Exploiting external microphones for speech enhancement algorithms in hearing aids

Prof. Dr. Simon Doclo

University of Oldenburg, Dept. of Medical Physics and Acoustics and
Cluster of Excellence Hearing4all

HADF – June 13, 2019

Introduction

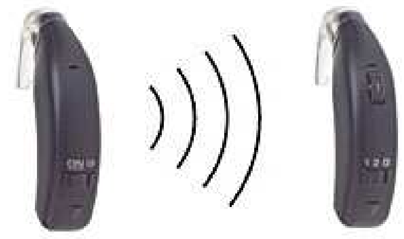
- Hearing impaired suffer from a loss of speech understanding in adverse acoustic environments with competing speakers, background noise and reverberation

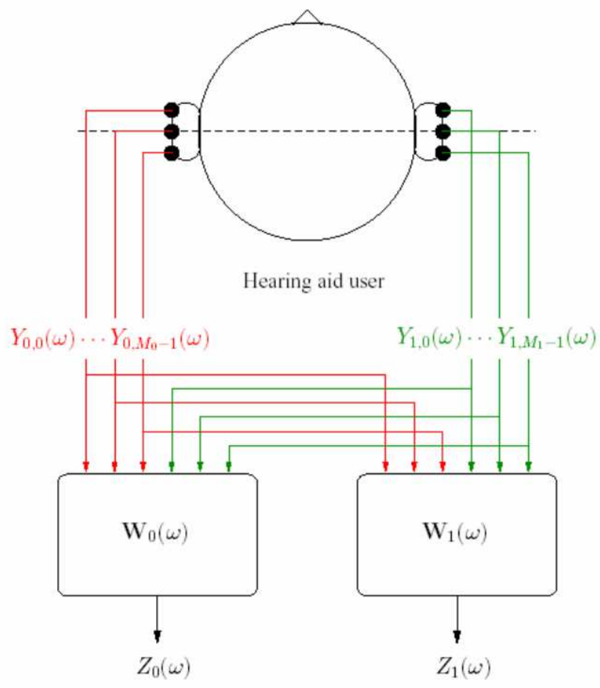
  Multiple microphones available → spatial + spectral processing

Monaural (2-3) *Binaural* *External microphones*

Introduction

- **This presentation:**
  - **Binaural noise reduction algorithms** based on minimum variance distortionless response (MVDR) beamformer
  - Integration with **external microphone(s)** that are spatially separated from the hearing aid microphones





- **Main objectives of algorithms:**
  - Improve speech intelligibility and avoid signal distortions
  - Preserve spatial awareness and directional hearing (binaural cues)

Binaural noise reduction

Binaural MVDR beamformer

Minimum-Variance-Distortionless-Response (MVDR) beamformer

Spatial filtering using **all** microphones (head-mounted and external)

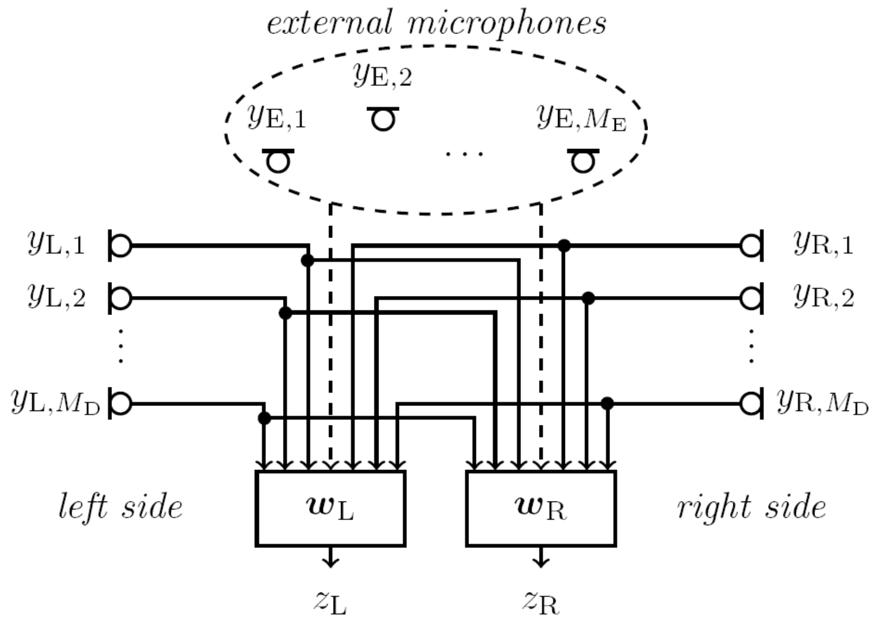
Goal: minimize noise power while preserving speech component in left and right reference microphone signals

