

# A Semi-Analytical Model for the Binaural Coherence of Noise Fields

Marco Jeub, *Student Member, IEEE*, Matthias Dörbecker, and Peter Vary, *Fellow, IEEE*

**Abstract**—A novel semi-analytical signal processing model for the binaural coherence of homogeneous isotropic noise fields is presented in this contribution. This is derived from a simplified geometrical model of the human head, where the shadowing between the left and right ear is modeled by two nonreflecting circular plates. Based on Kirchhoff's diffraction theory, it is shown how the corresponding coherence is calculated. This model can be used as part of various binaural signal processing algorithms, such as speech enhancement for digital hearing aids or binaural speech transmission systems. In experiments using an artificial head in a highly reverberant environment, it is confirmed that the proposed theoretical model shows a good match with the coherence obtained from measurements.

**Index Terms**—Binaural, head shadowing, head-related, Kirchhoff diffraction theory, noise field coherence.

## I. INTRODUCTION

THE noise field coherence is an essential acoustical parameter of multichannel speech enhancement algorithms. Therefore, it is required to have suitable signal processing models which accurately describe the acoustical environment. For homogeneous isotropic noise fields, well-known analytical models exist. In contrast to that, the influence of head shadowing on the coherence is usually modeled heuristically. This applies for binaural hearing aids and binaural speech transmission systems. Early studies in [1] showed that the influence of the head has a severe impact on the noise field coherence and proposed a modified coherence model which is basically a curve fitting procedure derived from measurements with an artificial head. The authors in [2] present a model for binaural sound synthesis which is based on the binaural cues. However, this model allows to reproduce binaural sound data with correct interaural time and level difference cues, but gives no information about the binaural coherence. In [3], the distance parameter of the free-field coherence model is simply scaled in order to take the modified coherence into account. This article is based on [4] and presents a semi-analytical model for the binaural noise field coherence. The derivation employs Kirchhoff's diffraction theory and Babinet's principle, cf. [5]. The main advantage compared to the previous models is that

Manuscript received November 18, 2010; revised January 12, 2011; accepted January 19, 2011. Date of publication January 24, 2011; date of current version January 31, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Constantine L. Kotropoulos.

M. Jeub and P. Vary are with the Institute of Communication Systems and Data Processing, RWTH Aachen University, Aachen 52074, Germany (e-mail: jeub@ind.rwth-aachen.de; vary@ind.rwth-aachen.de).

M. Dörbecker is with the Harman/Becker Automotive Systems GmbH, Karlsbad, Germany (e-mail: matthias.doerbecker@harman.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2011.2108284

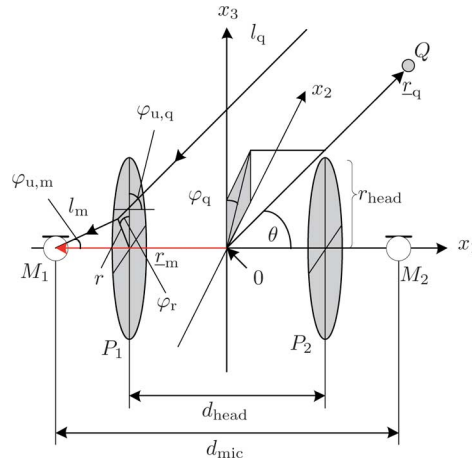


Fig. 1. Simplified geometrical model of the human head with two circular plates  $P_{1|2}$ . Microphones are denoted by  $M_{1|2}$ .

arbitrary dimensions for head and microphone distances can be employed and no acoustic measurements as in [1], [3] are required. A MATLAB reference implementation is available online<sup>1</sup>.

