\mathrm{surf}}(R,\omega) = \mathrm{e}^{-\mathrm{i}m\theta}\sum_{n=|m|}^{\infty} 4\pi\mathrm{i}^{-n} b_n\!\left(\omega\frac{R}{c},R\right)\left[Y_n^m(\pi/2,0)\right]^2. \tag{15}$$

Note that all terms that the sum in (15) is composed of are mathematical functions that can be evaluated with arbitrary precision.

We know from the 2D plane wave decomposition that any interior 2D, i.e., any height-invariant, sound field $S^{\mathrm{int}}(\vec{x},\omega)$ can be represented by a continuum of propagating plane waves as [Rabenstein *et al.* (2014), Sec. 4.32.3.4.4]

$$S^{\mathrm{int}}(\vec{x},\omega) = \frac{1}{2\pi}\int_0^{2\pi}\bar{S}(\theta,\omega)\;\mathrm{e}^{-\mathrm{i}\frac{\omega}{c}r\cos(\theta-\alpha)}\;\mathrm{d}\theta, \tag{16}$$

whereby $\bar{S}(\theta,\omega)$ are the plane wave coefficients and θ is the azimuthal propagation direction of a given plane wave. Despite the fact that (16) represents a sound field in terms of propagating plane waves, it is a general representation of 2D sound fields.

Equation (16) uses an infinite number of plane waves, each with a unique complex amplitude, to represent a sound field. We can reformulate (15) accordingly by employing a discrete sum over an infinite set of plane waves with complex amplitude $X_l(\omega)$ and propagation direction $(\pi/2,\theta_l)$ as

$$S_m^{\circ\,\mathrm{surf}}(R,\omega) = \underbrace{\sum_{l=1}^{\infty} X_l(\omega)\;\mathrm{e}^{-\mathrm{i}m\theta_l}}_{=\overset{\circ}{\bar{S}}_m(\omega)} \times \sum_{n=|m|}^{\infty} 4\pi\mathrm{i}^{-n}\, b_n\!\left(\omega\frac{R}{c},R\right)\left[Y_n^m(\pi/2,0)\right]^2. \tag{17}$$

We derive in the Appendix that the term marked by the curly brace is identical to the CH coefficients $\overset{\circ}{\bar{S}}_m(\omega)$ of the plane wave coefficients $\bar{S}(\theta,\omega)$ in (16). An important property of (17) is that the unknown quantity $\overset{\circ}{\bar{S}}_m$ is solely dependent on the SH mode m (and not on the SH degree n).

Equation (17) can be solved straightforwardly for $\overset{\circ}{\bar{S}}_m(\omega)$ as

$$\overset{\circ}{\bar{S}}_m(\omega) = \frac{S_m^{\circ\,\mathrm{surf}}(R,\omega)}{\sum_{n'=|m|}^{\infty} 4\pi\mathrm{i}^{-n'}\; b_{n'}\!\left(\omega\frac{R}{c},R\right)\left[Y_{n'}^m(\pi/2,0)\right]^2}. \tag{18}$$

$Y_{n'}^m(\pi/2,0) = 0\;\forall\; n'+m = \text{odd}$, which allows for ignoring certain terms in (18). The summation has to be bounded in practice. We tested different limits and found that the accuracy of the sound field reconstruction from the array signals that we evaluate in Sec. III B 1 does not increase for limits higher than the decomposition order N. We therefore choose this limit in the remainder.

The desired SH coefficients $\breve{S}_n^m(\omega)$ of the impinging sound field can be computed from $\overset{\circ}{\bar{S}}_m(\omega)$ as [cf. (3) and (17)]

$$\breve{S}_n^m(\omega) = \overset{\circ}{\bar{S}}_m(\omega)\; 4\pi\mathrm{i}^{-n}\; Y_n^m(\pi/2,0), \tag{19}$$

from which the ear signals can be computed via (9).

Similarly to Ahrens *et al.* (2021c) and Weller *et al.* (2011), (18) employs a 2D representation of the captured sound field. However, the fundamental difference from the previous solutions is that (18) accounts correctly for the 3D scattering as represented by the factor $b_n(\cdot)$ and therefore does not require the sound sources that produce the captured sound field to be located at specific distances for successful removal of the effect of the scatterer. Equation (18) inherently assumes height invariance of the captured field. This is indeed a fundamental limitation that generally prevents the

https://doi.org/10.1121/10.0005754

reconstruction of the sound pressure at locations off the equator. We analyze this aspect in Sec. III.

The denominator of (18) is always non-zero. Its inverse may be interpreted as the EMA equivalent of the so-called *radial filters*, which are applied with conventional SMAs [$1/b_n(\omega R/c, R)$ in (6)]. As depicted in Fig. 1 (left), the radial filters of SMAs exhibit large gains particularly at low frequencies that may require regularization, for example, through soft clipping (Bernschütz, 2016) as

$$g_{\mathrm{lim}}(\omega) = \frac{2g_{\mathrm{max}}}{\pi}\frac{g(\omega)}{|g(\omega)|}\arctan\left(\frac{\pi}{2g_{\mathrm{max}}}|g(\omega)|\right), \tag{20}$$

whereby $g(\omega)$ denotes the radial filter and g_{max} the user-defined maximum permitted magnitude. The effect of the regularization is an attenuation of the magnitude at higher orders n. This has been shown to have no perceptual consequences with binaural rendering of SMA signals when the captured sound field is rendered binaurally at orders of, say, $N = 8$ or higher and sound sources are in the horizontal plane (Ahrens and Andersson, 2019; Bernschütz, 2016; Zaunschirm *et al.*, 2018). The particular choice of the radial filter gain-limit is not important as long as the limit is chosen high enough so that the limitation does not affect that part of the filter transfer function where the magnitude increases toward higher frequencies. Then sufficient spatial information is preserved in the lower frequencies, while the global magnitude is not affected at high frequencies. The limit should not be chosen too high in order to prevent detrimental effects from sensor mismatch or additive noise from the hardware (Bernschütz, 2016; Helmholz *et al.*, 2020; Rafaely, 2005).