$$\min_{\mathbf{w}_L} \mathcal{E}\{|\mathbf{w}_L^H \mathbf{n}|^2\} \quad \text{subject to} \quad \mathbf{w}_L^H \mathbf{x} = X_{L,1}$$

$$\min_{\mathbf{w}_R} \mathcal{E}\{|\mathbf{w}_R^H \mathbf{n}|^2\} \quad \text{subject to} \quad \mathbf{w}_R^H \mathbf{x} = X_{R,1}$$

noise reduction ↑ distortionless constraint ↑

$$\mathbf{w}_L = \frac{\mathbf{R}_n^{-1}\mathbf{h}_L}{\mathbf{h}_L^H \mathbf{R}_n^{-1}\mathbf{h}_L}, \qquad \mathbf{w}_R = \frac{\mathbf{R}_n^{-1}\mathbf{h}_R}{\mathbf{h}_R^H \mathbf{R}_n^{-1}\mathbf{h}_R}$$

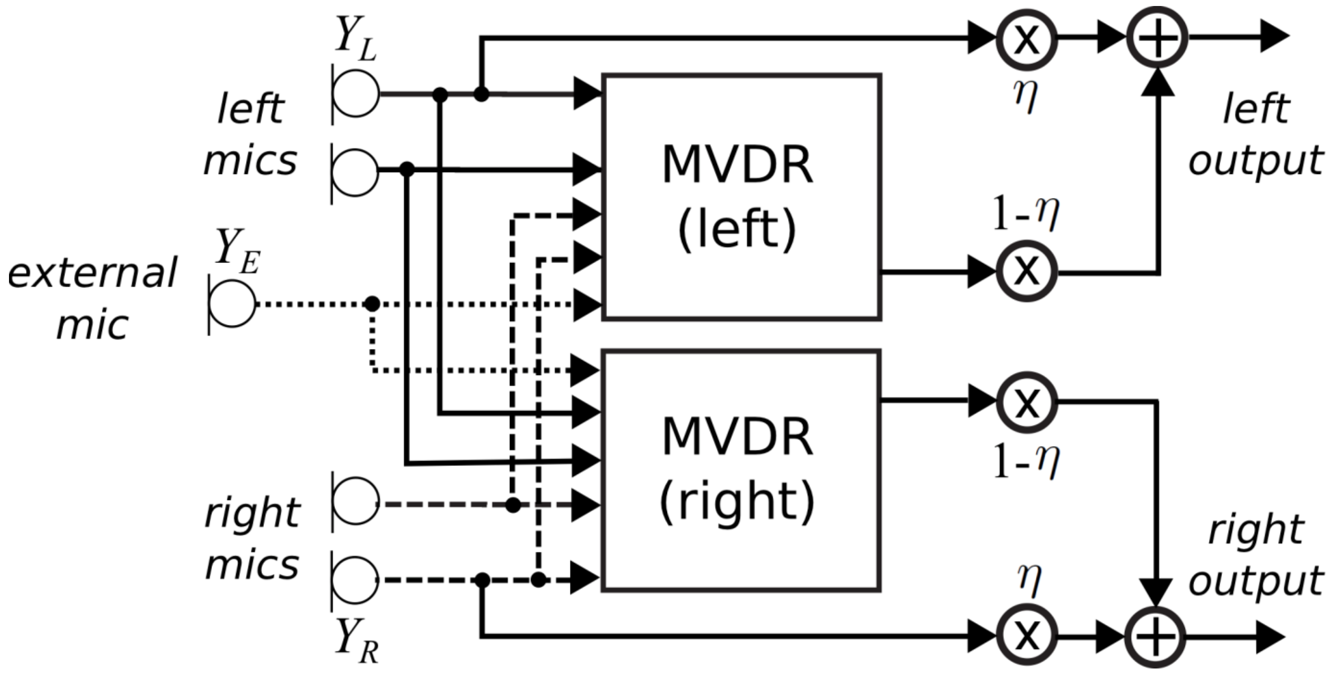


Requires estimate/model of noise covariance matrix (e.g. diffuse) and estimate/model of relative transfer function (RTF) of desired speech source

Preserves **binaural cues** of desired source, but distorts binaural cues of noise

Binaural MVDR beamformer with partial noise estimation

- **Goal:** preserve binaural cues of residual noise by **partly mixing** binaural MVDR output signals with reference microphone signals



- η = **0 :** binaural MVDR (optimal noise reduction, but no cue preservation)
- η = **1 :** reference microphone signals (perfect cue preservation, but no noise reduction)

Note: different procedures available to determine trade-off parameter η (frequency/signal-dependent, psycho-acoustically motivated)

