Comparison of binaural MVDR-based beamforming algorithms using an external microphone

Nico Gößling, **Simon Doclo**
Dept. of Medical Physics and Acoustics and Cluster of Excellence Hearing4all, University of Oldenburg, Germany

Problem Statement

Objective of **binaural noise reduction algorithm:**

- suppress background noise without introducing signal distortions
- preserve spatial impression of acoustic scene (binaural cues)

This poster: Exploit additional **external microphone** that is spatially separated from the head-mounted **local microphones** to:

1. improve noise reduction and binaural cue preservation performance
2. estimate relative transfer functions of desired source

Blind Binaural MVDR-based Beamforming

Spatial filtering using **all microphones** (local + external), assuming perfect wireless link

Extended binaural MVDR beamformer

Aim: minimize noise PSD while preserving speech component in left and right reference microphone signals [1,2]

$$\min_{\mathbf{w}} \mathcal{E}\{|\mathbf{w}^H\mathbf{n}|^2\} \text{ subject to } \mathbf{w}^H\mathbf{h} = 1 \quad \Rightarrow \quad \mathbf{w}_{\text{MVDR}} = \frac{\mathbf{R}_n^{-1}\mathbf{h}}{\mathbf{h}^H\mathbf{R}_n^{-1}\mathbf{h}}$$

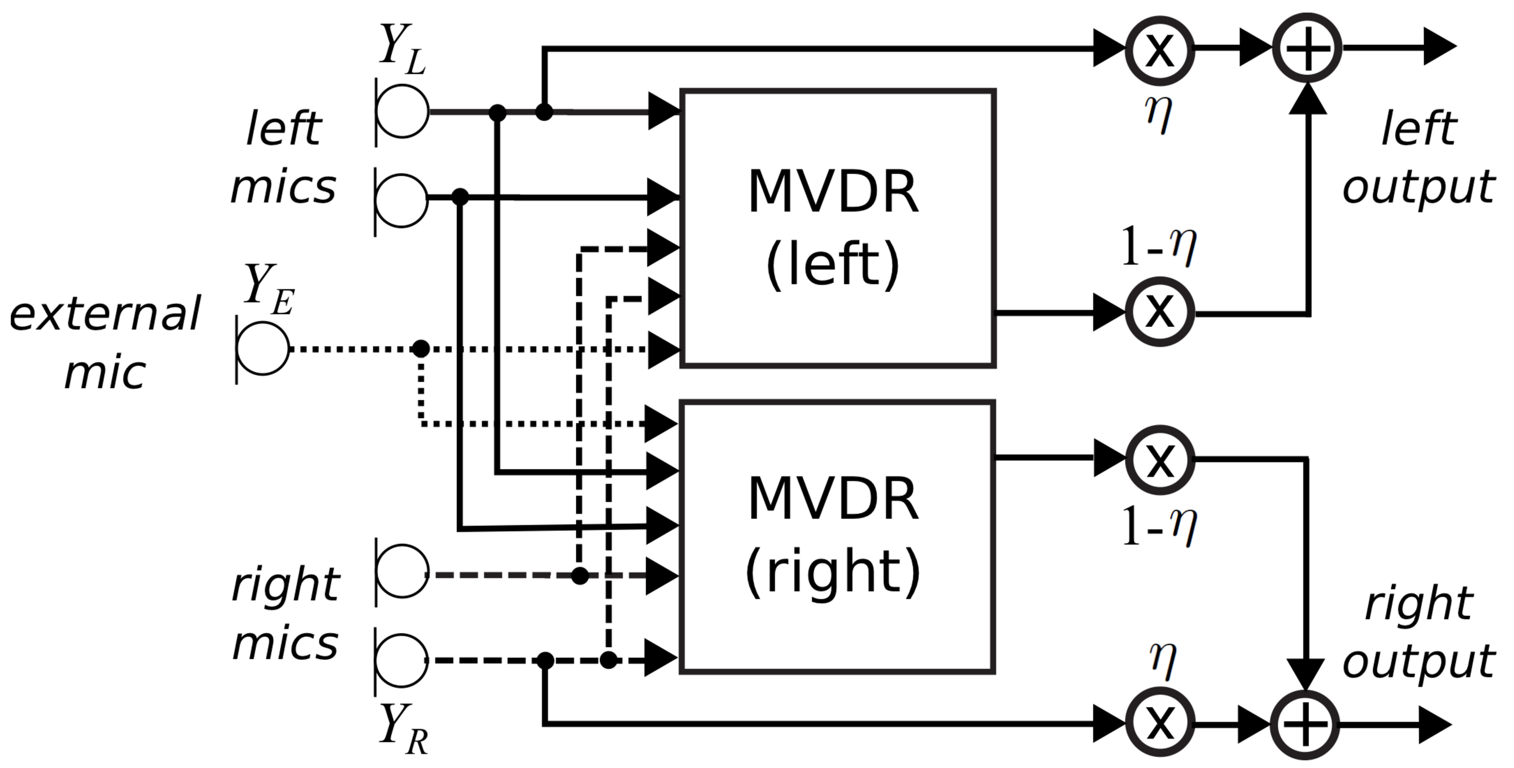
Requires

- Noise covariance matrix $\mathbf{R}_n$
- Relative transfer function (RTF) vector $\mathbf{h}$ of desired source