The EMA radial filters exhibit very similar qualitative properties as evident from Fig. 1 (right) so that all observations regarding SMA radial filters apply. We will employ soft-limiting in Sec. III B. Note that the definition of the SMA radial filters that we use in this article differs by a factor of 4π in terms of the magnitude from the one used in, for example, Ahrens and Andersson (2019), Bernschütz (2016), and Rafaely (2005). This corresponds to a difference of 22 dB, i.e., a clipping magnitude of 40 dB in our formulation is equivalent to 18 dB ibidem for SMAs.

III. RESULTS

We will use the example of a spherical scatterer with radius $R = 0.0875$ m for the evaluation of the proposed solution. This is similar to the radius of a human head and has proven useful when binaural rendering of the captured sound field is intended (Bernschütz, 2016). We will target an eighth-order SH decomposition of the captured sound field, which is also typical for binaural rendering applications. We use the minimum required number of microphones for the EMA of $L = 2N + 1 = 17$ [Ahrens (2012), Sec. 4.4] with equiangular spacing.

On many occasions, we will compare the performance of the EMA to that of a conventional SMA of the same order of $N = 8$. SMAs require at least $L = (N + 1)^2$ microphones depending on the microphone placement (Zotter, 2009). This leads to $L = 81$ for the present case. We use a Lebedev grid (Lebedev, 1977) with 110 nodes here for convenience as it is common in the SMA literature.

Cf. Fig. 2 for an illustration of the geometry. Note the different numbers of microphones (110 vs 17). Equation (7) tells us that for both arrays, SMA and EMA, the sound field representation will not be accurate above a frequency of approximately 5.1 kHz. We will comment on how our observations may be generalized to other array configurations in Sec. III A.

A. Sound field reconstruction

We evaluate the accuracy of the capabilities of SMA and EMA of reconstructing the captured sound field using the average normalized error $E(\omega)$ given by

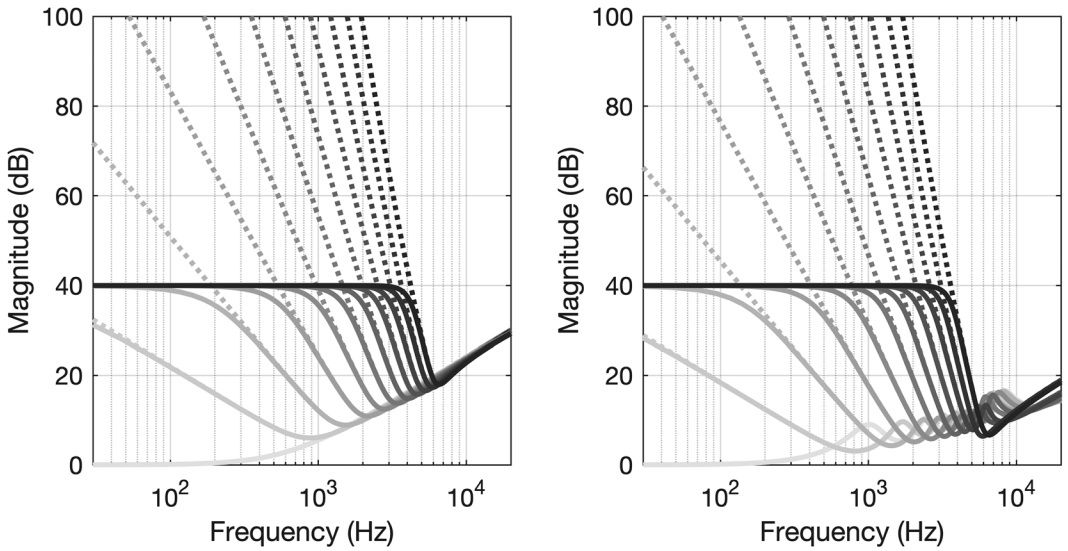

FIG. 1. Magnitude of the radial filters of SMA (*left*) and EMA (*right*) for orders $n = 0 \ldots 10$. The dotted lines represent the theoretical filters, and the solid lines represent the filters soft-clipped at 40 dB. Darker shading indicates higher order.

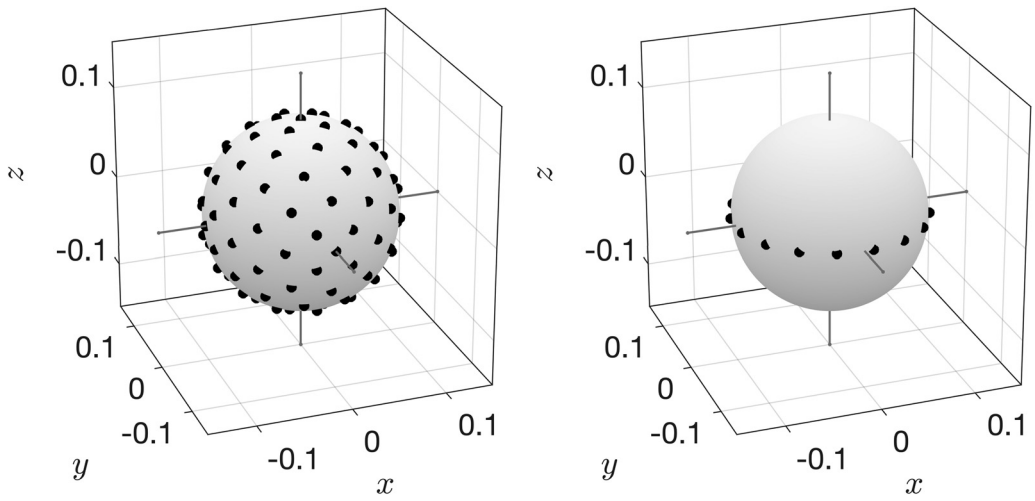

FIG. 2. Illustration of the microphone array geometries. Both support a maximum order of $N = 8$. Left: 110-node Lebedev grid. Right: $2N + 1 = 17$-node equatorial grid. The black dots denote the locations of the microphones.