External microphones

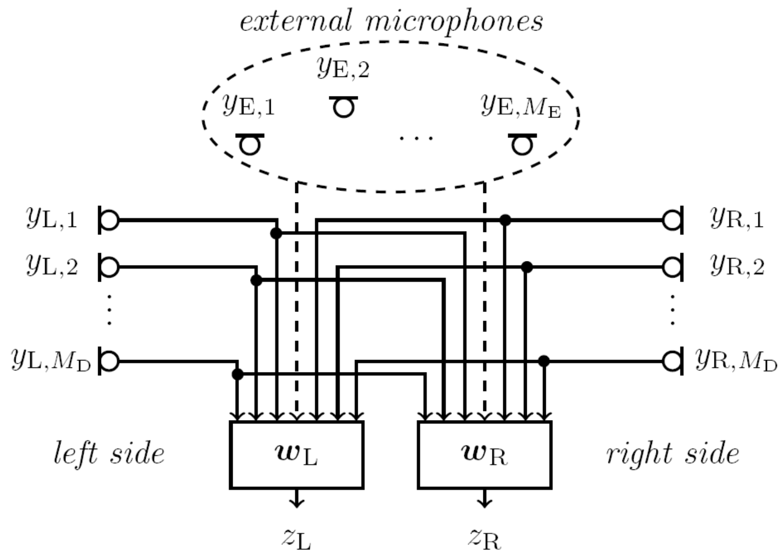
External microphones

- Exploit the availability of one or more external microphones **(acoustic sensor network)** with hearing aids

  [Bertrand 2009, Szurley 2016, Yee 2018, Farmani 2018, Kates 2018, Ali 2019, Gößling 2019]

- Integrating external microphone(s) with hearing aid microphones may lead to:
  - Low-complexity method to **estimate relative transfer function (RTF)** vector of target speaker
  - Improved **noise reduction** and **binaural cue preservation** performance

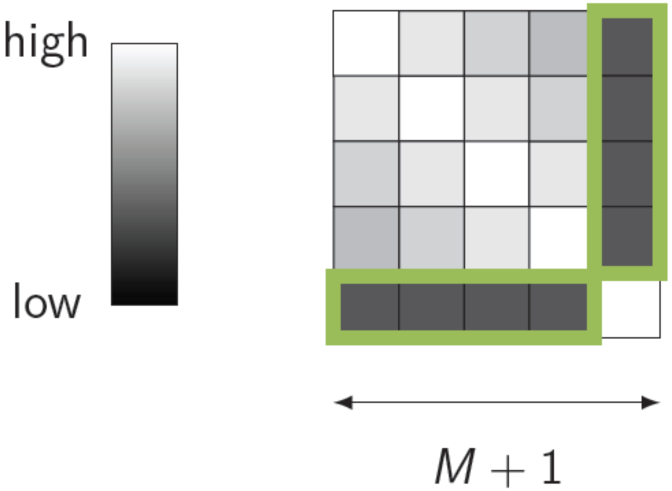
$$\mathbf{w}_L = \frac{\mathbf{R}_n^{-1}\mathbf{h}_L}{\mathbf{h}_L^H\mathbf{R}_n^{-1}\mathbf{h}_L}, \quad \mathbf{w}_R = \frac{\mathbf{R}_n^{-1}\mathbf{h}_R}{\mathbf{h}_R^H\mathbf{R}_n^{-1}\mathbf{h}_R}$$



One external microphone: RTF estimation

- **Estimate RTF vector of target speaker** to steer binaural MVDR beamformer
- **Spatial coherence (SC) method:** assume that noise components in external microphone and HA microphones are uncorrelated, e.g., when external microphone is spatially separated from HA microphones + diffuse noise field

  → correlate HA microphone signals with external microphone signals and normalize by reference element
- **Low computational complexity** with similar (even better in practice) performance than state-of-the-art covariance whitening approach

