BINAURAL COHERENT-TO-DIFFUSE-RATIO ESTIMATION FOR DEREVERBERATION USING AN ITD MODEL

Chengshi Zheng[*†], *Andreas Schwarz*[†], *Walter Kellermann*[†], *Xiaodong Li*[*]

[*] Communication Acoustics Laboratory
Institute of Acoustics, CAS
100190 Beijing, China

[†] Chair of Multimedia Communications and Signal Processing
Friedrich-Alexander-Universität Erlangen-Nürnberg
91058 Erlangen, Germany

ABSTRACT

Most previously proposed dual-channel coherent-to-diffuse-ratio (CDR) estimators are based on a free-field model. When used for binaural signals, e.g., for dereverberation in binaural hearing aids, their performance may degrade due to the influence of the head, even when the direction-of-arrival of the desired speaker is exactly known. In this paper, the head shadowing effect is taken into account for CDR estimation by using a simplified model for the frequency-dependent interaural time difference and a model for the binaural coherence of the diffuse noise field. Evaluation of CDR-based dereverberation with measured binaural impulse responses indicates that the proposed binaural CDR estimators can improve PESQ scores.

Index Terms— Binaural speech dereverberation, interaural time difference, coherent-to-diffuse-ratio

1. INTRODUCTION

Both speech quality and speech intelligibility may dramatically degrade in reverberant and noisy environments. Many different algorithms were proposed to suppress noise and the reverberation during the past decades (see [1–3] and references therein). This paper focuses on binaural speech dereverberation, where the binaural signals are recorded with two microphones located at two human ears.

Previous studies have already shown that it is important to preserve both the interaural time difference (ITD) and the interaural level difference (ILD) cues when applying binaural dereverberation methods for hearing aids [4–9], since, when binaural cues are distorted, localization of sound sources becomes difficult [10]. This condition is ensured by a two-channel postfiltering approach where the same gain is applied to both channels [6]. In [6], Jeub *et al.* took the shadowing effect of the head into account in the diffuse sound field model. In [8, 9], interaural coherence histograms were mapped to a gain function to suppress the reverberant components in each frequency channel.

Recently, coherent-to-diffuse-ratio (CDR) estimators have been proposed, which can be seen as an alternative formulation of coherence-based dereverberation approaches [12]. In [6], the assumption was made that binaural signals are time-aligned before calculating the spectral weights of the Wiener filter. In [11], two CDR estimators were proposed, where one requires knowledge on both the direction of arrival (DOA) of the desired speaker and the spatial coherence of the late reverberant speech, and the other does not need the DOA information. In [12], Schwarz and Kellermann proposed improved estimators both for the case of known and unknown DOA, which were shown to lead to improved dereverberation performance (see [12, Table III] for details). To the best of our knowledge, these CDR estimators have not been applied to binaural dereverberation and their performance has not been reported until now.

After briefly reviewing CDR estimators for free-field conditions, i.e., for a sound field with no obstructions close to the microphones, we describe models for the ITD and the coherence of diffuse noise under the influence of the head in a binaural scenario, and show that the direction-dependent CDR estimators based on a free-field assumption are not robust under this model. We propose to modify the CDR estimators to use binaural models. Experimental results confirm that the proposed estimators achieve higher PESQ scores than the free-field estimators when applied to coherence-based dereverberation. The proposed binaural CDR estimators have numerous applications, such as binaural hearing aids, robotics, or immersive audio communication systems.

2. FREE-FIELD SIGNAL MODEL AND CDR ESTIMATION

We model two reverberant and noisy microphone signals $x_i(t)$, $i = 1, 2$, as the sum of a desired speech component $x_{i,\mathrm{coh}}(t)$ and an undesired component $x_{i,\mathrm{diff}}(t)$ consisting of diffuse reverberation and/or noise:

$$x_i(t) = x_{i,\mathrm{coh}}(t) + x_{i,\mathrm{diff}}(t), \tag{1}$$

As in previous studies, we assume both microphones to be omnidirectional and the desired component to be a plane wave

This work was supported by the National Science Fund of China (NSFC) under Grants 61201403 and 61302126.