In contrast to external mic: **preserves binaural cues of desired source**

Extended binaural MVDR beamformer with partial noise estimation

By using an external microphone, a **better binaural cue preservation of the noise** can be achieved using the same mixing parameter, i.e., a smaller mixing parameter achieves the same binaural cue preservation [3]



RTF Estimation

RTF vector consists of part related to local microphones and part related to external microphone $\mathbf{h} = \begin{bmatrix} \bar{\mathbf{h}} \\ h_E \end{bmatrix}$

Local RTF vector $\bar{\mathbf{h}}$

- **Fixed** (anechoic), based on a-priori assumption about desired source position (e.g., in front)
- Estimated using **covariance whitening (CW)** method [4], requiring estimate of $\mathbf{R}_n$, high computational complexity
- Estimated using **spatial coherence (SC)** method [5], assuming coherence between external and local microphones to be zero, low computational complexity

$$\hat{\mathbf{h}}_{L/R} = \frac{\mathcal{E}\{\mathbf{y}Y_E^*\}}{\mathcal{E}\{Y_{L/R}Y_E^*\}}$$

External RTF h_E

- Needs to be estimated, e.g., using CW method

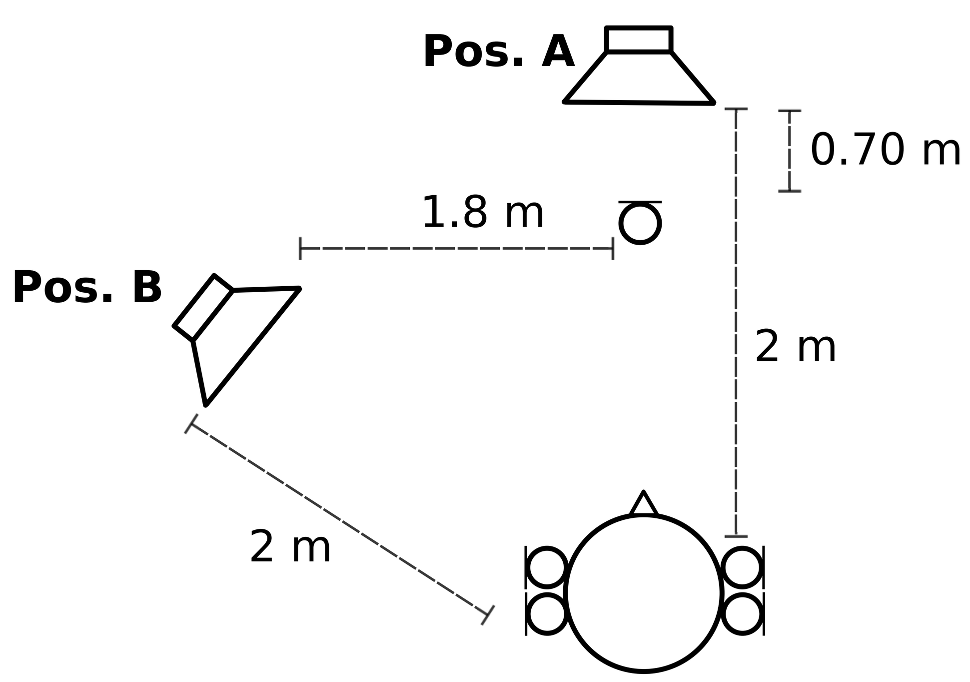
Experimental Setup

- Using **real-world recordings** ($T_{60} \approx 300$ ms) in an **online** implementation with **changing speaker position** (A, B)
- KEMAR with two BTE hearing aids (2 microphones each) and one external microphone
- German speaker (10 sec at position **A**, 10 sec at position **B**)
- Diffuse babble noise
- Intelligibility-weighted input SNR of 0 dB (left reference micropohone)