## II. FREE-FIELD COHERENCE MODEL

The complex coherence between the signals  $x_{1|2}(k)$  with discrete time index  $k$  is defined in the frequency domain as

$$\Gamma_{x_1 x_2}(\Omega) = \frac{\Phi_{x_1 x_2}(e^{j\Omega})}{\sqrt{\Phi_{x_1 x_1}(e^{j\Omega}) \cdot \Phi_{x_2 x_2}(e^{j\Omega})}} \quad (1)$$

where  $\Phi_{x_1 x_1|x_2 x_2}(e^{j\Omega})$  represent the auto-power spectral densities and  $\Phi_{x_1 x_2}(e^{j\Omega})$  the cross-power spectral density. The frequency variable  $f$  and sampling rate  $f_s$  are normalized to the normalized radian frequency  $\Omega = 2\pi f/f_s$ . The frequently used term magnitude squared coherence (MSC) is referred to the squared magnitude of (1). The coherence for spherically isotropic (3-D, diffuse) and cylindrically isotropic (2-D) fields read for the far-field [6]

$$\Gamma_{x_1 x_2}^{(\text{diff})}(\Omega) = \text{sinc}\left(\frac{\Omega f_s d_{\text{mic}}}{c}\right), \quad (2)$$

$$\Gamma_{x_1 x_2}^{(\text{cyl})}(\Omega) = J_0\left(\frac{\Omega f_s d_{\text{mic}}}{c}\right) \quad (3)$$

with distance  $d_{\text{mic}}$  between two omnidirectional microphones,  $c$  the speed of sound and the zero-order Bessel function of first kind  $J_0(\cdot)$ .

## III. GEOMETRIC HEAD DIFFRACTION MODEL

A simplified geometric model for the complex head geometry according to Fig. 1 is used in the following. It is assumed

<sup>1</sup><http://www.ind.rwth-aachen.de/~bib/jeub11>

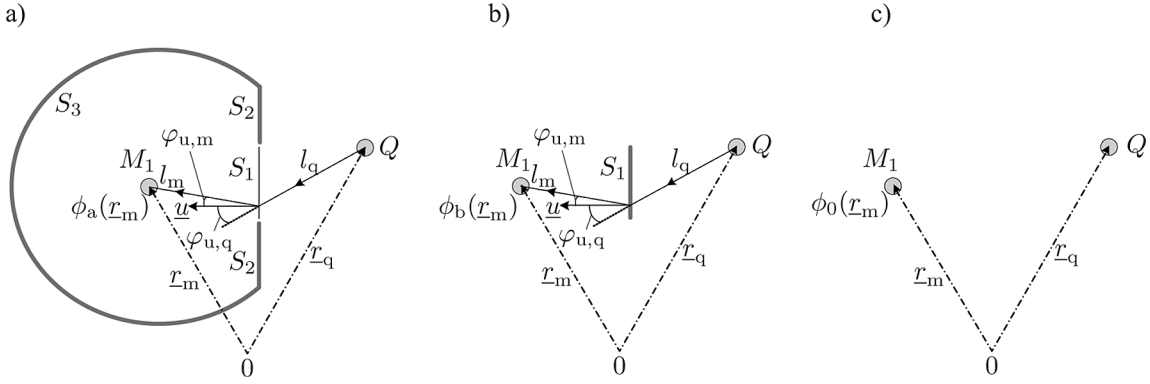


Fig. 2. Illustration of Kirchhoff's diffraction theory and Babinet's principle. (a) Pinhole arrangement with potential  $\phi_a(\underline{r}_m)$  at microphone  $M_1$ , (b) inverse arrangement with obstacle and corresponding potential  $\phi_b(\underline{r}_m)$ , (c) free-field arrangement with potential  $\phi_0(\underline{r}_m)$  (6) as the sum of  $\phi_a(\underline{r}_m)$  and  $\phi_b(\underline{r}_m)$ . The surface  $S_1$  corresponds to plate  $P_1$  in Fig. 1.

