CARL VON OSSIETZKY universität OLDENBURG
INSTITUTE OF PHYSICS
MEDICAL PHYSICS
Universität Bremen

Objective Perceptual Quality Assessment for Self-Steering Binaural Hearing Aid Microphone Arrays

Thomas Rohdenburg   Volker Hohmann   Birger Kollmeier
Medical Physics, Faculty V, Carl von Ossietzky University of Oldenburg
thomas.rohdenburg@uni-oldenburg.de

Stefan Goetze   Karl-Dirk Kammeyer
Dept. Of Communications Engineering, University of Bremen, Germany
goetze@uni-bremen.de

Introduction

- A novel ***self-steering beamformer with binaural output*** for a head-worn hearing aid microphone array is presented.
- Direction of arrival estimation (**DOA**) and noise reduction **including** several different **head models**
- Performance evaluation and test in a **realistic sound** scenario using **real-world recordings**
- **Benchmark test** for multi-channel noise reduction schemes with binaural output using **objective quality measures** based on **perceptual models of the auditory system** is proposed for performance evaluation.

Binaural Noise Reduction

- Minimum variance distortionless response (MVDR) beamformers with binaural output stage are promising for hearing aid applications.

→ Spatial information can still be exploited by the listener
→ Target signal direction estimation is needed ⇨ DOA estimation

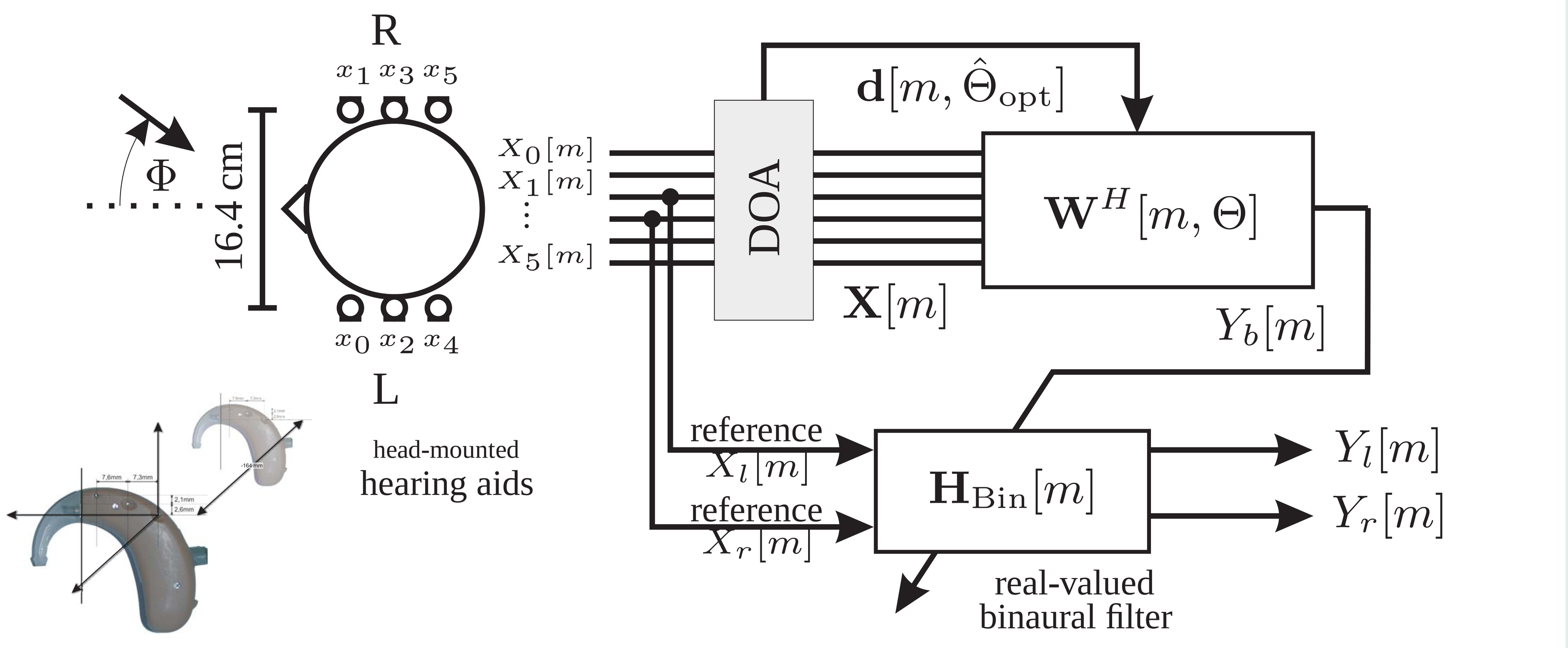


Fig. 1. 6-channel beamformer with head mounted microphone array, DOA estimator and binaural post-filter