$$E(\omega) = \frac{1}{Q}\sum_{q=1}^{Q}\left|\left(\hat{S}(\vec{x}_q,\omega) - S(\vec{x}_q,\omega)\right)/S(\vec{x}_q,\omega)\right|, \quad (21)$$

whereby $\hat{S}(\vec{x}_q,\omega)$ denotes the sound pressure at each of the Q evaluated locations $\vec{x}_q$ on the scatterer as reconstructed by a given array and method using (3) for $n, |m| \leq N$. $S(\vec{x}_q,\omega)$ denotes the true sound pressure. We used 50 evaluation points $\vec{x}_q$ equally spaced along the equator and 1250 points distributed over the whole spherical scattering surface based on the Gauss–Legendre grid.

For the SMA, the SH coefficients are computed via discretizing the integral in (6). For the EMA, they are computed using (19), which uses a discretization of the integral in (10). We do not apply a limitation of the radial filter gain in this section.

Figure 3 depicts $E(\omega)$ for SMA and EMA for an impinging horizontally propagating plane wave as well as for spherical waves originating from the horizontal plane at different distances. The impinging sound field was of order $N_2 = 45$, which fulfills (8) so that the error due to the order truncation is negligible. Figure 3 (bottom) depicts the results for fifth-order SMA and EMA for comparison.

Both SMA and EMA exhibit the same fundamental limitations toward high frequencies that are represented by (7). Setting $r_{\max} = R$ in (7) and solving it for f yields 5.1 kHz for the eighth-order arrays and 3.2 kHz for the fifth-order arrays. The SMA is equally accurate anywhere on the sphere for both plane and spherical waves (cf. Fig. 3, blue lines). Only the horizontally propagating plane wave fulfills the assumption of height invariance that the EMA solution is based on [cf. Eq. (17)]. In this case, the EMA sound field reconstruction is at least as accurate as the SMA reconstruction anywhere on the sphere.

Spherical waves violate the assumption of height invariance stronger the closer the point from which they originate is to the EMA. This reduces the accuracy of the EMA sound field reconstruction as evident from Fig. 3 (middle; orange, yellow, and purple lines). The accuracy is lower off the equator than on it, where it is reduced only mildly. Figure 4 (top) depicts the sound pressure distribution on the spherical scatterers due to a spherical wave originating from a distance of 1 m. Figure 4 (middle) depicts the sound pressure as

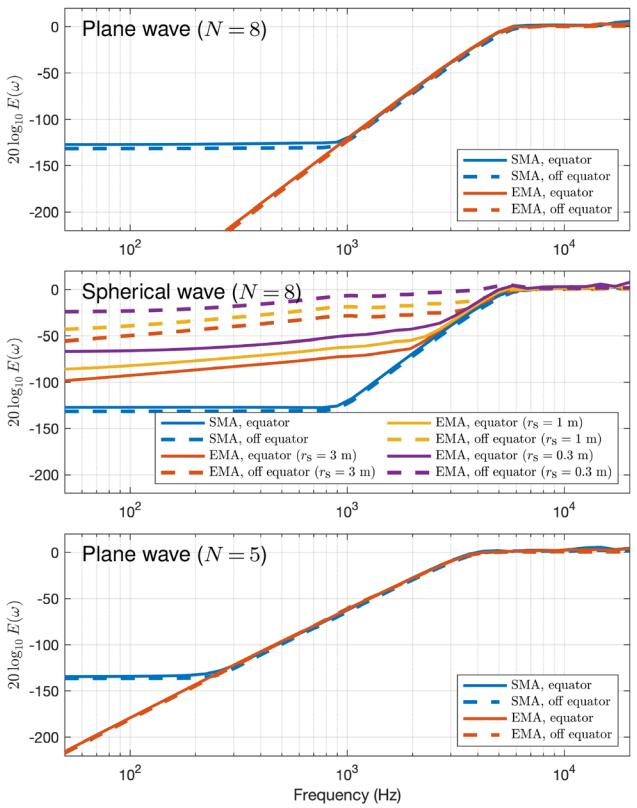

FIG. 3. (Color online) Top: $20\log_{10}E(\omega)$ according to (21) for an impinging horizontally propagating plane wave for an eighth-order conventional SMA (blue) and an eighth-order EMA (orange). The solid lines depict $20\log_{10}E(\omega)$ evaluated exclusively along the equator of the spherical scatterer. The dashed lines depict $20\log_{10}E(\omega)$ evaluated at the remaining spherical surface. Middle: Same as the top plot but for impinging spherical waves originating from distances of $r_s = 3$, 1, and 0.3 m. The SMA result is independent of the source distance. Bottom: Same as the top plot but for a fifth-order EMA with 11 microphones and a fifth-order SMA with the microphones placed on a 50-node Lebedev grid with the same radius like in the top plot.