that the two microphones  $M_1$  and  $M_2$  are placed at distance  $d_{\text{mic}}$  next to the pinna of each side. The resulting shadowing is modeled by two nonreflecting circular plates ( $P_1$  and  $P_2$ ) with radius  $r_{\text{head}}$  and distance  $d_{\text{head}}$ . We further define a punctual sound source  $Q$  by its position vector  $\underline{r}_q$  or angles  $\theta$  (azimuth) and  $\varphi_q$  (elevation) in the far-field and  $H_1$  and  $H_2$  the transfer functions between  $Q$  and the two microphones  $M_1$  and  $M_2$ . The position vector of  $M$  is denoted by  $\underline{r}_m$ . The point of origin are marked with 0 in the figures. Besides, any point on the plate is given either by position vector  $\underline{r}$  or  $r = \|\underline{r}\|$  and  $\varphi_r$ . All further variables will be introduced below successively. Assuming two omnidirectional microphones, the spherically isotropic coherence can be calculated for a homogeneous noise field by the integration over all possible directions of incident of a directional sound source. Since this procedure requires knowledge about the transfer functions  $H_{1|2}$ , a derivation based on optical principles will be given in the following.

#### IV. KIRCHHOFF'S DIFFRACTION THEORY

Kirchhoff's diffraction theory was initially developed to explain optical phenomena in terms of diffraction. However, this general theory can also be applied to sound waves. Consider a monochromatic wave which propagates from a punctual source  $Q$  through the opening of a screen (pinhole)  $S_1$  as depicted in Fig. 2(a), where  $M_1$  is the point at which the disturbance is to be determined. The distances between opening and  $Q$  and opening and  $M_1$  are denoted by  $l_q$  and  $l_m$  respectively. In the following we assume that the opening is small compared to the distance of both  $Q$  and  $M_1$  from the obstacle and show how the corresponding transfer function between  $Q$  and  $M_1$ , i.e.,  $H_{qm}(f)$  is derived. Based on the Kirchhoff diffraction theory, the potential at location  $M_1$  can be expressed by the *integral theorem of Helmholtz and Kirchhoff* as

$$\phi(\underline{r}_m) = \frac{1}{4\pi} \int_S \left[ \phi(\underline{r}) \frac{\partial}{\partial u} \left( \frac{e^{-j\beta l_m}}{l_m} \right) - \frac{e^{-j\beta l_m}}{l_m} \frac{\partial \phi(\underline{r})}{\partial u} \right] dS \quad (4)$$

where  $\int_S \dots dS$  gives the integration over a nonreflecting surface  $S$  which encloses the point  $M_1$ ,  $\partial/\partial u$  denotes differentiation along the inward normal  $\underline{u}$  to the surface of integration and  $l_m = \|\underline{l}_m\|$  is the distance of the element  $dS$  from  $M_1$ . The wave number is denoted by  $\beta = 2\pi f/c$ . The surface  $S$  is formed by three partial surfaces which together form a closed surface, the opening  $S_1$ , a portion  $S_2$  at the backside of the ob-

stacle and a large sphere, centered at  $M_1$ ,  $S_3$ . This decomposition allows for some important simplifications as shown later. So far the wave propagation between  $Q$  and  $M_1$  was disturbed by a pinhole. Now in order to calculate the potential and hence, the transfer function with an obstacle between  $Q$  and  $M_1$ , (4) could be applied for the case depicted in Fig. 2(b). However, since this is inherently difficult, the *Babinet principle* is applied in the following, cf. [5]. This theorem gives the relation between the free-field potential  $\phi_0(\underline{r}_m)$  [Fig. 2(c)] and a superposition of the potential  $\phi_a(\underline{r}_m)$  for the pinhole arrangement [Fig. 2(a)] with the potential  $\phi_b(\underline{r}_m)$  of the complementary arrangement [Fig. 2(b)] according to

$$\phi_0(\underline{r}_m) = \phi_a(\underline{r}_m) + \phi_b(\underline{r}_m). \quad (5)$$