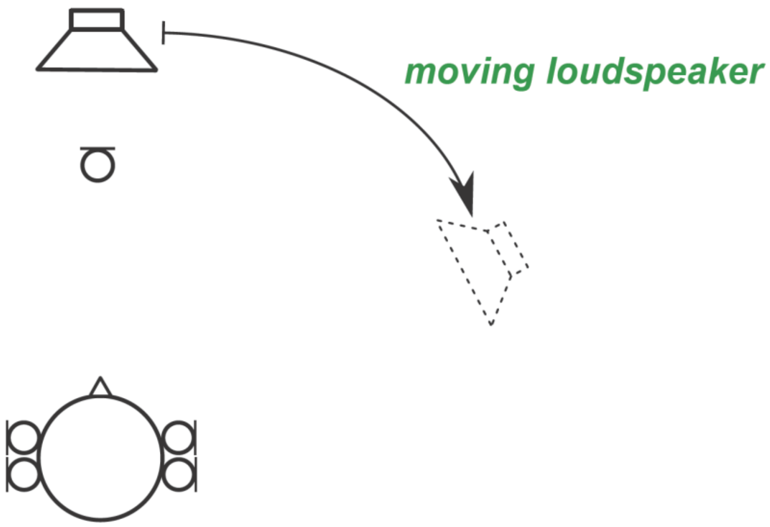


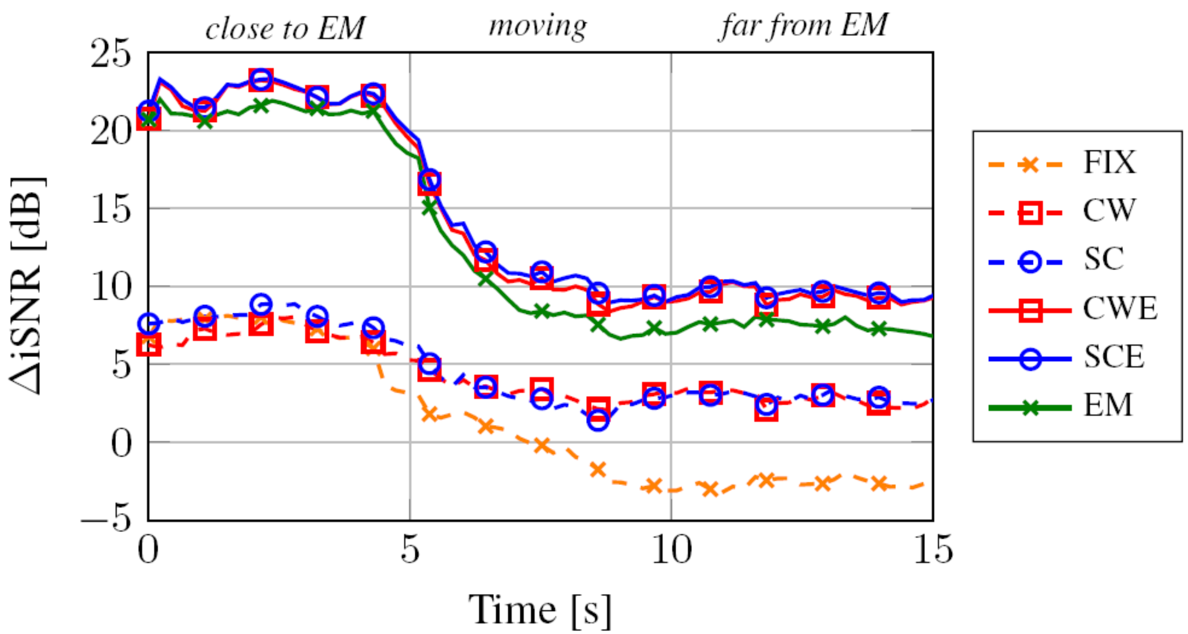
Low spatial coherence between noise components in local array signals and external microphone signal

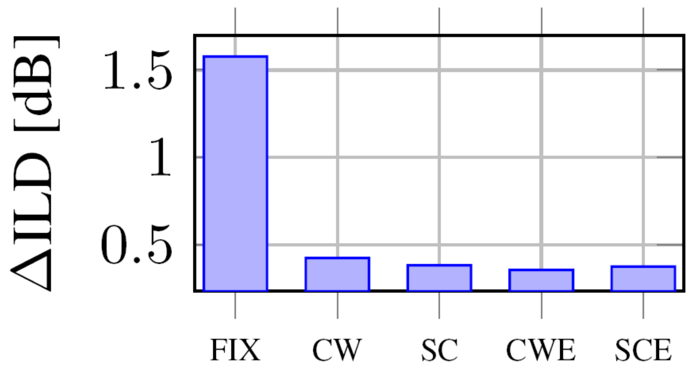
$$\bar{\mathbf{a}}_{\mathrm{L}}^{\mathrm{SCE}} = \frac{\bar{\mathbf{R}}_{\mathrm{y}}\mathbf{e}_{\mathrm{E}}}{\mathbf{e}_{\mathrm{L}}^{T}\bar{\mathbf{R}}_{\mathrm{y}}\mathbf{e}_{\mathrm{E}}}, \quad \bar{\mathbf{a}}_{\mathrm{R}}^{\mathrm{SCE}} = \frac{\bar{\mathbf{R}}_{\mathrm{y}}\mathbf{e}_{\mathrm{E}}}{\mathbf{e}_{\mathrm{R}}^{T}\bar{\mathbf{R}}_{\mathrm{y}}\mathbf{e}_{\mathrm{E}}}$$

$$\bar{\mathbf{w}}_{\mathrm{L}}^{\mathrm{SCE}} = \begin{bmatrix} \alpha \cdot [\mathbf{I}_{2M}, \mathbf{0}_{2M\times 1}]\,\bar{\mathbf{w}}_{\mathrm{L}} \\ \alpha(1+\beta)\cdot \mathbf{e}_{\mathrm{E}}^{T}\bar{\mathbf{w}}_{\mathrm{L}} \end{bmatrix}$$