https://doi.org/10.1121/10.0005754

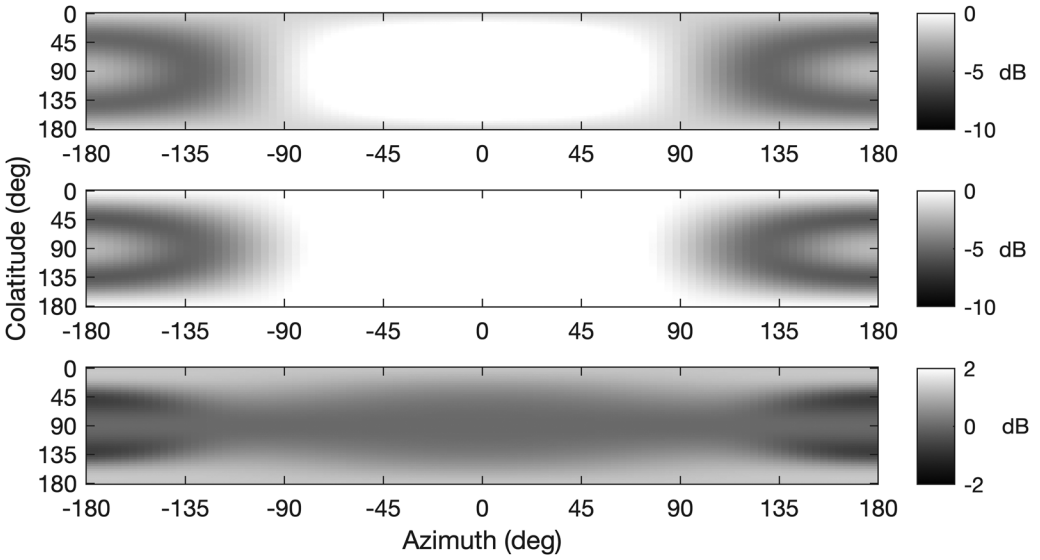

FIG. 4. Top: Cylindrical projection of the actual magnitude sound pressure distribution on a logarithmic scale on the surface of the rigid scattering object due to an impinging spherical wave of frequency $f = 1000$ Hz originating from an azimuth of 0° in the horizontal plane from a distance of 1 m. Middle: Reconstruction of the sound pressure distribution by an eighth-order EMA. Bottom: Difference between the two.

reconstructed by the EMA, and Fig. 4 (bottom) depicts the difference between the two.

Figure 3 (top) depicts the accuracy of the sound field reconstruction by the EMA for the relatively high order of $N = 8$. Figure 3 (bottom) depicts the same data but for fifth-order arrays. The evolution of the average relative error is qualitatively identical to the eighth-order array. We therefore conclude that qualitative properties of SMAs and EMAs with respect to spatial aliasing and order limitation are identical.

B. Binaural rendering

In the following, we evaluate the binaural signals when the sound field reconstruction from SMA and EMA is rendered according to (9). We use the term binaural signals to refer to the transfer function from the sound field through the microphone array and the binaural rendering. We depict only the data for the left ear for convenience. We use the HRTFs from Bernschütz (2013), which were obtained from a torso-less Neumann (Berlin, Germany) KU100 dummy head (DH) on a 2702-node Lebedev grid [the data can be downloaded from Bernschütz (2020)].

We employ soft clipping of the radial filters at 40 dB [cf. Eq. (20)], which corresponds to clipping the SMA radial filters at 18 dB with the normalization used in Ahrens and Andersson (2019), Bernschütz (2016), and Rafaely (2005). Audio examples for all analyzed scenarios are available at Ahrens *et al.* (2021a).

1. Accuracy of the binaural signals

Figure 5 compares the binaural signals from an EMA, an SMA, and the HRTFs of the DH for a horizontally propagating plane impinging from straight ahead of the DH. The DH HRTFs that we employ were measured at a distance of 3.25 m. This distance is sufficient to assume that the HRTFs do not change for farther source distances apart from the overall amplitude (Wierstorf *et al.*, 2011) so that we can interpret them as the free-field acoustic response of the DH's ears to a plane wave.

The EMA and SMA signals are virtually identical below approximately 4 kHz but depart slightly from the DH signals. This is most likely due to the SH order-limitation of the HRTFs that is inherent to Eq. (9). The limitation smooths the data so that certain details can get lost.

Above 4 kHz, both the SMA and the EMA depart from the DH ear signals, which is expected as the order-limited SH representation of the captured sound field is not accurate in this frequency range. The two arrays sample the sound field differently so that the manifestation of spatial aliasing also differs.

2. Non-horizontally propagating sound fields

Figure 6 depicts the binaural signals from the EMA and the SMA as well as the corresponding HRTFs for a plane wave impinging from straight ahead from different

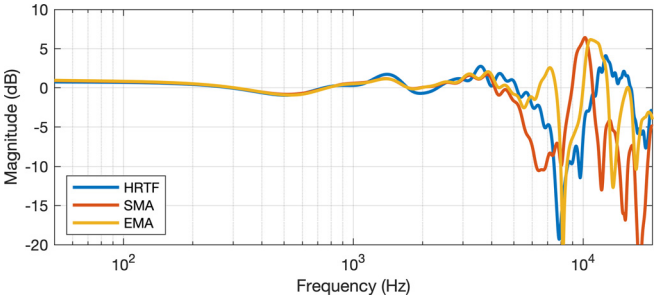

FIG. 5. (Color online) Binaural signals from an eighth-order SMA and an eighth-order EMA as well as the HRTF of the DH for a plane wave impinging from straight ahead.