Therefore, the potential  $\phi_a(\underline{r}_m)$  is calculated first, followed by the use of (5) to obtain  $\phi_b(\underline{r}_m)$ . The corresponding free-field potential (without obstacle) at the microphone reads [5]:

$$\phi_0(\underline{r}_m) = C \frac{e^{-j\beta \|\underline{r}_q - \underline{r}_m\|}}{\|\underline{r}_q - \underline{r}_m\|} \quad (6)$$

with constant  $C$ . A further difficulty is encountered that the values of  $\phi_a(\underline{r}_m)$  and  $\partial/\partial u$  on the partial surfaces  $S_1$ ,  $S_2$  and  $S_3$  are never known exactly. Therefore, the following approximations are made which are referred to as the *Kirchhoff boundary conditions*. For  $S_1$  it is assumed that the rim of the opening can be neglected and hence, that the potential will not considerably differ from the values obtained in the absence of the plate (free-field). Hence, it can be written

$$\phi_a^{(S_1)}(\underline{r}) = \phi_0(\underline{r}) = C \frac{e^{-j\beta l_q}}{l_q}, \quad (7)$$

$$\frac{\partial \phi_a^{(S_1)}(\underline{r})}{\partial u} = C \frac{e^{-j\beta l_q}}{l_q} \left( -j\beta - \frac{1}{l_q} \right) \cos \varphi_{u,q}. \quad (8)$$

Furthermore, the potential and hence, the derivative vanish on  $S_2$ , i.e.,  $\phi_a^{(S_2)}(\underline{r}) = 0$  and  $\partial \phi_a^{(S_2)}(\underline{r})/(\partial u) = 0$ . Additionally, the integral over  $S_3$  will vanish by letting the radius increase indefinitely (see [5] for details). With such simplifications and

$$\frac{\partial}{\partial u} \left( \frac{e^{-j\beta l_m}}{l_m} \right) = \frac{e^{-j\beta l_m}}{l_m} \left( -j\beta - \frac{1}{l_m} \right) \cdot (-\cos \varphi_{u,m}), \quad (9)$$

the potential for the pinhole arrangement [Fig. 2(a)] can be given with (4) as

$$\phi_a(\underline{r}_m) = \frac{jfC}{2c} \int_{S_1} \frac{e^{-j\beta(l_q+l_m)}}{l_q l_m} \left[ \left(1 - \frac{jc}{2\pi f l_q}\right) \cos \varphi_{u,m} + \left(1 - \frac{jc}{2\pi f l_q}\right) \cos \varphi_{u,q} \right] dS \quad (10)$$

which is known as the *Fresnel-Kirchhoff diffraction formula*. The desired transfer function between  $Q$  and  $M_1$  arises from the potential by means of a normalization as

$$H_{qm}(\Omega) = l_0 \cdot \phi_b(\underline{r}_m) \quad (11)$$

where  $l_0$  denotes a scaling factor such that  $\|H_{qm}\| = 1$  holds in the free-field. Finally, by means of (5), (6), (10), this transfer function reads

$$\begin{aligned} H_{qm}(\Omega) &= \frac{e^{-j(\Omega f_s/c)\|\underline{r}_q - \underline{r}_m\|}}{\|\underline{r}_q - \underline{r}_m\|} l_0 \\ &\quad - \frac{j\Omega f_s l_0}{4\pi c} \int_{S_1} \frac{e^{-j(\Omega f_s/c)(l_q+l_m)}}{l_q l_m} \\ &\quad \times \left[ \left(1 - \frac{jc}{\Omega f_s l_q}\right) \cos \varphi_{u,q} + \left(1 - \frac{jc}{\Omega f_s l_m}\right) \cos \varphi_{u,m} \right] dS. \end{aligned} \quad (12)$$

The geometric interpretation of  $l_q$  and  $l_m$  will be given in Section V.