One external microphone: Simulation results







- MVDR with external microphone (SCE) leads to **better SNR** compared to MVDR using only HA microphones (SC,FIX) and external microphone (EM)
- MVDR using estimated RTFs (SCE, SC) **preserves binaural cues of target speaker** compared to fixed MVDR (FIX) and external microphone (EM)

Oldenburg Varechoic Lab ($T_{60} \approx$ 350ms), M=4 + 1 external mic (1.5m/0.5m), moving speaker, pseudo-diffuse babble noise, iSNR=0dB (right HA)
STFT: 32 ms, 50% overlap, sqrt-Hann; SPP in external microphone; smoothing: 100 ms (speech), 1 s (noise)

Multiple external microphones

- Each external microphone yields (different) RTF estimate
- **Linear combination/selection** of RTF estimates

$$\boldsymbol{a}_{\mathrm{L}}^{\mathrm{SC-C}} = \frac{\boldsymbol{A}_{\mathrm{L}}^{\mathrm{SC}}\boldsymbol{c}}{\boldsymbol{e}_{\mathrm{L}}^{T}\boldsymbol{A}_{\mathrm{L}}^{\mathrm{SC}}\boldsymbol{c}}$$

1. Input SNR-based selection

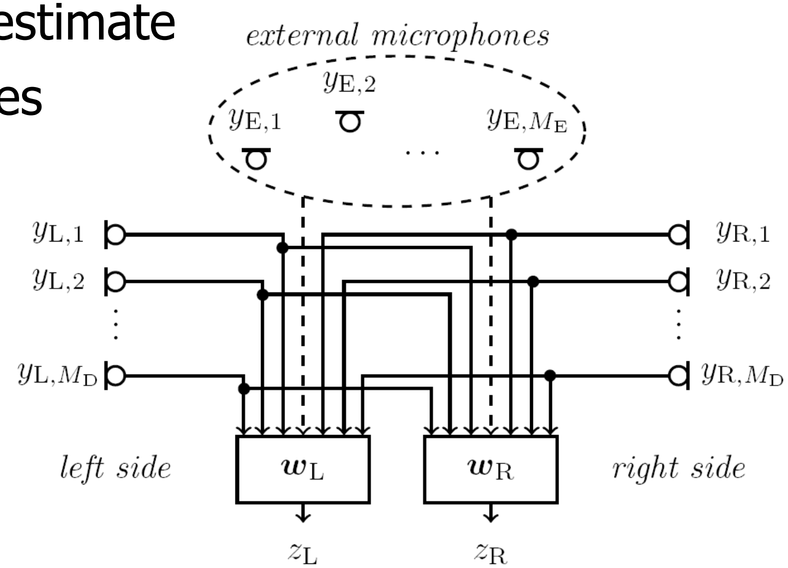
$$\boldsymbol{c}^{\mathrm{iSNR}} = \boldsymbol{e}_{\mathrm{E},\hat{i}}\,, \quad \hat{i} = \arg\max_{i} \frac{\boldsymbol{e}_{\mathrm{E},i}^{T}\boldsymbol{R}_{\mathrm{y}}\boldsymbol{e}_{\mathrm{E},i}}{\boldsymbol{e}_{\mathrm{E},i}^{T}\boldsymbol{R}_{\mathrm{n}}\boldsymbol{e}_{\mathrm{E},i}}$$

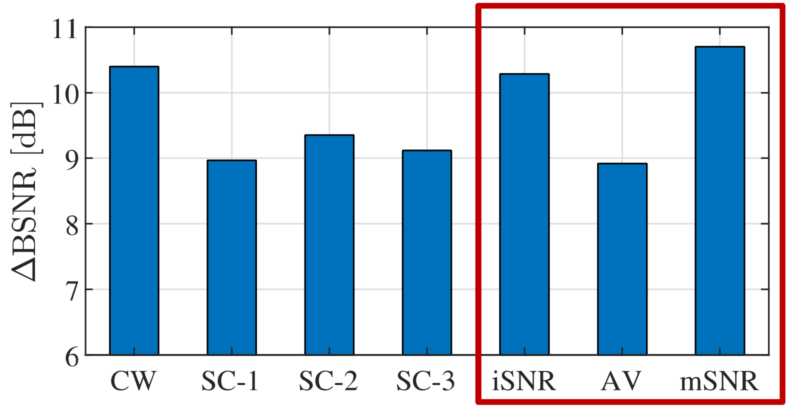
2. Simple averaging

$$\boldsymbol{c}^{\mathrm{AV}} = \left[\frac{1}{M_{\mathrm{E}}}, \dots, \frac{1}{M_{\mathrm{E}}}\right]^{T}$$