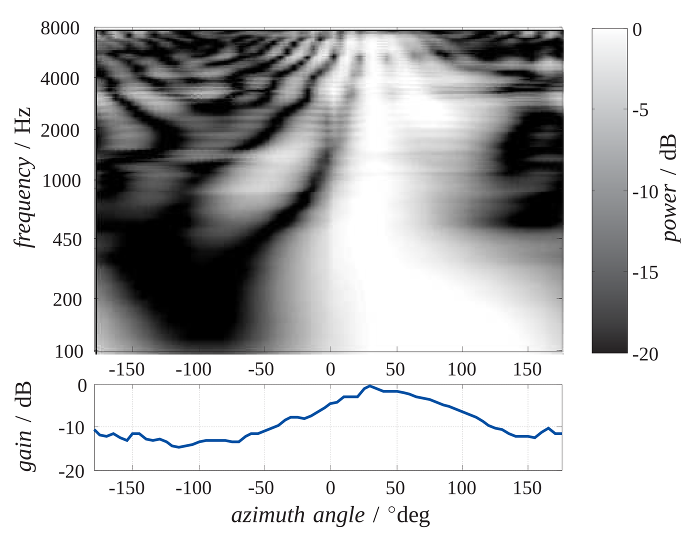


Fig. 2. Beampattern for a 6-channel beamformer as seen on the head

- Classical MVDR beamformer [9,10]

$$\mathbf{W}[m,\Theta] = \frac{\mathbf{\Gamma}_{NN}^{-1}[m]\mathbf{d}[m,\Theta]}{\mathbf{d}^H[m,\Theta]\mathbf{\Gamma}_{NN}^{-1}[m]\mathbf{d}[m,\Theta]}.$$

$$\mathbf{d}[m,\Theta] = [d_0[m,\Theta], d_1[m,\Theta], \ldots, d_{N-1}[m,\Theta]]^T$$
$$d_i[m,\Theta] = |d_i[m,\Theta]| e^{-j2\pi m \frac{f_s}{M}\tau_i[m,\Theta]}, \quad i = 0..N-1$$

- Binaural output stage adapted from [5]

$$H_{\text{Bin}}[m] = \frac{\left(|d_l[m,\Theta]|^2 + |d_r[m,\Theta]|^2\right)\Phi_{Y_bY_b}[m]}{\Phi_{X_lX_l}[m] + \Phi_{X_rX_r}[m]}$$
$$Y_l[m] = H_{\text{Bin}}[m]X_l[m]$$
$$Y_r[m] = H_{\text{Bin}}[m]X_r[m]$$

Signals

- Head related transfer functions (HRTFs) were recorded in anechoic room and in an office environment.
- 3 different signal mixes were evaluated:

	Noise Field	Moving Speaker
Condition 1:	Diffuse noise	anechoic HRTF
Condition 2:	Office noise	Office HRTF
Condition 3:	Cafeteria noise	Office HRTF

Direction of Arrival Estimation (DOA)

- DOA estimation is based on the Generalized Cross Correlation approach (GCC-PHAT) [11]

$$\tau_d = \arg\max_k R_{x_lx_r}[k]$$

$$R_{x_lx_r}[k] = \frac{1}{L_{\text{DFT}}}\sum_{m=0}^{L_{\text{DFT}}-1} \frac{\Phi_{x_lx_r}[m]}{|\Phi_{x_lx_r}[m]|} e^{j\frac{2\pi}{M}mk}$$