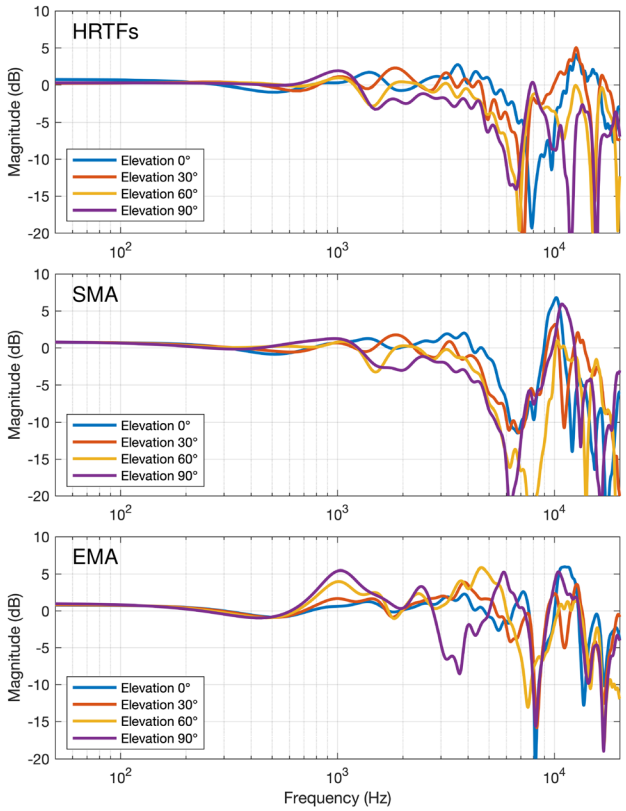

FIG. 6. (Color online) HRTFs (top) and binaural output signals of an eighth-order SMA (middle) and an eighth-order EMA (bottom) for plane wave incidence from straight ahead from different elevations.

elevations. It is clear that the EMA cannot preserve elevation-related localization cues as it necessarily produces a representation of a height-invariant sound field, yet the effect of the violation of the requirement for a height-invariant sound field on the binaural signals is moderate. An undesired amplification by a few dB occurs around 1 kHz for all non-zero elevations of sound incidence, and an undesired notch between 3 and 4 kHz occurs for sound incidence from 90° elevation (straight above). For sound impinging from straight above, all microphones of the EMA capture the exact same signal so that the output contains only the zeroth SH order (Thomas *et al.*, 2014; Weller *et al.*, 2011).

Most static HRTF elevation queues are apparent in the frequency range above 4 kHz (Blauert, 1996), where neither EMA nor SMA are accurate for the given maximum SH order N and radius R of the array. The deviations of the ear signals are caused by spatial aliasing and SH order truncation. Usually, the maximum order of the sound field representation that can be obtained from the microphone array is lower than the maximum order at which the HRTFs are available. This means that, if the binaural rendering is performed according to (9), the potentially available modes $(n,m)\ \forall\ n, |m| > N$ of the HRTFs are not used because $\breve{S}_n^m(\omega) = 0\ \forall\ n, |m| > N$. Using virtual loudspeakers instead of the modal rendering (9) may potentially mitigate this limitation so that more elevation cues are preserved (Thresh *et al.*, 2019).

3. Sources at short distances

We compare the EMA output to the SMA output in the following to evaluate the EMA's performance for sources at finite distances while keeping in mind that we cannot draw detailed conclusions for frequencies above approximately 4 kHz from this. The SMA signals may be considered exact below 4 kHz [cf. Fig. 3 (middle)]. This is true even with the present radial filter gain limitation applied. We tested different gain limitations that are even higher than the present one and did not find a noticeable effect on the binaural signals.

Figure 7 compares the SMA and EMA binaural signals. All data are normalized with respect to the source distance. A small amplification at very low frequencies is apparent in the SMA signals, particularly for the closest source distance of $r_s = 0.3$ m compared to the farther distances. This is expected and is a general phenomenon with the scattering off rigid bodies (Wierstorf *et al.*, 2011).

The EMA signals are almost identical to the corresponding SMA signals below 4 kHz. Some small deviations are apparent, particularly for the very close source at $r_s = 0.3$ m, for example, in that the low-frequency amplification is slightly higher for the EMA. This can be explained by the circumstance that spherical waves originating from finite distances violate the assumption of height invariance that the EMA solution is based on [recall Fig. 3 (middle)].

This demonstrates that the presented solution overcomes the limitations of our previous approach for the EMA from Ahrens *et al.* (2021c), which produced an excessive low-frequency boost for source distances shorter than 4 m for the array geometry under consideration.

4. Head rotations about the roll axis

So far, we have been assuming that the listener's ears are in the horizontal plane or close to it. The remaining

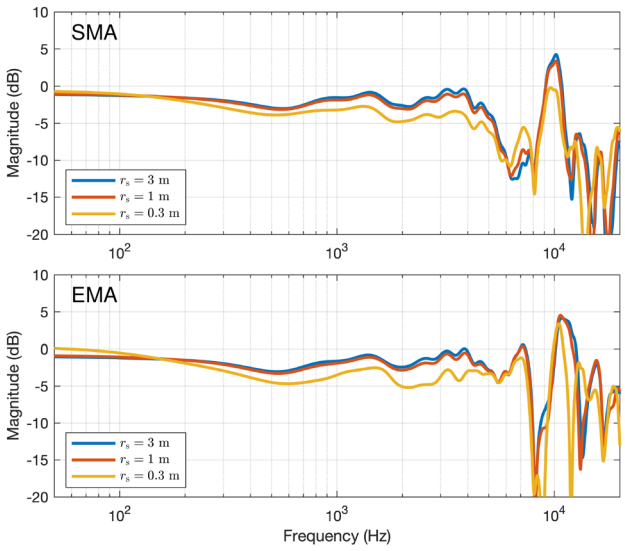

FIG. 7. (Color online) Binaural output signals for a listener oriented toward an incoming spherical wave originating from the horizontal plane from different distances r_s. Top: Eighth-order SMA; bottom: eighth-order EMA. All data are amplitude normalized.

https://doi.org/10.1121/10.0005754

aspect to investigate is how the binaural signals of the EMA are affected by tracked roll rotations of the head, i.e., rotations about the look direction, if sound sources are present at finite distances. Figure 3 showed that the accuracy of the EMA is lower off the equator so that the ear signals may be affected more strongly if the ears are not located on the equator.