in the free (locally unobstructed) field, so that $x_{2,\mathrm{coh}}(t)$ is a time-shifted version of $x_{1,\mathrm{coh}}(t)$ [6,11,12]:

$$x_{2,\mathrm{coh}}(t) = x_{1,\mathrm{coh}}(t - \tau_{12}), \tag{2}$$

where τ_{12} is the time difference of arrival (TDOA) of the desired sound between the first and the second microphone. The free-field model for the spatial coherence between the desired speech component at both microphones, $x_{1,\mathrm{coh}}(t)$ and $x_{2,\mathrm{coh}}(t)$, is given by

$$\Gamma_{\mathrm{coh}}^{\mathrm{FF}}(f) = \exp(j\tau_{12}). \tag{3}$$

If $\theta = 0°$ corresponds to broadside direction, the TDOA in the free field can be expressed as

$$\tau_{12} = d \sin\theta / c, \tag{4}$$

where d is the distance of the two microphones and c is the speed of sound.

The spatial coherence between the reverberation/noise components $x_{1,\mathrm{diff}}(t)$ and $x_{2,\mathrm{diff}}(t)$ is given by the spatial coherence function of two omnidirectional sensors in a diffuse (spherically isotropic), locally unobstructed sound field:

$$\Gamma_{\mathrm{diff}}^{\mathrm{FF}}(f) = \frac{\sin(2\pi f d/c)}{2\pi f d/c}, \tag{5}$$

where f is the frequency in Hz. For the cylindrically isotropic field, the spatial coherence can be given by

$$\Gamma_{\text{2D-iso}}^{\mathrm{FF}}(f) = J_0(2\pi f d/c). \tag{6}$$

Generally, (5) often fits better than (6) in practical applications [12], therefore we use $\Gamma_{\mathrm{diff}}^{\mathrm{FF}}(f)$ in this paper, although $\Gamma_{\text{2D-iso}}^{\mathrm{FF}}(f)$ may be applied analogously.

The CDR at the i-th microphone can be given by

$$CDR_i(k,f) = \frac{\Phi_{i,\mathrm{coh}}(k,f)}{\Phi_{i,\mathrm{diff}}(k,f)}, \tag{7}$$

where $\Phi_{i,\mathrm{coh}}(k,f)$ and $\Phi_{i,\mathrm{diff}}(k,f)$ are the short-time power spectra of $x_{i,\mathrm{coh}}(t)$ and $x_{i,\mathrm{diff}}(t)$, respectively, with the frame index k and frequency f (we will omit both k and f in the following for brevity). We further assume that the power spectra are identical at the two microphones for both the desired and undesired component, i.e., $\Phi_{\mathrm{coh}} = \Phi_{1,\mathrm{coh}} = \Phi_{2,\mathrm{coh}}$ and $\Phi_{\mathrm{diff}} = \Phi_{1,\mathrm{diff}} = \Phi_{2,\mathrm{diff}}$, and therefore

$$CDR = CDR_1 = CDR_2 = \frac{\Phi_{\mathrm{coh}}}{\Phi_{\mathrm{diff}}}. \tag{8}$$

Using the models for the coherence of the desired and diffuse signal components given above, and a short-time estimate of the coherence between $x_1(t)$ and $x_2(t)$, which is in the following denoted as $\hat{\Gamma}_x(k,f)$ and which may be obtained by recursive averaging, it is possible to estimate the time- and frequency-dependent CDR, as described in detail in [12]. The CDR estimators which are evaluated in this paper are summarized in Table 1.

Table 1: Summary of CDR estimators evaluated in this paper. Γ_{coh} and Γ_{diff} indicate the model coherence functions used for desired signal and diffuse noise, respectively, $\hat{\Gamma}_x$ indicates the estimated coherence of the mixed sound field. $\Re\{\bullet\}$ extracts the real part of a complex value and * denotes the complex conjugate.