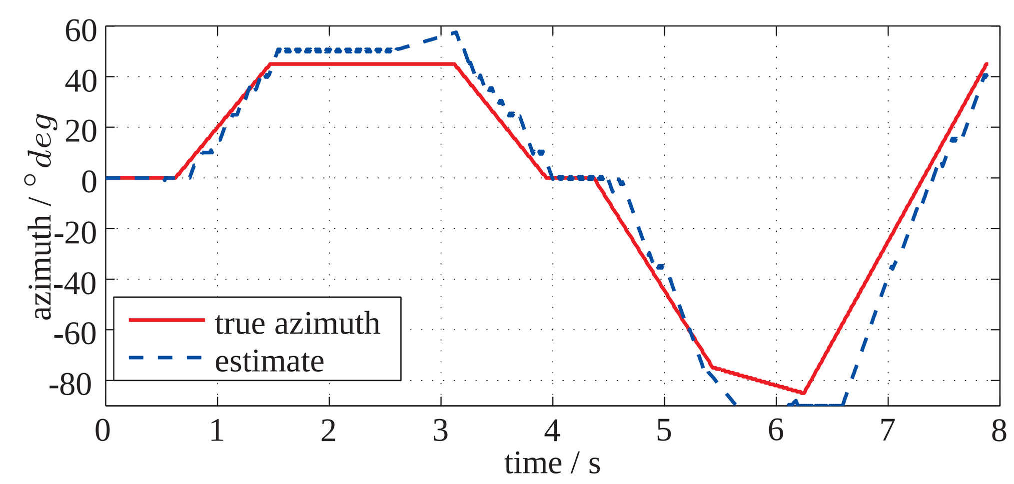


Fig. 3. Virtual azimuth path of a moving speech source and its estimate for HM2 at 12dB SNR. The movement is simulated by time variant filtering of the signal with HRTFs from different directions.

Fig. 4. HRTF Measurements

Propagation Models

- Binaural cues for DOA estimation with head-worn microphone arrays:

→ Interaural level difference (ILD)
→ Interaural time difference (ITD)

- HRTFs should be incorporated into design of DOA estimator (HRTFs user dependent ⇨ head models)

Head Models

- Head models incorporate both, changes of ILD and ITD, due to head shadow and diffraction effects. (Two head models were evaluated HM1 [12] / HM2 [13])
- Interaural Time Difference (**ITD**)