Algorithm implementation details

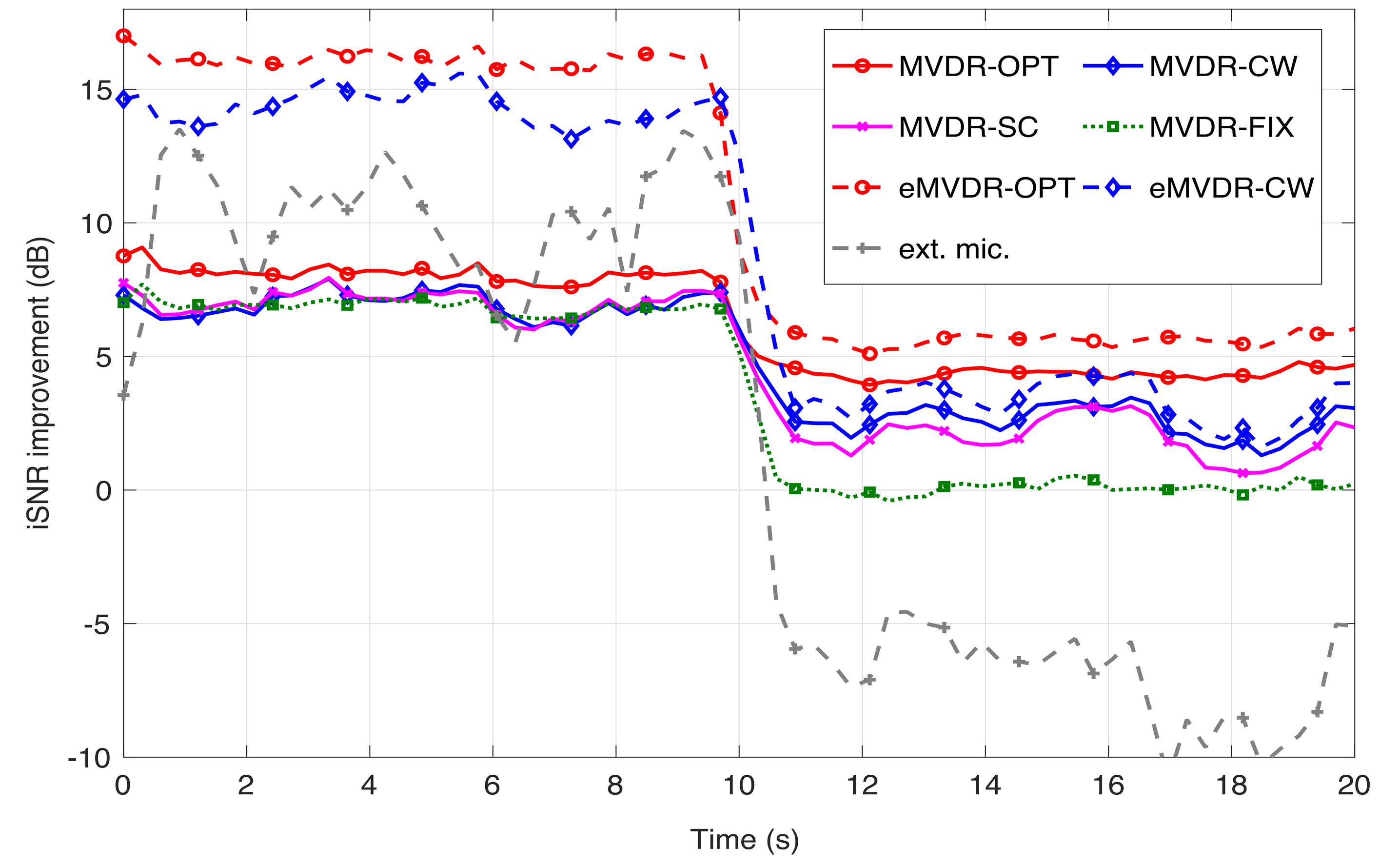
- STFT framework: f_s = 16 kHz, 32 ms frame length, 50% overlap
- SPP-based voice activity detection [6] (threshold 0.6 and 0.4)
- Recursive smoothing time constants: $\tau_y = 150$ ms, $\tau_n = 1.5$ s

Filter	RTF estimation	eMic
MVDR-OPT	oracle	no
MVDR-CW	covariance whitening	no
MVDR-SC	spatial coherence	no
MVDR-FIX	front (anechoic)	no
eMVDR-OPT	oracle	yes
eMVDR-CW	covariance whitening	yes