Figure 8 depicts the SMA and EMA ear signals due to a spherical wave impinging from straight ahead and originating from the horizontal plane from different distances for different roll angles of the listener's head. Ideal binaural signals are not affected by the roll rotations because of the rotational symmetry of the sound field.

The same small alterations of the SMA signals with changing source distance like in Fig. 7 are apparent. The SMA ear signals are indeed independent of the roll angle below approximately 4 kHz. Spatial aliasing occurs at higher frequencies, which causes changes in the ear signals with changing roll angle. Also, the EMA ear signals are independent of the roll angle below approximately 4 kHz for plane waves (i.e., sources at infinite distance). Roll rotations cause small deviations on the order of 1–2 dB below 4 kHz for a source distance of $r_s = 1$ m. Very close sources at $r_s = 0.3$ m cause magnitude deviations on the order of several dB below 4 kHz. Note that the source position is at a distance of only approximately 0.2 m from the surface of the scatterer in this case.

The deviations above 4 kHz tend to be larger for the EMA than for the SMA due to the limitations to predicting the sound pressure on the surface of the EMA outside of the equator (cf. Fig. 3, middle).

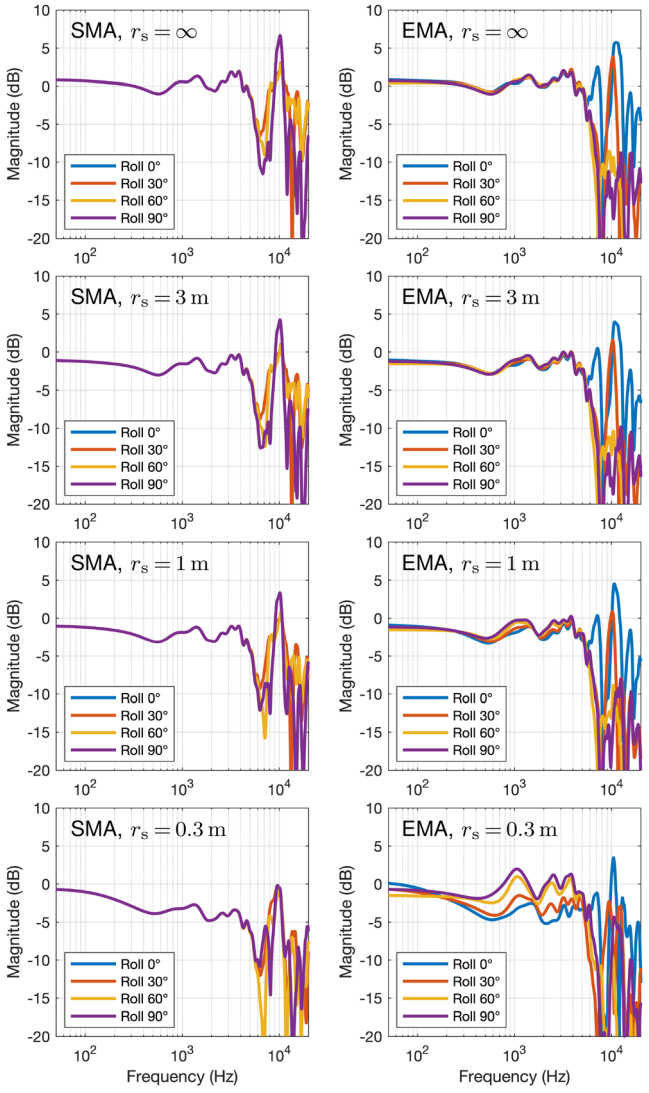

FIG. 8. (Color online) Binaural signals for eighth-order SMA (left column) and EMA (right column) due to an impinging spherical wave originating from the horizontal plane from straight ahead from different distances r_s. The listener's head was rotated in the roll direction by different angles γ_{roll}. All data are amplitude normalized.

IV. CONCLUSIONS

We presented a method for obtaining a SH representation of a sound field that is captured by a microphone array along the equator of a rigid spherical object. Our method reduces the required number of microphones by almost an order of magnitude compared to conventional spherical microphone arrays [$2N+1$ vs $(N+1)^2$ for SH order N]. The lower required number of microphones comes with a reduction of the computational complexity.

The solution requires that the captured sound field be height invariant. The consequences of violating this assumption are alterations of the magnitude spectrum if binaural rendering of the captured sound field is targeted. The alterations may be audible. Audio examples that accompany this article are available (Ahrens *et al.*, 2021a).

The primary theoretical limitation of the proposed equatorial array compared to conventional spherical microphone arrays with respect to binaural rendering is the fact that the equatorial array always outputs a height-invariant sound field that, consequently, propagates parallel to the horizontal plane. As to our awareness, it has not been proven that practical spherical microphone arrays are actually capable of preserving perceptually relevant elevation information in a modal binaural rendering. It is therefore unclear at this point if equatorial arrays actually exhibit a disadvantage in practice. This has to be clarified in a perceptual experiment.

ACKNOWLEDGMENTS

We thank Facebook Reality Labs for funding the presented work.