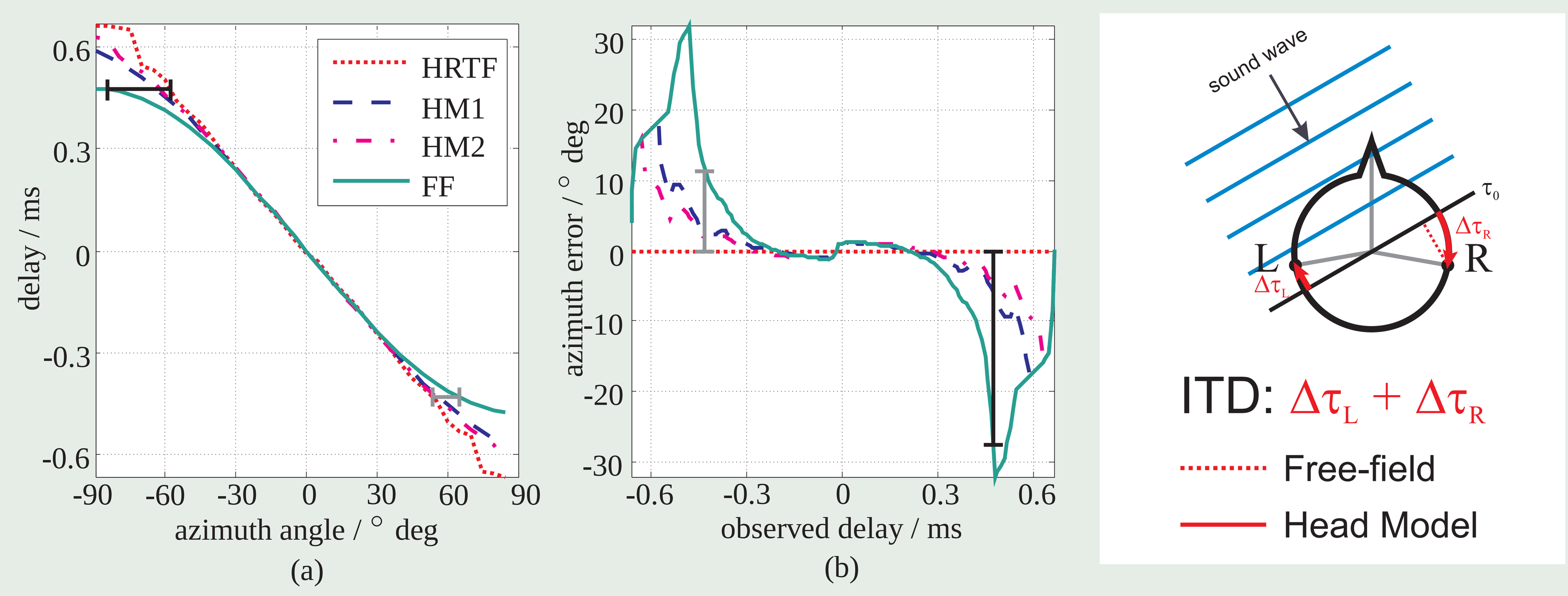


Fig. 5. Comparison of interaural time differences for free-field assumptions, HRTFs, and head models (HM1 / HM2)

- Interaural Level Difference (**ILD**)

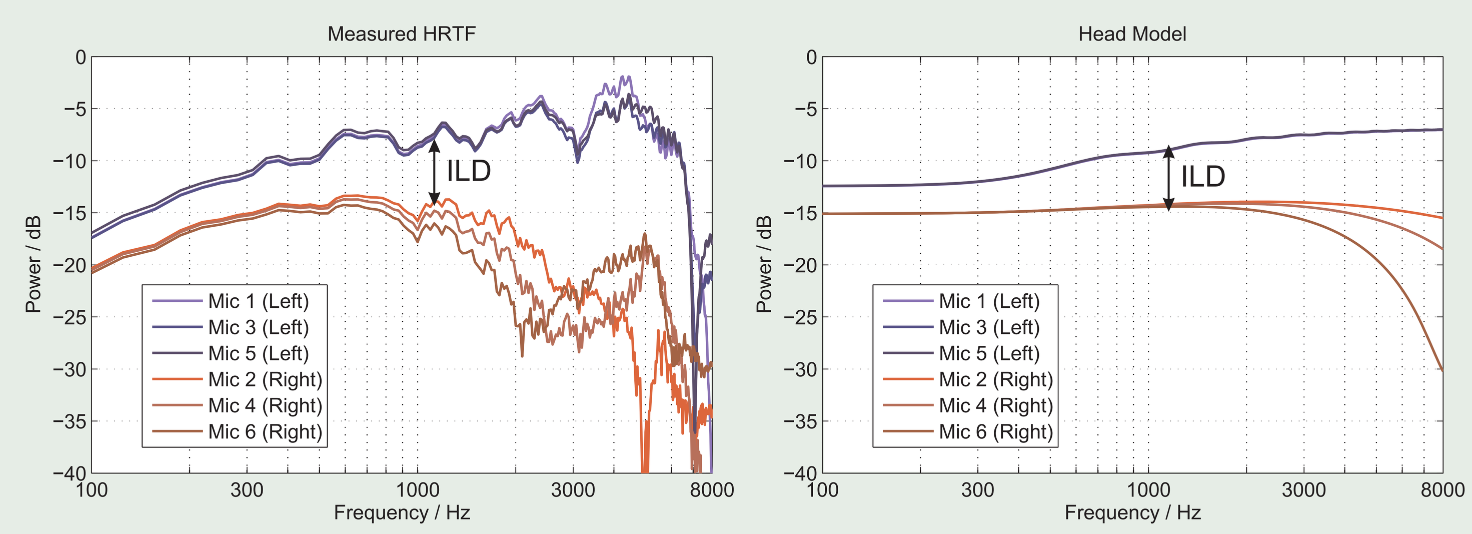


Fig. 6. Comparison of interaural level differences for HRTFs (left) and head models (right, HM2)

References

[1] T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Objective perceptual quality measures for the evaluation of noise reduction schmes," in *9th International Workshop on Acoustic Echo and Noise Control*, Eindhoven, 2005, pp. 169-172.