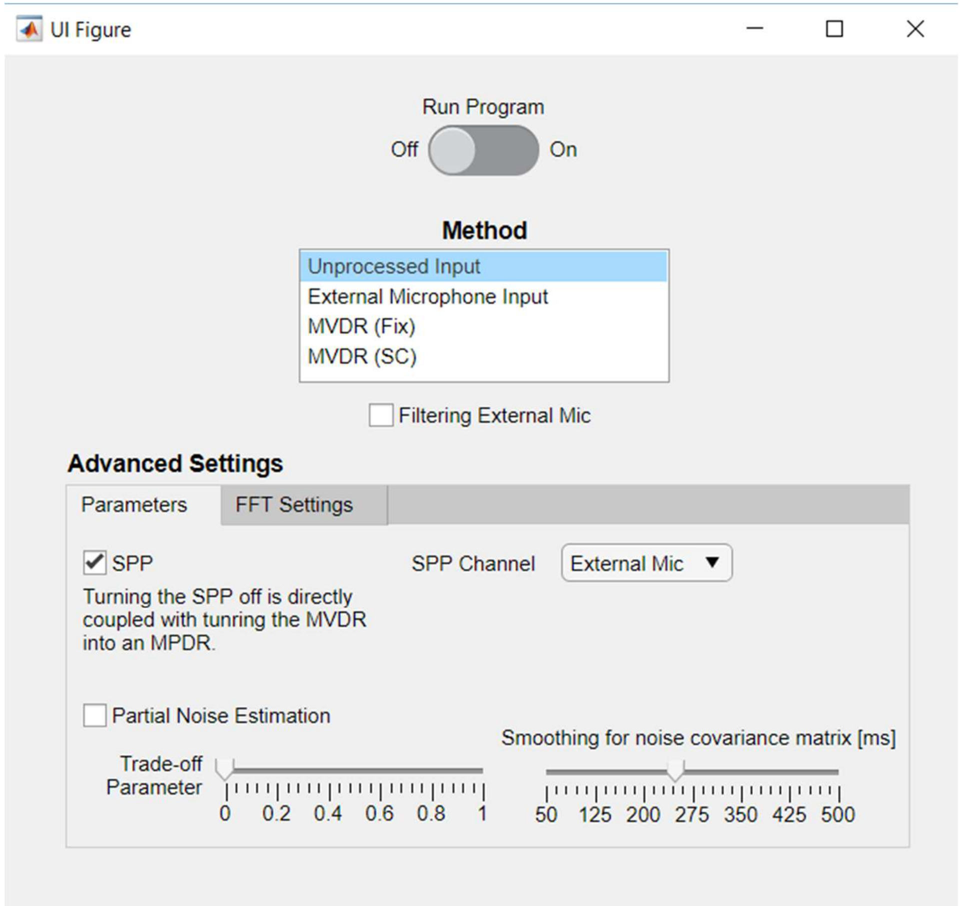
3. Output SNR-maximizing combination

$$\boldsymbol{c}^{\mathrm{mSNR}} = \arg\max_{\boldsymbol{c}} \mathrm{SNR}_{\mathrm{BMVDR,L}}^{\mathrm{out}} = \mathcal{P}\{\boldsymbol{\Lambda}_2^{-1}\boldsymbol{\Lambda}_1\}$$