## V. BINAURAL COHERENCE MODEL

For each microphone, the transfer function is calculated and, due to the symmetry, the diffraction at the corresponding nearest plate is taken into account. First, it is considered that the angle  $\theta$  lies in the range  $0 \leq \theta < (\pi/2)$  (see Fig. 1). According to (12), the transfer function between  $Q$  and  $M_1$  is given by

$$\begin{aligned} H_1(\Omega) &= \frac{e^{-j(\Omega f_s/c)\|\underline{r}_q - \underline{r}_m\|}}{\|\underline{r}_q - \underline{r}_m\|} l_0 \\ &\quad - \frac{j\Omega f_s l_0}{4\pi c} \int_{P_1} \frac{e^{-j(\Omega f_s/c)(l_q+l_m)}}{l_q l_m} \\ &\quad \cdot \left[ \left(1 - \frac{jc}{\Omega f_s l_q}\right) \cos \varphi_{u,q} + \left(1 - \frac{jc}{\Omega f_s l_m}\right) \cos \varphi_{u,m} \right] dP. \end{aligned} \quad (13)$$

As previously  $\int_{P_1} \dots dP$  indicates the integration over a surface, here, of plate  $P_1$ .  $\underline{r}_q$  and  $\underline{r}_m$  are the position vectors, such that the distance between source and microphone is given by  $\|\underline{r}_q - \underline{r}_m\|$ . According to Fig. 1, it can be written

$$\|\underline{r}_q - \underline{r}_m\| = r_q + \frac{d_{mic}}{2} \cos \theta \quad (14)$$

where  $r_q = \|\underline{r}_q\|$  is the distance from  $Q$  to the point of origin. This equation holds since the distance from the source is assumed large compared to the distance of the microphones, i.e.,  $r_q \gg d_{mic}$ . The distance  $l_q$  in (12) and (13), which corresponds

to the distance of the sound wave from the source to the point specified by  $(r, \phi_r)$  of the plate  $P_1$ , reads

$$l_q = r_q + \frac{d_{head}}{2} \cos \theta - r \sin \theta \cos(\varphi_r - \varphi_q). \quad (15)$$

The distance between this specific point on the plate and microphone  $M_1$  can be expressed by

$$l_m = \sqrt{r^2 + \left(\frac{d_{mic} - d_{head}}{2}\right)^2}. \quad (16)$$

Additionally, the incident and emergent angles in (13) can be written according to Fig. 1 due to the far-field assumption as

$$\begin{aligned} \cos \varphi_{u,q} &= \cos \theta \quad \text{and} \\ \cos \varphi_{u,m} &= \frac{d_{mic} - d_{head}}{2l_m}. \end{aligned} \quad (17)$$

Besides, the distances  $l_q$  and  $\|\underline{r}_q - \underline{r}_m\|$  can be replaced by  $r_q$ , unless for such arguments in the exponential function. Taking further into account that  $1 - (jc/\Omega f_s l_q) \approx 1$ , the transfer function can be expressed by means of polar coordinates with  $dP = r \cdot d\varphi_r \cdot dr$  and after rearranging as

$$\begin{aligned} H_1(\Omega, \theta) &= \frac{e^{-j(\Omega f_s/c)r_q}}{\frac{r_q}{l_0}} \\ &\quad \cdot \left[ e^{-j(\Omega f_s/c)(d_{mic}/2) \cos \theta} \right. \\ &\quad \left. - \frac{j\Omega f_s}{4\pi c} \int_0^{r_{head}} \int_0^{2\pi} e^{-j(\Omega f_s/c)((d_{head}/2) \cos \theta - r \sin \theta \cos \varphi + l_m)} \right. \\ &\quad \left. \cdot \left[ \cos \theta + \left(1 - \frac{jc}{\Omega f_s l_m}\right) \frac{d_{mic} - d_{head}}{2l_m} \right] \frac{r}{l_m} d\varphi dr \right] \end{aligned} \quad (18)$$

where  $l_m$  is dependent on  $r$  as in (16). The derivation of (12) was performed by means of Babinet's principle where the diffraction at the obstacle was replaced by a pinhole. There it was assumed that the potential field inside the pinhole [surface  $S_1$  in Fig. 2(a)] is equal to the free-field potential, i.e.,  $\phi_0(\underline{r}_m)$ . However, this assumption does not hold if the source is located at the same side as the obstacle (or opening). This can be explained since in this case, in the inverse arrangement the surfaces  $S_2$  and  $S_3$  would block the line-of-sight between source and the opening. Hence, in our geometrical approximation of the head as two plates, (12) is valid only for  $0 \leq \theta < (\pi/2)$ , where for the special case  $\theta \approx (\pi/2)$  the equation is also only an approximation. Therefore, (12) cannot be used to calculate the transfer function between source and  $M_2$  for  $0 \leq \theta < (\pi/2)$ . Since in this case  $Q$  and  $M_2$  are located at the same side of the plate, no diffraction occurs. Hence, the frequency response corresponds to the free-field condition for  $0 \leq \theta < (\pi/2)$ , obtained from (6), (14) and the same normalization as in (12) by