[2] R. Huber and B. Kollmeier, "PEMO-Q - a new method for objective audio quality assessment using a model of auditory perception.," *IEEE Trans. on Audio, Speech and Language Proc.*, 2006 Special Issue on Objective Quality Assessment of Speech and Audio

[3] R. Beutelmann and T. Brand, "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *Journal of the Acoustical Society of America*, vol. 120, no.1, pp. 331-342, 2006

[4] T. Van den Bogaert, J.Wouters, S. Doclo, and M. Moonen, "Binaural cue preservation for hearing aids using an interaural transfer function multichannel wiener filter," *in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2007.

[5] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. Article ID 63 297, 14 pages, 2006.

[6] S. Goetze, T. Rohdenburg, V. Hohmann, B. Kollmeier and K.-D. Kammeyer, "Direction of Arrival Estimation based on a Dual Delay Line Approach for Binaural Hearing Aid Microphone Arrays, " in *Proc. Int. Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*, Xiamen, China, Nov. 2007.

[7] J. Desloge,W. Rabinowitz, and P. Zurek, "Microphone-array hearing aids with binaural output .i. fixed-processing systems," *IEEE Trans. on Speech and Audio Processing*, vol. 5, no. 6, pp. 529–542, Nov 1997.

[8] D. Welker, J. Greenberg, J. Desloge, and P. Zurek, "Microphonearray hearing aids with binaural output. ii. a two-microphone adaptive system," *IEEE Trans. on Speech and Audio Processing*, vol. 5, no. 6, pp. 543–551, Nov. 1997

CARL VON OSSIETZKY universität OLDENBURG

INSTITUTE OF PHYSICS
MEDICAL PHYSICS

Universität Bremen

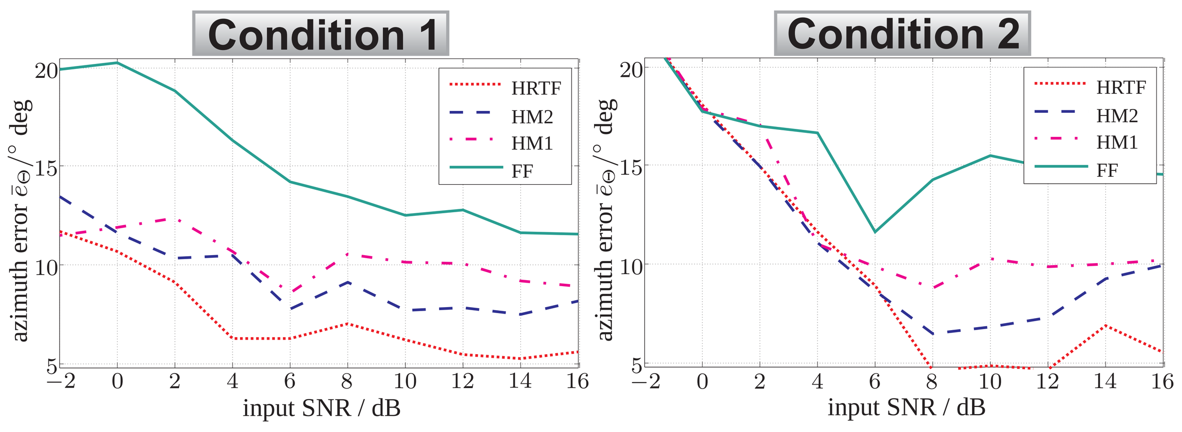
Objective Perceptual Quality Assessment for Self-Steering Binaural Hearing Aid Microphone Arrays

Thomas Rohdenburg Volker Hohmann Birger Kollmeier
Medical Physics, Faculty V, Carl von Ossietzky University of Oldenburg
thomas.rohdenburg@uni-oldenburg.de

Stefan Goetze Karl-Dirk Kammeyer
Dept. Of Communications Engineering, University of Bremen, Germany
goetze@uni-bremen.de