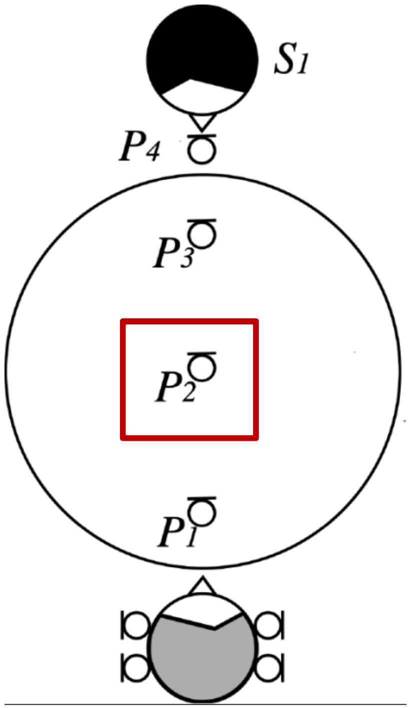


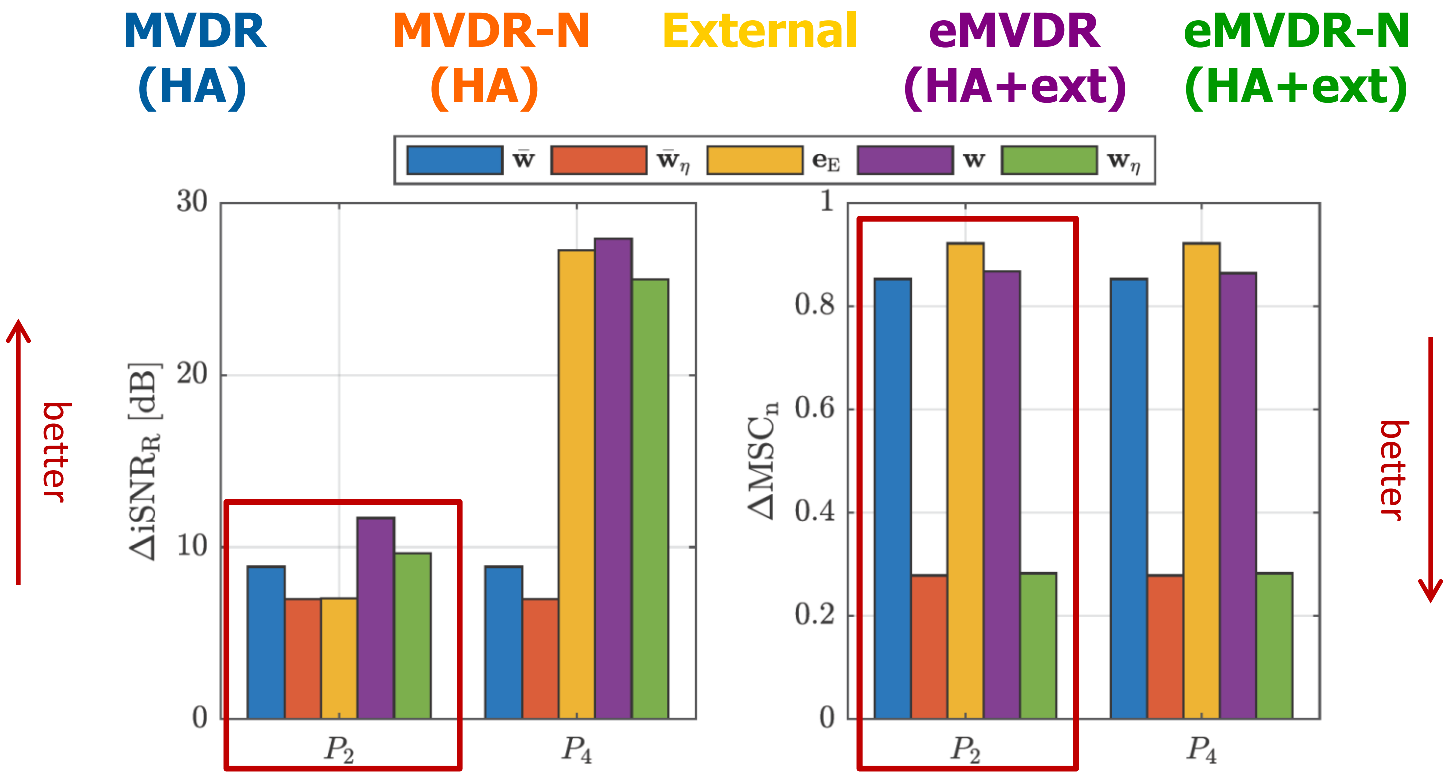
Real-time demonstrator (during coffee break)



Binaural MVDR-N beamformer

- Including external microphone in **binaural MVDR-N beamformer** leads to:
  - **Larger output SNR** for same trade-off parameter η
  - Same output SNR with larger trade-off parameter η → **better cue preservation**





Starkey database with real-world recordings ($T_{60} \approx 620$ms), M=4, target speaker S_1, multi-talker babble noise, 0 dB input iSNR (right hearing aid)
MVDR: perfectly estimated noise correlation matrix, RTF of target speaker estimated using covariance whitening method