Estimator	Direction-dependent
$\tilde{\eta}_{\mathrm{Schwarz1}}$	$\Re\left\{\Gamma_{\mathrm{coh}}^*\left(\Gamma_{\mathrm{diff}} - \hat{\Gamma}_x\right)\right\} \Big/ \left(\Re\left\{\Gamma_{\mathrm{coh}}^*\hat{\Gamma}_x\right\} - 1\right)$
$\tilde{\eta}_{\mathrm{Schwarz2}}$	$\left\lvert\Gamma_{\mathrm{coh}}^*\left(\Gamma_{\mathrm{diff}} - \hat{\Gamma}_x\right) \Big/ \left(\Re\left\{\Gamma_{\mathrm{coh}}^*\hat{\Gamma}_x\right\} - 1\right)\right\rvert$
Estimator	**Direction-independent**
$\tilde{\eta}_{\mathrm{Thiergart2}}$	$\Re\left\{\left(\Gamma_{\mathrm{diff}} - \hat{\Gamma}_x\right) \Big/ \left(\hat{\Gamma}_x - \exp\left(j\angle\hat{\Gamma}_x\right)\right)\right\}$
$\tilde{\eta}_{\mathrm{Schwarz3}}$	[12, (25)]

3. BINAURAL SIGNAL MODEL

When the two microphones are placed at the two ears, the ITD is the propagation delay of the desired sound from the left ear to the right ear and the ILD measures the power level difference between the two microphones. Both the ITD and the ILD have already been widely studied, and various models can be found in [13, 14] and references therein. As in [6], the impact of the ILD is neglected in the following, i.e., we maintain the assumption of equal power at both microphones. Based on this assumption, both the CDRs and the postfilter gain functions are the same at the two microphones placed at the two ears.

In this section, we first describe a simplified model for the frequency-dependent ITD and use it to derive a coherence model for the desired signal component. Then, we describe appropriate models for the diffuse sound field coherence which account for the effect of the head. Finally, we describe the application of these models for binaural CDR estimation and compare the robustness of CDR estimators based on the free-field model to the binaural CDR estimators.

3.1. ITD and Desired Signal Coherence Model

Previous studies have shown that, unlike the TDOA in the free-field case given by (4), the ITD is highly dependent on the frequency, the azimuth angle, the elevation angle and the distance of the desired speaker from the head [14–16]. Here, we use a simplified ITD model to make it applicable for practical application to binaural dereverberation. We assume that the distance of the desired speaker from the head is larger than 1m, and thus does not have a significant effect on the ITD [15, Fig. 9]. Furthermore, we neglect separate consideration of elevation and azimuth angles, and instead model the ITD as a function of the angle θ, which we define as the angle between the direction of the desired speaker and the forward median plane of the head. According to the head-related spherical coordinate system [15, Fig. 7], $\theta = 0$ and $\theta = \pm\pi$ correspond to the forward and the backward median planes of the head, respectively.

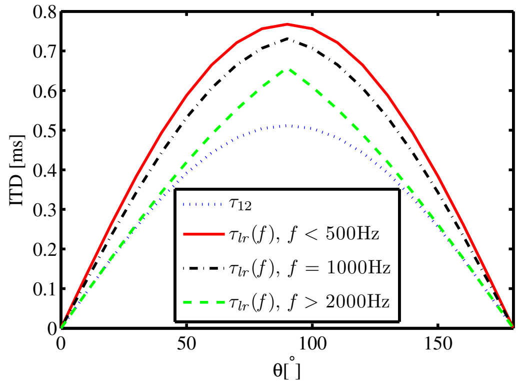

Fig. 1: Comparison of τ_{12} and $\tau_{lr}(f)$ versus the angle of the desired sound for different frequencies.

Kuhn has shown that the ITD is frequency-dependent [16], which can be approximately summarized as

$$\tau_{lr}^{L} = \frac{1.5d\sin(\theta)}{c}, \text{for } f \leq 500\text{Hz}, \tag{9}$$

and, for $f \geq f_H = 2000$,

$$\tau_{lr}^{H} = \begin{cases} 0.5d\,(\sin(\theta)+\theta)/c, \theta \in [-\pi/2, \pi/2], \\ 0.5d\,(\sin(\theta)-(\pi+\theta))/c, \theta \in [-\pi, -\pi/2], \\ 0.5d\,(\sin(\theta)+(\pi-\theta))/c, \theta \in [\pi/2, \pi], \end{cases} \tag{10}$$

where (9) and (10) are identical to [16, (7) and (12)], respectively. However, for the middle frequency range, there is not an explicit expression. We propose to use a linear interpolation to model the ITD in the middle frequency range, which agrees well with the measurement results [16] and is given by

$$\tau_{lr}^{\text{Mid}}(f) = \tau_{lr}^{L} + \frac{\tau_{lr}^{H} - \tau_{lr}^{L}}{f_H - f_L}(f - f_L), \tag{11}$$

where $f = f_{\text{Mid}} \in [500 \;\; 2000]$Hz.