Results

DOA Estimation Error

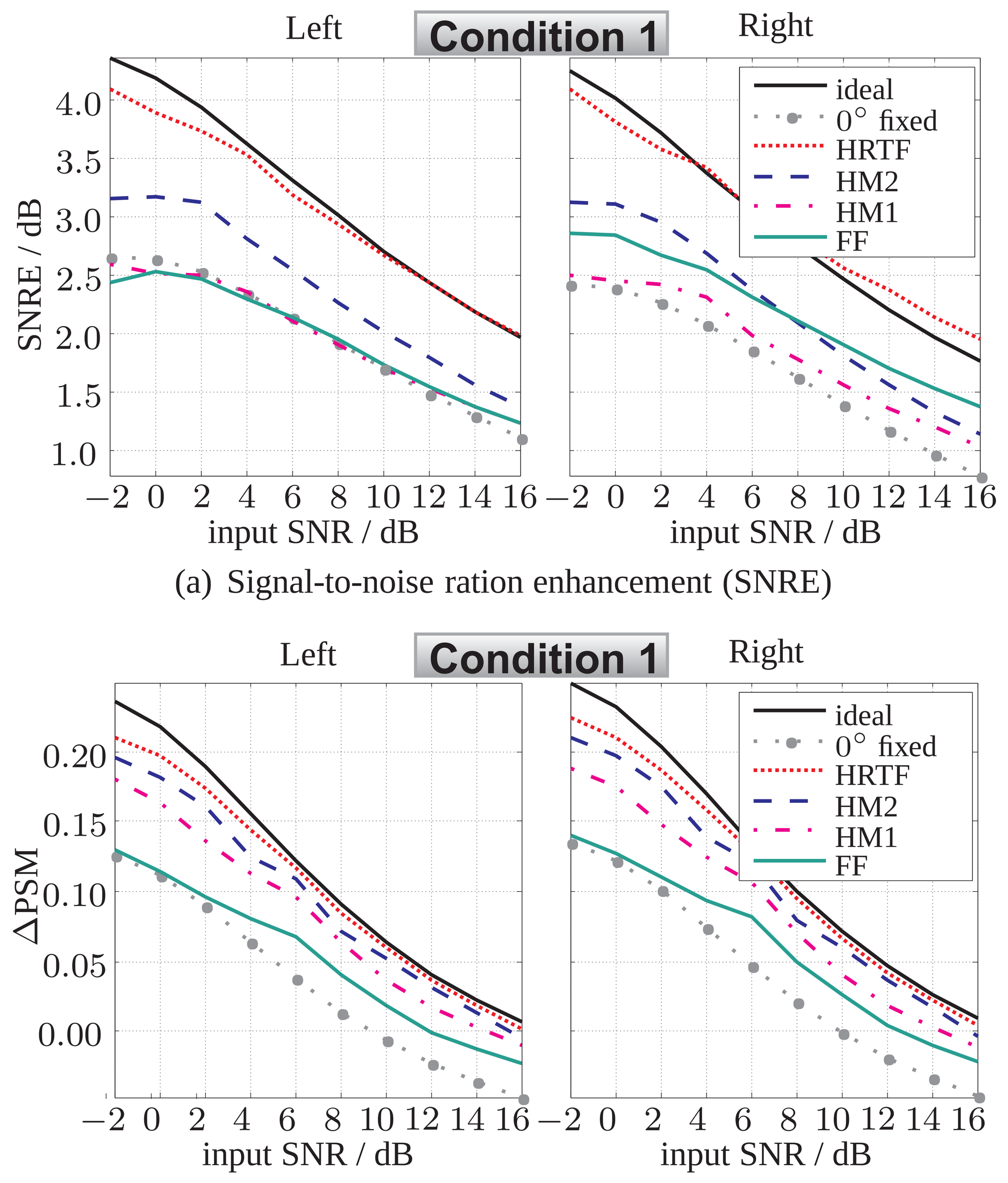


(a) Articial diffuse noise (b) Office noise

Fig. 7. Average error of the direction of arrival (DOA) estimation in different noise conditions.

- Depending on the directivity $\bar{e}_\theta < 10°\text{deg}$ is an acceptable value.
- At least a coarse head model is needed.