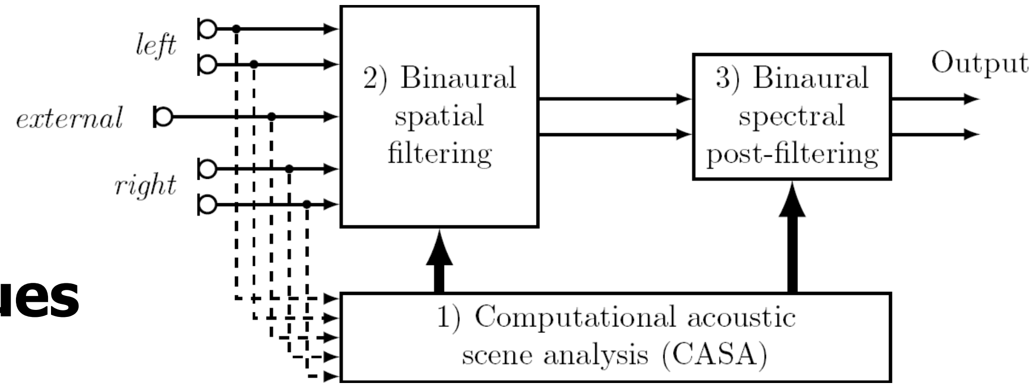
Current/future work

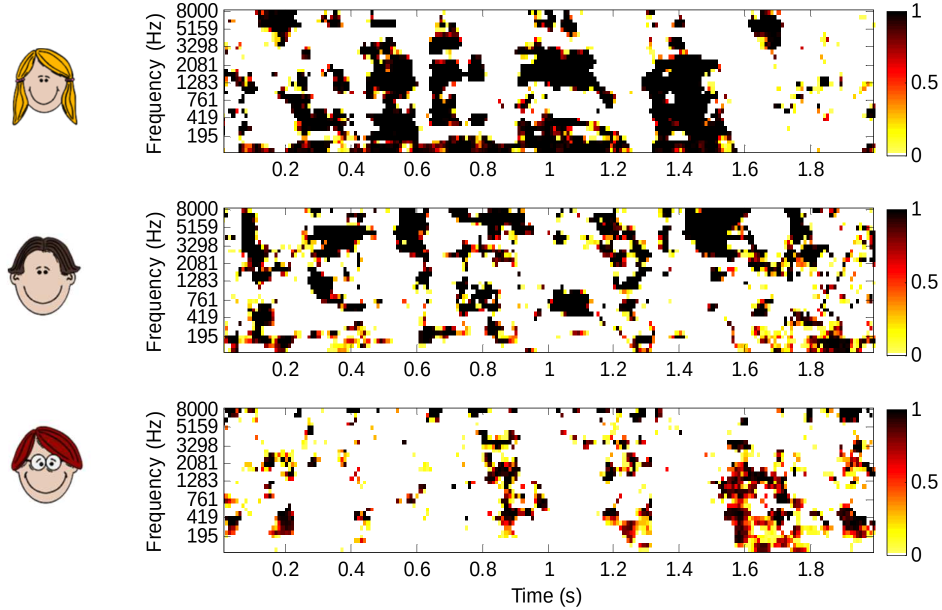
- **Performance analysis** for different acoustic scenarios (interfering speakers)

- **Synchronization/latency issues**

- **Complex and time-varying scenarios**: incorporate computational acoustic scene analysis (CASA) into control path of developed algorithms

- **Subjective evaluation** of binaural speech enhancement algorithms with **HA/CI users** ongoing





Acknowledgments

Dr. Daniel Marquardt

Nico Gößling

Wiebke Middelberg

Dr. Elior Hadad

Prof. Sharon Gannot

- **Funding:**
  - Cluster of Excellence Hearing4all (DFG)
  - Collaborative Research Centre SFB 1330 Hearing acoustics (DFG)
  - Joint Lower-Saxony Israel Project "Acoustic scene aware speech enhancement for binaural hearing aids" (Partner: Bar-Ilan University, Israel)

DFG Deutsche Forschungsgemeinschaft

VolkswagenStiftung

Recent publications

- **S. Doclo, S. Gannot, D. Marquardt, E. Hadad, "Binaural Speech Processing with Application to Hearing Devices", Chapter 18 in *Audio Source Separation and Speech Enhancement* (E. Vincent, T. Virtanen, S. Gannot, eds.), Wiley, 2018.**
- **S. Doclo, W. Kellermann, S. Makino, S. Nordholm, *Multichannel signal enhancement algorithms for assisted listening devices*, *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 18-30, Mar. 2015.**
- D. Marquardt, V. Hohmann, S. Doclo, *Interaural Coherence Preservation in Multi-channel Wiener Filtering Based Noise Reduction for Binaural Hearing Aids*, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 23, no. 12, pp. 2162-2176, Dec. 2015.
- J. Thiemann, M. Müller, D. Marquardt, S. Doclo, S. van de Par, *Speech Enhancement for Multimicrophone Binaural Hearing Aids Aiming to Preserve the Spatial Auditory Scene*, *EURASIP Journal on Advances in Signal Processing*, 2016:12, pp. 1-11.
- D. Marquardt, S. Doclo, *Interaural Coherence Preservation in Binaural Hearing Aids using Partial Noise Estimation and Spectral Postfiltering*, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 26, no. 7, pp. 1257-1270, Jul. 2018.
- 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 (HSCMA)*, San Francisco, USA, Mar. 2017, pp. 1-5.
- 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, pp. 146-150.
- N. Gößling, S. Doclo, *RTF-steered Binaural MVDR Beamforming Incorporating an External Microphone for Dynamic Acoustic Scenarios*, in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, May 2019, pp. 416-420.
- N. Gößling, W. Middelberg, S. Doclo, *RTF-steered Binaural MVDR Beamforming Incorporating Multiple External Microphones*, submitted to IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, USA, Oct. 2019.

http://www.sigproc.uni-oldenburg.de → Publications

Questions ?


House of Hearing, Oldenburg