APPENDIX

We reformulate (16) by expanding the plane wave term $e^{-i\omega/c\, r\cos(\theta-\alpha)}$ into SHs using (14) as

$$S^{\text{int}}(\vec{x},\omega) = \frac{1}{2\pi}\int_0^{2\pi} \bar{S}(\theta,\omega)\sum_{n=0}^{\infty}\sum_{m=-n}^{n} 4\pi i^{-n}\, e^{-im\theta} \times Y_n^m\left(\frac{\pi}{2},0\right) j_n\left(\omega\frac{r}{c}\right) Y_n^m(\beta,\alpha)\, d\theta. \quad \text{(A1)}$$

Equation (A1) represents the sound field that impinges on the scatterer. Similarly as in (1) and (3), we replace the radial term $j_n(r,\omega)$ with $b_n(\omega r/c, R)$ to account for the effect of the scatterer as

$$S(\vec{x},\omega) = \frac{1}{2\pi}\int_0^{2\pi} \bar{S}(\theta,\omega) \sum_{n=0}^{\infty}\sum_{m=-n}^{n} 4\pi \mathrm{i}^{-n}\, \mathrm{e}^{-\mathrm{i}m\theta} \times Y_n^m\left(\frac{\pi}{2},0\right) b_n\left(\omega\frac{r}{c},R\right) Y_n^m(\beta,\alpha)\,\mathrm{d}\theta. \quad \text{(A2)}$$

Equation (A2) represents a 3D sound field that is produced by a 2D impinging sound field. We change the order of integration and summation to yield

$$S(\vec{x},\omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \underbrace{\frac{1}{2\pi}\int_0^{2\pi} \bar{S}(\theta,\omega)\, \mathrm{e}^{-\mathrm{i}m\theta}\,\mathrm{d}\theta}_{=\mathring{\bar{S}}_m(\omega)} \times 4\pi\mathrm{i}^{-n}\, Y_n^m(\pi/2,0)\; b_n\left(\omega\frac{r}{c},R\right) Y_n^m(\beta,\alpha), \quad \text{(A3)}$$

whereby we used (10) to identify $\mathring{\bar{S}}_m(\omega)$, the CH coefficients of the plane wave coefficients $\bar{S}(\theta,\omega)$. Evaluating (A3) along the equator of the scatterer and changing the order of summations yields

$$S(\vec{x},\omega) = \sum_{m=-\infty}^{\infty} \mathring{\bar{S}}_m(\omega) \times \sum_{n=|m|}^{\infty} 4\pi\mathrm{i}^{-n}\; b_n\left(\omega\frac{r}{c},R\right) \left[Y_n^m(\pi/2,0)\right]^2 \mathrm{e}^{\mathrm{i}m\alpha}, \quad \text{(A4)}$$

which brings out the similarity to (17).

Equation (A4) was derived in Koyama *et al.* (2016) [Eq. (A20)] by combining spherical and cylindrical wave functions in one formulation and solving an integral over the product of a cylindrical Bessel function and an associated Legendre function.

Abhayapala, T. D., and Ward, D. B. (**2002**). "Theory and design of high order sound field microphones using spherical microphone array," in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 13–17, Orlando, FL, Vol. 2, pp. 1949–1952.

Ahrens, J. (**2012**). *Analytic Methods of Sound Field Synthesis* (Springer, Heidelberg, Germany).

Ahrens, J., and Andersson, C. (**2019**). "Perceptual evaluation of headphone auralization of rooms captured with spherical microphone arrays with respect to spaciousness and timbre," J. Acoust. Soc. Am. **145**(4), 2783–2794.

Ahrens, J., Helmholz, H., Alon, D., and Amengual Gari, S. (**2021a**). "Audio examples accompanying the article 'Spherical harmonic decomposition of a sound field based on observations along the equator of a rigid spherical scatterer', JASA (2021)," https://zenodo.org/record/4805266/export/hx#.YPWDuehKhPY (Last viewed May 31, 2021).

Ahrens, J., Helmholz, H., Alon, D., and Amengual Gari, S. (**2021b**). "Audio examples accompanying the paper 'The far-field equatorial array for binaural rendering', IEEE ICASSP (2021)," https://zenodo.org/record/4807499#.YPWEOOhKhPY (Last viewed May 31, 2021).

Ahrens, J., Helmholz, H., Alon, D., and Amengual Gari, S. (**2021c**). "The far-field equatorial array for binaural rendering," in *Proceedings of ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, June 6–11, Toronto, Canada.

Ahrens, J., and Spors, S. (**2008**). "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," Acta Acust. united Acust. **94**(6), 988–999.

Ahrens, J., and Spors, S. (**2012**). "Wave field synthesis of a sound field described by spherical harmonics expansion coefficients," JASA **131**(3), 2190–2199.

Algazi, V. R., Duda, R. O., and Thompson, D. M. (**2004**). "Motion-tracked binaural sound," J. Audio Eng. Soc. **52**(11), 1142–1156.

Bernschütz, B. (**2013**). "A spherical far field HRIR/HRTF compilation of the Neumann KU100," in *Proceedings of the 40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (DAGA)*, March 18–21, Meran, Italy, pp. 592–595.

Bernschütz, B. (**2016**). "Microphone arrays and sound field decomposition for dynamic binaural recording," Ph.D. thesis, Technische Universität Berlin, Berlin, Germany.

Bernschütz, B. (**2020**). "Spherical far-field HRIR compilation of the Neumann KU100," https://zenodo.org/record/3928297##.X-8-zKpKhBw (Last viewed May 31, 2021).

Betlehem, T., and Poletti, M. (**2019**). "Measuring the spherical-harmonic representation of a sound field using a cylindrical array," in *Proceedings of ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 955–959.

Blauert, J. (**1996**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA).

Chen, H., Abhayapala, T., and Zhang, W. (**2015**). "Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis," J. Acoust. Soc. Am. **138**(5), 3081–3092.