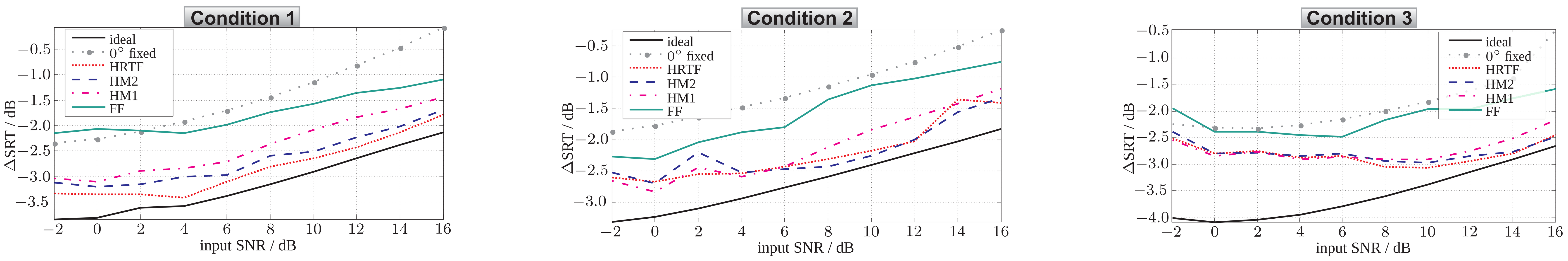
Overall Quality



(a) Signal-to-noise ration enhancement (SNRE)

(b) Perceptual quality enhancement (ΔPSM)

Fig. 8. (a)-(c) Objective quality assessment of DOA plus beamformer system with different propagation models in artificial diffuse noise. Ideal system has a priori information about the direction of arrival and uses the measured HRTF as a propagation model.
The 0°fixed system uses the measured HRTF for beamforming but the DOA is fixed to the look direction (0°deg).



(c) Binaural SRT reduction due to the algorithm (ΔSRT) (d) Binaural SRT reduction in office ambient noise (ΔSRT) (e) Binaural SRT reduction in office with cafeteria noise (ΔSRT)

Fig. 8. (c)-(d) Objective binaural ΔSRT measure shows the expected gain in speech intelligibility due to noise reduction and preservation of binaural information. A lower ΔSRT leads to a higher speech intelligibility. (d)+(e) show the performance results in difficult noise conditions: still the self-steering systems that include at least a coarse head model are superior to the fixed system.