Compared to τ_{12}, the ITD $\tau_{lr}(f)$ is not only a function of the DOA but also of the frequency. τ_{12} and $\tau_{lr}(f)$ versus θ are plotted in Fig. 1 for different frequencies. Fig. 1 shows that the difference between $|\tau_{lr}(f)|$ and $|\tau_{12}|$ is largest for $f \leq 500$Hz. For $f > 2000$Hz, $\tau_{lr}(f)$ is close to τ_{12} when $|\theta|$ or $|\pi \pm \theta|$ is smaller than $\pi/4$, while $|\tau_{lr}(f)|$ is much larger than $|\tau_{12}|$ for $|\theta|$ close to $\pi/2$.

Without the shadowing effect of the head, the free-field coherence model of the desired signal is given by (3). Based on the frequency-dependent ITD model which accounts for the head effect, we can now define the coherence of the desired component for the binaural case as:

$$\Gamma_{\text{coh}}^{\text{Binaural}}(f) = \exp(j\tau_{lr}(f)). \tag{12}$$

3.2. Diffuse Noise Coherence Model

The shadowing effect of the head also has an impact on the spatial coherence of the two microphone signals in a diffuse sound field. Both theoretical results and experimental results can be found in [17, 18]. Here we use the analytic representation of the binaural correlation function proposed by Lindevald and Benade [17], given by

$$\Gamma_{\text{diff}}^{\text{Binaural}}(f) = \frac{1}{\sqrt{1+(\beta 2\pi f d/c)^4}} \frac{\sin(\alpha 2\pi f d/c)}{(\alpha 2\pi f d/c)}, \tag{13}$$

where $\alpha = 2.2$ and $\beta = 0.5$.

The binaural CDR estimators are now obtained by inserting the binaural coherence models $\Gamma_{\text{coh}}^{\text{Binaural}}$ and $\Gamma_{\text{diff}}^{\text{Binaural}}$ into the estimators given in Table 1. This extension makes the direction-dependent CDR estimators suitable for binaural dereverberation. The corresponding estimators are denoted as $\tilde{\eta}_{\bullet}^{\text{Binaural}}$ in the following, where $\bullet$ represents the name of the technique that is being used.

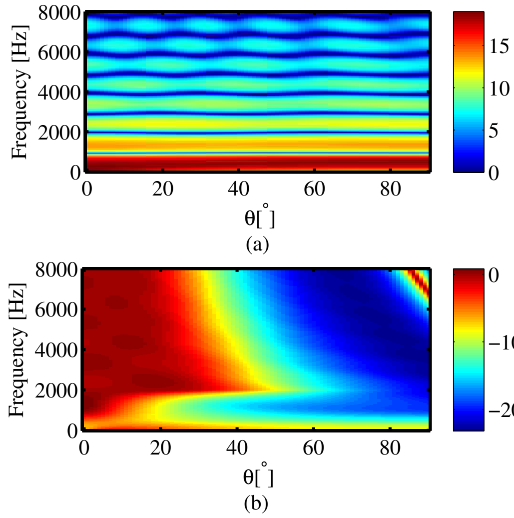

Fig. 2: CDR estimation error of the free-field estimator, Δ, versus f and θ for (a) the input CDR $\eta_{\text{in}} = -20$dB, (b) the input CDR $\eta_{\text{in}} = 20$dB.