$$H_2(\Omega, \theta) = \frac{e^{-j(\Omega f_s/c)r_q}}{\frac{r_q}{l_0}} \cdot e^{j(\Omega f_s/c)(d_{mic}/2) \cos \theta}. \quad (19)$$

For the case  $(\pi/2) < \theta \leq \pi$ , the opposite effect occurs, i.e.,  $M_2$  is blocked by  $P_2$  while a free-field condition can be assumed between source  $Q$  and  $M_1$ . Due to symmetry, the transfer function can be hence formulated as

$$\begin{aligned}
H_1(\Omega, \theta) &= H_2(\Omega, \pi - \theta) \\
&\text{for } \frac{\pi}{2} < \theta \leq \pi \\
H_2(\Omega, \theta) &= H_1(\Omega, \pi - \theta).
\end{aligned} \quad (20)$$

The coherence definition of (1) can, under the assumption of two omnidirectional microphones, generally be expressed as

$$\begin{aligned}
\Gamma_{x_1 x_2}(\Omega) &= \frac{\int_0^{\pi/2} \left( H_1(\Omega, \theta) H_2^*(\Omega, \theta) + H_2(\Omega, \theta) H_1^*(\Omega, \theta) \right) \sin \theta d\theta}{\int_0^{\pi/2} \left( |H_1(\Omega, \theta)|^2 + |H_2(\Omega, \theta)|^2 \right) \sin \theta d\theta}.
\end{aligned} \quad (21)$$

Since  $H_1$  and  $H_2$  are independent of  $\varphi_q$ , the double integral simplifies to the integration over  $\theta$ . With (20), (21) reads

$$\begin{aligned}
&= \frac{2 \int_0^{\pi/2} \text{Re} \{ H_1(\Omega, \theta) H_2^*(\Omega, \theta) \} \sin \theta d\theta}{\int_0^{\pi/2} \left( |H_1(\Omega, \theta)|^2 + |H_2(\Omega, \theta)|^2 \right) \sin \theta d\theta}.
\end{aligned} \quad (22)$$

Finally, with (22), (18), and (19), the coherence can be calculated. It can be observed that the coherence is independent of scaling factor  $l_0$  and distance  $r_q$  since the prefactor  $1/(r_q/l_0)e^{-j(\Omega f_s/c)r_q}$  eliminates. Since a closed-form solution of the integral in (18) cannot be obtained, the coherence given by (22) has to be solved numerically. One solution is to calculate the integral by summation. The intervals  $\Delta r$  and  $\Delta \tilde{\varphi}$  in (18) have to be chosen such that the corresponding surface elements  $\Delta S = r \cdot \Delta \tilde{\varphi} \cdot \Delta r$  are small compared to the sound wavelength  $\lambda$ . Here, it is proposed that the maximum length of every surface element should be one-tenth of the wavelength. Details are given within the reference implementation. In order to calculate the binaural coherence for a 2-D noise field, where the noise sources are distributed in the same horizontal plane as the head, the  $\sin \theta$ -terms in (22) have to be disregarded.