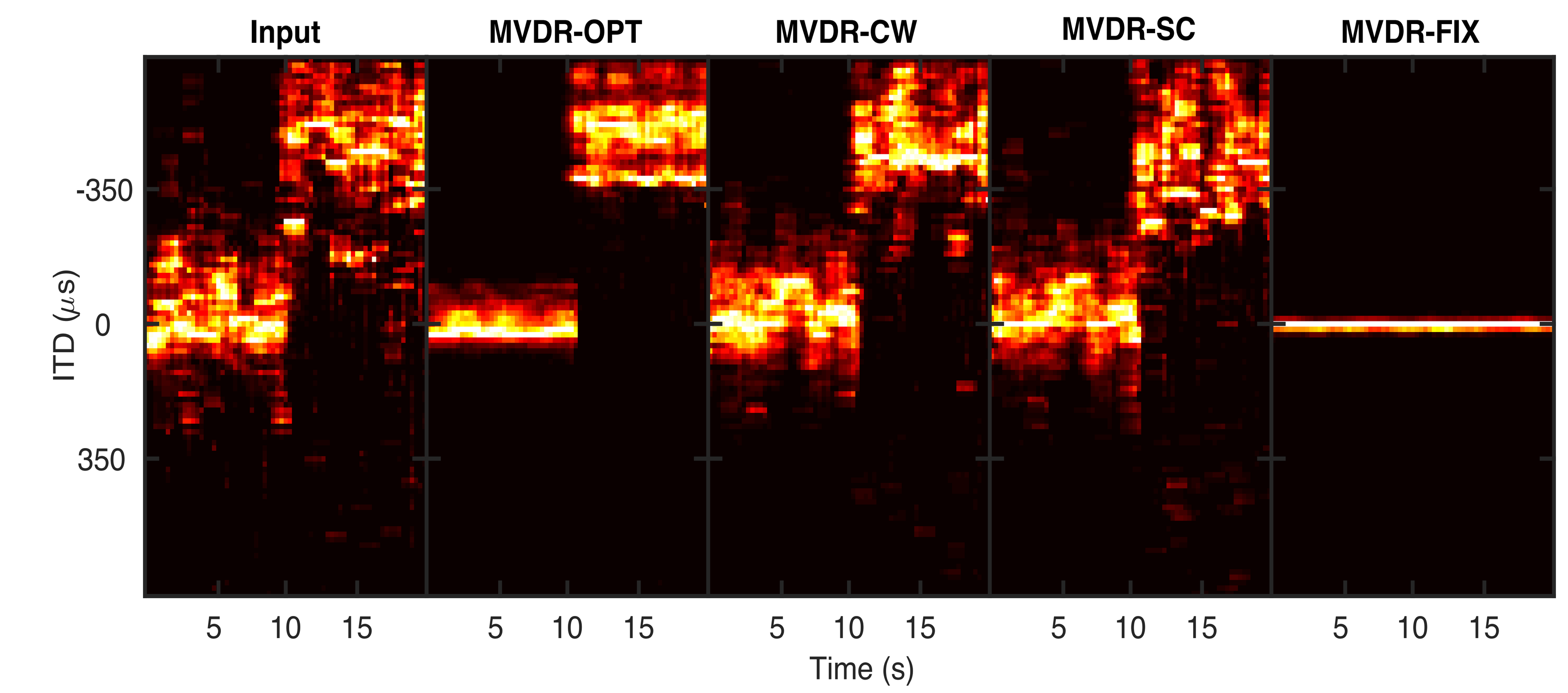
Results

Intelligibility-weighted SNR improvement



- **MVDR with external microphone** leads to better SNR compared to MVDR using only local microphones and **external microphone** alone
- MVDR using local microphones: similar SNR for **SC method** as for CW method at much lower computational complexity, better SNR than **fixed MVDR** (especially for position B)

Binaural cue distribution for desired source



- **External microphone** leads to in-head localisation (no binaural cues)
- **Fixed MVDR** does not preserve binaural cues (especially for position B)
- **MVDR using estimated RTFs** (with/without external microphone): similar binaural cues as in reference microphone signals

Outlook

- Partial noise estimation with external microphone
- Advanced RTF estimation using more than one external microphone

References

[1] S. Doclo, W. Kellermann, S. Makino, S.E. Nordholm, "Multichannel Signal Enhancement Algorithms for Assisted Listening Devices: Exploiting spatial diversity using multiple microphones," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 18–30, Mar. 2015.

[2] D. Marquardt, S. Doclo, "Interaural Coherence Preservation for Binaural Noise Reduction Using Partial Noise Estimation and Spectral Postfiltering," *IEEE/ACM Trans. on Audio, Speech and Language Processing*, vol. 26, no. 7, pp. 1257–1270, Jul. 2018.

[3] N. Gößling, D. Marquardt, S. Doclo, "Performance analysis of the extended binaural MVDR beamformer with partial noise estimation in a homogeneous noise field," in *Proc. Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, San Francisco, USA, 2017.

[4] S. Markovich, S. Gannot, I. Cohen, "Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, pp. 1071–1086, Aug. 2009.

[5] N. Gößling, S. Doclo, "Relative transfer function estimation exploiting spatially separated microphones in a diffuse noise field," in *Proc. International Workshop on Acoustic Signal Enhancement*, Tokyo, Japan, Sep. 2018. (Accepted for publication).

[6] T. Gerkmann, R. C. Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," *IEEE/ACM Trans. on Audio, Speech and Language Processing*, vol. 20, no. 4, pp. 1383–1393, May 2012.