3.3. Robustness of the Free-Field Estimators in the Binaural Scenario

This part evaluates the robustness of the direction-dependent CDR estimators using the free-field model against the shadowing effect of the head. For the limited space of this paper, only $\tilde{\eta}_{\text{Schwarz2}}^{\text{Binaural}}$ is chosen to compare with $\tilde{\eta}_{\text{Schwarz2}}^{\text{FF}}$, since a previous study [12] has already shown that $\tilde{\eta}_{\text{Schwarz2}}^{\text{FF}}$ has the best performance among the direction-dependent CDR estimators in Table 1 (see [12, Table III] for details). For the comparison, we generate values of the mixture coherence $\hat{\Gamma}_x$ for a certain input CDR η_{in} and different angles and frequencies according to the binaural coherence models defined above, and insert these coherence values into the free-field estimator. We then define the estimation error of the free-field CDR

Table 2: PESQ scores averaged over all angles for CDR estimators in Table 1, using free-field ($\tilde{\eta}_{\bullet}^{\mathrm{FF}}$) or binaural coherence models ($\tilde{\eta}_{\bullet}^{\mathrm{Binaural}}$).

AIR	Unprocessed	Direction-dependent				Direction-independent			
Distance	Left/Right	$\tilde{\eta}_{\mathrm{Schwarz1}}^{\mathrm{FF}}$	$\tilde{\eta}_{\mathrm{Schwarz1}}^{\mathrm{Binaural}}$	$\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{FF}}$	$\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{Binaural}}$	$\tilde{\eta}_{\mathrm{Thiergart2}}^{\mathrm{FF}}$	$\tilde{\eta}_{\mathrm{Thiergart2}}^{\mathrm{Binaural}}$	$\tilde{\eta}_{\mathrm{Schwarz3}}^{\mathrm{FF}}$	$\tilde{\eta}_{\mathrm{Schwarz3}}^{\mathrm{Binaural}}$
1 m	2.24/2.25	2.40/2.41	2.65/2.68	2.57/2.59	**2.69/2.71**	2.66/2.67	2.64/2.65	2.65/2.67	2.64/2.65
2 m	1.88/1.90	2.00/2.00	2.12/2.13	2.10/2.10	**2.17/2.18**	2.16/2.17	2.15/2.15	2.16/2.17	2.15/2.16
3 m	1.77/1.77	1.85/1.84	1.91/1.90	1.92/1.91	**1.97/1.96**	1.95/1.95	1.95/1.95	1.96/1.96	1.95/1.95

estimator compared to the true CDR η_{in} as

$$\Delta = 10\log_{10}\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{FF}} - 10\log_{10}\eta_{\mathrm{in}}. \quad (14)$$

Fig. 2 plots Δ versus f and θ for the true input CDR $\eta_{\mathrm{in}} = -20$dB (a) and $\eta_{\mathrm{in}} = 20$dB (b). Only $\theta \in [0\ \pi/2]$ is considered due to the symmetry of the scenario. Fig. 2 shows that the CDR is somewhat overestimated for low input CDR, while for high input CDR, the CDR is seriously underestimated for angles larger than 45°. The influence of the head on the coherence, especially the one of the desired speech component $\Gamma_{\mathrm{coh}}^{\mathrm{Binaural}}(f)$, is significant enough to deteriorate the performance of the free-field CDR estimator considerably.

4. EVALUATION

This section evaluates the application of the CDR estimators in Table 1 with the free-field and binaural coherence models to the problem of dereverberation. We use the Aachen Impulse Response (AIR) database [19], which consists of binaural RIRs measured by a dummy head with azimuth angles from -90° to 90° with 15° increments and source-head distances from 1 m to 3 m with 1 m increments.

Ten clean speech samples (five female and five male speakers) are taken from the TIMIT database [20]. The reverberant speech samples are generated by convolving the clean speech with the "stairway" RIRs from the AIR database. We use the same filterbank, postfilter gain function and parameters as in [12, (29)], with the CDR estimators in Table 1. Knowledge of the true DOA is assumed for computation of the desired signal coherence models. The gain function is applied to the two microphone signals separately. PESQ is chosen as evaluation measure since it was found to be highly correlated with speech quality for the evaluation of noise and reverberation suppression methods [21, 22]. Here, we give raw MOS scores obtained by wideband PESQ. The PESQ scores of the two microphone signals and those of the processed signals are given separately. Note that the average PESQ scores for both ears are very similar, due to the symmetry of the scenario. The experimental results for the different distances are presented in Table 2. From these results, we can make the following observations:

(1) Using the ITD and binaural diffuse coherence model can improve all of the direction-dependent CDR estimators.