Galindo, M. B., Coleman, P., and Jackson, P. J. (**2020**). "Microphone array geometries for horizontal spatial audio object capture with beamforming," J. Audio Eng. Soc. **68**(5), 324–337.

Gumerov, N., and Duraiswami, R. (**2005**). *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier, Amsterdam).

Helmholz, H., Alon, D. L., Gari, S. A., and Ahrens, J. (**2020**). "Instrumental evaluation of sensor self-noise in binaural rendering of spherical microphone array signals," in *Proceedings of e-Forum Acusticum*, December 7–11, Lyon, France.

Kaiser, F., Pomberger, H., and Zotter, F. (**2012**). "Investigations on cylindrical microphone arrays," in *Proceedings of Spatial Audio in Today's 3D World: 25th AES UK Conference*, March 25–27, York, UK.

Koyama, S., Furuya, K., Wakayama, K., Shimauchi, S., and Saruwatari, H. (**2016**). "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," J. Acoust. Soc. Am. **139**(3), 1024–1036.

Lebedev, V. I. (**1977**). "Spherical quadrature formulas exact to orders 25–29," Sib. Math. J. **18**(1), 99–107.

Li, Z., and Duraiswami, R. (**2006**). "Headphone-based reproduction of 3D auditory scenes captured by spherical/hemispherical microphone arrays," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 14–19, Toulouse, France.

Lübeck, T., Helmholz, H., Arend, J. M., Pörschmann, C., and Ahrens, J. (**2020**). "Perceptual evaluation of mitigation approaches of impairments due to spatial undersampling in binaural rendering of spherical microphone array data," J. Audio Eng. Soc. **68**(6), 428–440.

Meyer, J. (**2001**). "Beamforming for a circular microphone array mounted on spherically shaped objects," J. Acoust. Soc. Am. **109**, 185–193.

Meyer, J., and Elko, G. (**2002**). "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 13–17, Orlando, FL, pp. 1781–1784.

Parthy, A., Epain, N., van Schaik, A., and Jin, C. T. (**2011**). "Comparison of the measured and theoretical performance of a broadband circular microphone array," J. Acoust. Soc. Am. **130**(6), 3827–3837.

Rabenstein, R., Spors, S., and Ahrens, J. (**2014**). "Sound field synthesis," edited by in *Academic Press Library in Signal Processing*, Vol. 4, edited by S. Theodoridis and R. Chellappa (Academic, Chennai, India), Chap. 32, pp. 915–979.

Rafaely, B. (**2004**). "Plane-wave decomposition of the sound field on a sphere by spherical convolution," J. Acoust. Soc. Am. **116**, 2149–2157.

Rafaely, B. (**2005**). "Analysis and design of spherical microphone arrays," IEEE Trans. Speech Audio Proc. **13**(1), 135–143.
Rafaely, B., and Avni, A. (**2010**). "Interaural cross correlation in a sound field represented by spherical harmonics," J. Acoust. Soc. Am. **127**(2), 823–828.
Teutsch, H., and Kellermann, W. (**2006**). "Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays," J. Acoust. Soc. Am. **120**(5), 2724–2736.
Thomas, M. R. P., Ahrens, J., and Tashev, I. (**2014**). "A method for converting between cylindrical and spherical harmonic representations of sound fields," in *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 4–9, Florence, Italy, pp. 4723–4727.
Thresh, L., Armstrong, C., and Kearney, G. (**2019**). "A direct comparison of localisation performance when using first, third and fifth order ambisonics for real loudspeaker and virtual loudspeaker rendering," in *Proceedings of the 143rd Convention of the AES*, October 18–21, New York, Paper 9864.
Tiana-Roig, E., Jacobsen, F., and Fernandez-Grande, E. (**2011**). "Beamforming with a circular array of microphones mounted on a rigid sphere (L)," J. Acoust. Soc. Am. **130**(3), 1095–1098.
Trevino, J., Koyama, S., Sakamoto, S., and Suzuki, Y. (**2014**). "Mixed-order ambisonics encoding of cylindrical microphone array signals," Acoust. Sci. Technol. **35**(3), 174–177.
Ward, D. B., and Abhayapala, T. D. (**2001**). "Reproduction of a plane-wave sound field using an array of loudspeakers," IEEE Trans. Speech Audio Process. **9**(6), 697–707.
Weller, T., Favrot, S., and Buchholz, J. M. (**2011**). "Application of a circular 2D hard-sphere microphone array for higher-order ambisonics auralization," in *Proceedings of Forum Acusticum*, June 26–July 1, Aalborg, Denmark, pp. 2269–2274.
Wierstorf, H., Geier, M., and Spors, S. (**2011**). "A free database of head related impulse response measurements in the horizontal plane with multiple distances," in *Proceedings of the 130th Convention of the AES*, May 13–16, London, UK, e-Brief 6.
Williams, E. (**1999**). *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, New York).
Zaunschirm, M., Schörkhuber, C., and Höldrich, R. (**2018**). "Binaural rendering of ambisonic signals by head-related impulse response time alignment and a diffuseness constraint," J. Acoust. Soc. Am. **143**(6), 3616–3627.
Zaunschirm, M., and Zotter, F. (**2014**). "Measurement-based modal beamforming using planar circular microphone arrays," in *Proceedings of the EAA Joint Symposium on Auralization and Ambisonics*, April 3–5, Berlin, Germany, pp. 75–80.
Zotkin, D. N., Duraiswami, R., and Gumerov, N. A. (**2010**). "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," IEEE Trans. Audio Speech Lang. Process. **18**(1), 2–16.
Zotter, F. (**2009**). "Sampling strategies for acoustic holography/holophony on the sphere," *Proceedings of NAG/DAGA*, March 23–26, Rotterdam, Netherlands, pp. 1107–1110.