## VI. EXPERIMENTS

For the verification of the proposed coherence model, room impulse responses of a reverberant noise field ( $T_{60} = 0.72$  s) are measured with an artificial head ( $d_{\text{head}} = 0.15$  m,  $r_{\text{head}} = 0.075$  m). The two microphones are positioned close to the pinna at 1 cm from the ear canal, i.e.,  $d_{\text{mic}} = 0.17$  m. The measurements were repeated in an otherwise unchanged experimental setup after the head was removed to examine the influence of the head. In order to evaluate the noise field coherence of the late reverberant part only, we removed the coherent direct and early parts from the impulse response. Fig. 3 shows the corresponding curves for two microphones at a distance of  $d_{\text{mic}} = 0.17$  m (gray). The theoretical curves represent the ideal 3-D and 2-D noise field without head (free-field) by (2) and (3) respectively (red). The corresponding 3-D and 2-D curves with head shadowing (22) (with and without the  $\sin \theta$ -terms) are plotted in black. Regarding the figure, we conclude that the proposed coherence model greatly approximates the measured

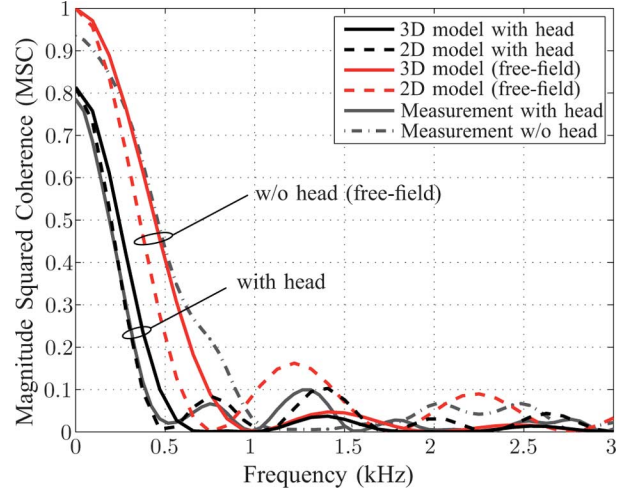


Fig. 3. MSC of ideal diffuse noise field and shadowing influence. Plotted are the theoretical curves and results from measurements in a reverberant environment ( $T_{60} = 0.72$  s).

coherence. Especially for hearing-aid algorithms, assuming a cocktail-party environment, the proposed 2-D model is the most appropriate one.

## VII. CONCLUSIONS

In this letter, a novel binaural coherence model is proposed which takes shadowing effect of the head into account. This was derived by using Kirchhoff's diffraction theory and a simplified geometrical model of the head. Experiments have shown that the new model is a good approximation with measurements obtained in a reverberant environment. This coherence model can be used to investigate the influence of head-shadowing on coherence-based speech enhancement algorithms, in binaural noise reduction or dereverberation algorithms where the binaural coherence is exploited explicitly, e.g., [3], [7]; and to generate realistic binaural noise fields for simulations, e.g., using the approach of [8].

## REFERENCES

- [1] I. Lindevald and A. Benade, "Two-ear correlation in the statistical sound fields of rooms," *J. Acoust. Soc. Amer.*, vol. 80, no. 2, pp. 661–664, 1986.
- [2] C. P. Brown and R. O. Duda, "A structural model for binaural sound synthesis," *IEEE Trans. Speech Audio Process.*, vol. 6, pp. 476–488, 1998.
- [3] A. H. Kamkar-Parsi and M. Bouchard, "Improved noise power spectrum density estimation for binaural hearing aids operating in a diffuse noise field environment," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 4, pp. 521–533, 2009.
- [4] M. Dörbecker, "Mehrkanalige Signalverarbeitung Zur Verbesserung Akustisch Gestörter Sprachsignale am Beispiel Elektronischer Hörhilfen," Ph.D., RWTH Aachen Univ., Aachen, Germany, 1998.
- [5] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [6] H. Kuttruff, *Room Acoustics*, 5th ed. Oxon, U.K.: Spon, 2009.
- [7] M. Jeub, M. Schäfer, T. Esch, and P. Vary, "Model-based dereverberation preserving binaural cues," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 7, pp. 1732–1745, 2010.
- [8] E. Habets, I. Cohen, and S. Gannot, "Generating nonstationary multi-sensor signals under a spatial coherence constraint," *J. Acoust. Soc. Amer.*, vol. 124, no. 5, pp. 2911–2917, 2008.