(2) The direction-independent CDR estimators, which do not rely on a model of the desired signal coherence, are robust in the binaural case, even when using the free-field diffuse coherence model. This indicates that the choice of the diffuse coherence model is not critical, and the main effect of the head is on the ITD.

(3) The direction-independent estimators almost reach the performance of the best binaural direction-dependent estimator.

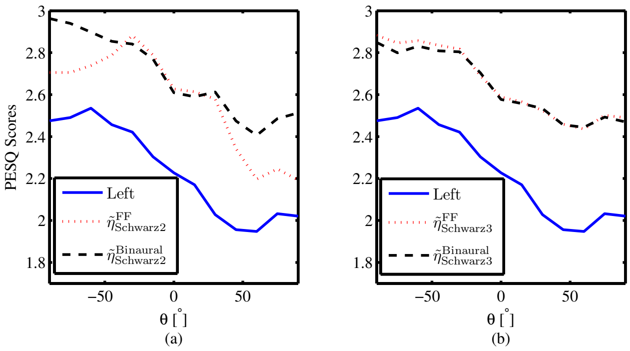

Fig. 3: PESQ scores versus DOA for the 1 m distance case: (a) direction-dependent CDR estimators; (b) direction-independent CDR estimators. *Left* represents the unprocessed signal recorded by the microphone located at the left ear.

To reveal the mechanism of the better performance of the proposed direction-dependent binaural CDR estimators, the PESQ scores versus θ are plotted in Fig. 3 (due to symmetry, only PESQ scores for the left microphone are shown). As can be seen from this figure, $\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{Binaural}}$ is much better than $\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{FF}}$ for $|\theta| \geq 45°$. This phenomenon can be explained by the robustness analysis results in Fig. 2, where it was found that the estimation error of $\tilde{\eta}_{\mathrm{Schwarz2}}^{\mathrm{FF}}$ becomes significant for $|\theta| > 45°$. However, $\tilde{\eta}_{\mathrm{Schwarz3}}^{\mathrm{FF}}$ and $\tilde{\eta}_{\mathrm{Schwarz3}}^{\mathrm{Binaural}}$ nearly have the same performance for all angles, which confirms that the effect of using $\Gamma_{\mathrm{diff}}^{\mathrm{FF}}(f)$ or $\Gamma_{\mathrm{diff}}^{\mathrm{Binaural}}(f)$ is not critical for the direction-independent CDR estimators.

The estimators $\tilde{\eta}_{\mathrm{Thiergart2}}$ and $\tilde{\eta}_{\mathrm{Schwarz3}}$ show similar behavior in this scenario, although the former is biased [12]. This can be explained by the fact that the bias is roughly proportional to the noise coherence and disappears for $\Gamma_{\mathrm{diff}} \to 0$; since, for binaural signals, the noise coherence is lower than for the setup investigated in [12], due to the large spacing of the sensors and the shadowing effect of the head, the practical impact of the bias is not significant here.

5. CONCLUSIONS

This paper extends previously proposed free-field CDR estimators to binaural dereverberation by using a simplified model for the ITD. Experimental results show that this extension is important for the direction-dependent CDR estimators, where PESQ scores for dereverberation can be significantly improved. It is further shown that the direction-independent CDR estimators, which do not require a model of the desired signal coherence, can achieve similar performance and are robust towards the shadowing effect of the head. Further work could concentrate on studying the impact of the ILD on binaural dereverberation and the theoretical limits of the CDR estimators by using statistical analysis [23].

REFERENCES

[1] M. Brandstein, and D. Ward, *Microphone arrays: signal processing techniques and applications*. Berlin: Springer-Verlag, 2001.

[2] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*. Berlin: Springer-Verlag, 2005.

[3] P. A. Naylor, and N. D. Gaubitch, *Speech dereverberation*. London: Springer-Verlag, 2010.

[4] J. B. Allen, D. A. Berkley, and J. Blauert. “Multimicrophone signal-processing technique to remove room reverberation from speech signals.” *J. Acoust. Soc. Am.*, vol. 62, pp. 912-915, 1977.