Quality Measures

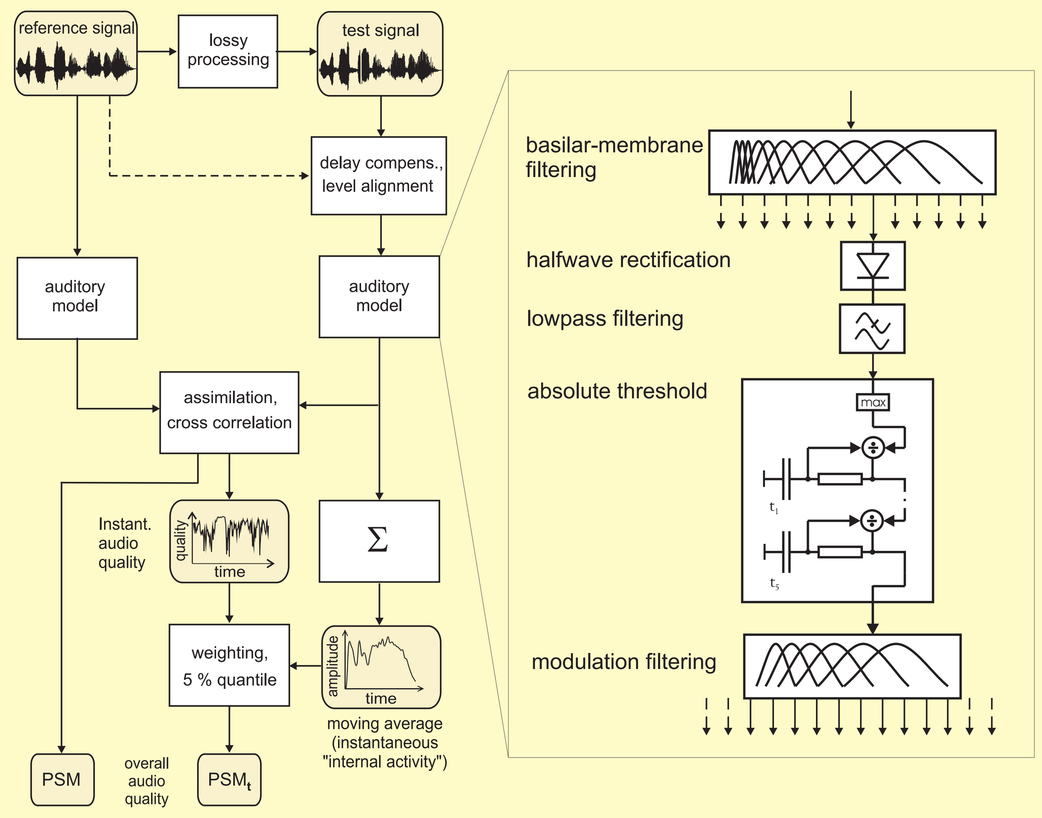


Fig. 9. The quality prediction method PEMO-Q [2]

- Objective performance evaluation is realized in a simulation system with the *shadow filtering* method (i.e. signal and noise are processed separately using the filter coefficients as calculated for the mix).
- Signals can be transformed by models of auditory perception and are compared in the internal representation domain.
- Individual hearing loss can be integrated into auditory models.
- The SNR-Enhancement (SNRE) is the difference of the signal-to-noise ratio (SNR) at the output of the noise reduction system and a reference input SNR, both measured in dB. For binaural systems the SNRE must be calculated bilaterally. By simply taking the average SNRE, the better-ear effect would be ignored.
- The quality measure PSM from PEMO-Q [2] estimates the perceptual similarity between the processed signal and the clean speech source signal. For monaural noise reduction schemes this measure has shown a high correlation with subjective overall quality ratings according to [1,14]. For binaural outputs PSM is measured bilaterally [16].
- The speech reception threshold (SRT) is defined as the SNR at 50% speech intelligibility. In [3] a binaural model of speech intelligibility based on the equalization cancelation (EC) processing by Durlach had been defined which is able to predict the SRT with high accuracy. Here, we are interested in the ΔSRT, i.e. the difference between input and output SRT. The binaural speech intelligibility measure provides an integrative measure of binaural unmasking and can identify differences in the estimated speech-reception threshold (SRT) if binaural information is distorted [16].

Conclusion

- Self-steering beamformer for the application in a binaural hearing aid system shows promising results.
- Performance evaluation under realistic conditions is only realiable using perceptual models of the auditory system [1,16].
- The results show the importance of the propagation model for the DOA estimation and beamforming on a head-worn binaural system.
- The DOA-beamformer system performs best in diffuse or office noise conditions. In adverse noise conditions, such as cafeteria noise, the achievable performance gain is lower compared to a system with perfect knowlegde about the direction of arrival.
- For signal-to-noise ratios greater -2dB self-steering systems are superior to systems that assume a fixed look direction if a certain complexity of the propagation model is met.
- The individual HRTFs are not needed, head models are sufficient.

References

[9] J. Bitzer and K. U. Simmer, "Superdirective microphone arrays," in Microphone Arrays, Brandstein and Ward, Eds. Springer, 2001, ch. 2, pp. 19–38.2005, pp. 169-172.

[10] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. on Antennas Propagation, vol. 30, pp. 27–34, 1982.

[11] C.H. Knapp and G.C. Carter, " The Generalized Correlation Method for Estimation of Time Delay," IEEE Trans. on Acoustics, Speech ans Signal Processing, vol. 24, no. 4, pp. 320-327,1976

[12] P. C. Brown and R. O. Duda, "A structural model for binaural sound synthesis," IEEE Trans. on Speech and Audio Processing, vol. 6, no. 5, pp. 476–488, Sep 1998.

[13] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *Journal of the Acoustical Society of America (JASA)*, vol. 104, no. 5, pp. 3048–3058, 1998.

[14] ITU-T, "Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm," *ITU, Series P: Telephone Transmission Quality Recommendation P.835*, Nov. 2003.

[15] O. Hoshuyama and A. Sugiyama, "Robust adaptive beamforming," *in Microphone Arrays*, Brandstein and Ward, Eds. Springer, 2001, ch. 5, pp. 87–106.

[16] T. Rohdenburg V. Hohmann, and B.Kollmeier, "Robustness Analysis of Binaural Hearing Aid Beamformer Algorithms by means of Objective Perceptual Quality Measures," *in Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, Oct. 2007