[5] K. Lebart, J. Boucher, and P. Denbigh. “A binaural system for the suppression of late reverberation.” in *Proc. EUSIPCO*, Island of Rhodes, Greece, 1998.

[6] M. Jeub, M. Schafer, T. Esch, and P. Vary. “Model-based dereverberation preserving binaural cues.” *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 18, pp. 1732-1745, 2010.

[7] A. Kuklasinski, S. Doclo, S. H. Jensen, and J. Jensen. ”Maximum likelihood based multi-channel isotropic reverberation reduction for hearing aids.” in *Proc. EUSIPCO*, Lisbon, Portugal, 2014.

[8] A. Westermann, J. M. Buchholz, and T. Dau. “Binaural dereverberation based on interaural coherence histograms.” *J. Acoust. Soc. Am.*, vol. 133, pp. 2767-2777, 2013.

[9] A. Tsilfidis, A.Westermann, J. M. Buchholz, E. Georganti and J. Mourjopoulos. *Binaural Dereverberation*. Berlin: Springer-Verlag, 2013.

[10] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, U. Rass. “Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends.” *EURASIP J. on Adv. in Signal Process.*, vol. 18, pp. 2915-2929, 2005.

[11] O. Thiergart, G. Del Galdo, and E. A. P. Habets. “Signal-to-reverberant ratio estimation based on the complex spatial coherence between omnidirectional microphones.” in *Proc. ICASSP*, Kyoto, Japan, 2012.

[12] A. Schwarz, and W. Kellermann. “Coherent-to-diffuse power ratio estimation for dereverberation.” *IEEE/ACM Trans. on Audio, Speech and Lang. Process.*, vol. 23, pp. 1006-1018, 2015.

[13] J. Blauert. *Spatial Hearing*. The MIT Press: Harvard MA, 1997.

[14] J. Blauert. *The Technology of Binaural Listening*. Berlin-Heidelberg-New York: Springer-Verlag, 2013.

[15] T. Qu, Z. Xiao, M. Gong, Y. Huang, X. Li, and X. Wu. ”Distance-dependent head-related transfer functions measured with high spatial resolution using a spark gap.” *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 17, no. 6, pp. 1124-1132, Aug. 2009.

[16] G. F. Kuhn. “Model for the interaural time differences in the azimuthal plane.” *J. Acoust. Soc. Am.*, vol. 62, pp. 157-167, 1977.

[17] I. M. Lindevald and A. H. Benade. “Two-ear correlation in the statistical sound fields of rooms.” *J. Acoust. Soc. Am.*, vol. 80, pp. 661-664, 1986.

[18] M. Jeub, M. Dorbecker and P. Vary. “A semi-analytical model for the binaural coherence of noise fields.” *IEEE Signal Process. Letters*, vol. 18, pp. 197-200, 2011.

[19] M. Jeub, M. Schafer, and P. Vary. “A binaural room impulse response database for the evaluation of dereverberation algorithms.” in *Proc. Int. Conf. Digital Signal Process. (DSP)*, Santorini, Greece, 2009.

[20] J. S. Garofolo. “Getting Started With the DARPA TIMIT CD-ROM: An Acoustic-Phonetic Continous Speech Database.” *Nat. Inst. of Standards and Technology (NIST)*, Gaithersburg, MD, 1993.

[21] Y. Hu and P. C. Loizou. “Evaluation of objective quality measures for speech enhancement.” *IEEE Trans. Audio, Speech and Lang. Process.*, vol. 16, pp. 229-238, 2008.

[22] S. Goetze, A. Warzybok, I. Kodrasi, J. O. Jungmann, B. Cauchi, J. Rennies, E. A. P. Habets, A. Mertins, T. Gerkmann, S. Doclo, and B. Kollmeier. “A Study on Speech Quality and Speech Intelligibility Measures for Quality Assessment of Single-Channel Dereverberation Algorithms,.” in *Proc. IWAENC*, Antibes, France, 2014.

[23] C. Zheng, H. Liu, R. Peng,and X. Li. “A Statistical Analysis of Two-Channel Post-Filter Estimators in Isotropic Noise Fields.” *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 21, pp. 336-342, 2013.