Subjective Evaluation of Signal-Dependent Partial Noise Preservation Algorithms for Binaural Hearing Aids

Jonas Klug[1], Daniel Marquardt[2], Simon Doclo[1]

[1]Dept. of Medical Physics and Acoustics and Cluster of Excellence Hearing4all, University of Oldenburg, Germany

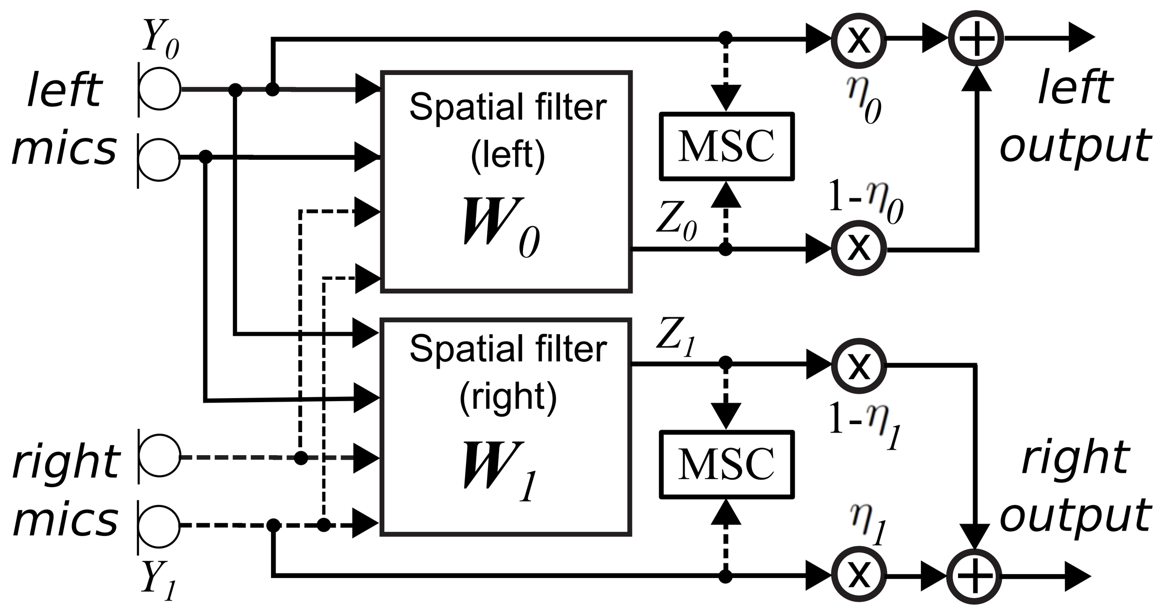
[2]Starkey Hearing Technologies, Eden Prairie MN, USA

Problem Statement

- Objective of **binaural noise reduction algorithm**:
  - increase speech intelligibility: suppress undesired sound sources
  - preserve spatial impression of acoustic scene (binaural cues of all sound sources)
- Binaural minimum variance distortionless response beamformer with partial noise estimation (BMVDR-N)
  - preserves binaural cues of target speaker
  - parameter allows to **trade off noise reduction and binaural cue preservation of background noise**

- **This poster:** signal-dependent method to determine trade-off parameter based on coherence between noisy input signals and output signals of BMVDR beamformer

Binaural Noise Reduction



- **BMVDR beamformer** [1, 2]: minimize PSD of the noise component while preserving speech component in reference microphone signals at left and right hearing aid

$$\min_{\mathbf{W}_0} \mathcal{E}\{|\mathbf{W}_0^H \mathbf{V}|^2\} \quad \text{subject to} \quad \mathbf{W}_0^H \bar{\mathbf{A}}_0 = 1$$

$$\min_{\mathbf{W}_1} \mathcal{E}\{|\mathbf{W}_1^H \mathbf{V}|^2\} \quad \text{subject to} \quad \mathbf{W}_1^H \bar{\mathbf{A}}_1 = 1$$

$$\mathbf{W}_{\text{BMVDR},0} = \frac{\mathbf{R}_v^{-1}\bar{\mathbf{A}}_0}{\bar{\mathbf{A}}_0^H \mathbf{R}_v^{-1}\bar{\mathbf{A}}_0} \qquad \mathbf{W}_{\text{BMVDR},1} = \frac{\mathbf{R}_v^{-1}\bar{\mathbf{A}}_1}{\bar{\mathbf{A}}_1^H \mathbf{R}_v^{-1}\bar{\mathbf{A}}_1}$$

Requires:

- $\mathbf{R}_v$: noise covariance matrix (estimate or model)
- $\bar{\mathbf{A}}_0$ and $\bar{\mathbf{A}}_1$: relative transfer function (RTF) of target speaker for left and right hearing aid
  - \+ preserves binaural cues of target speaker
  - \- distorts interaural coherence (IC) of noise component

- **BMVDR-N beamformer** [1, 3]: also preserve portion of noise component

$$\min_{\mathbf{W}_0} \mathcal{E}\{|\mathbf{W}_0^H \mathbf{V} - \eta_0 V_0|^2\} \quad \text{subject to} \quad \mathbf{W}_0^H \bar{\mathbf{A}}_0 = 1$$

$$\min_{\mathbf{W}_1} \mathcal{E}\{|\mathbf{W}_1^H \mathbf{V} - \eta_1 V_1|^2\} \quad \text{subject to} \quad \mathbf{W}_1^H \bar{\mathbf{A}}_1 = 1$$

$$\mathbf{W}_{\text{BMVDR-N},0} = (1-\eta_0)\mathbf{W}_{\text{BMVDR},0} + \eta_0 \mathbf{e}_0$$

$$\mathbf{W}_{\text{BMVDR-N},1} = (1-\eta_1)\mathbf{W}_{\text{BMVDR},1} + \eta_1 \mathbf{e}_1$$

- **η_0 and η_1: frequency-dependent trade-off parameters between noise reduction and binaural cue preservation of background noise**

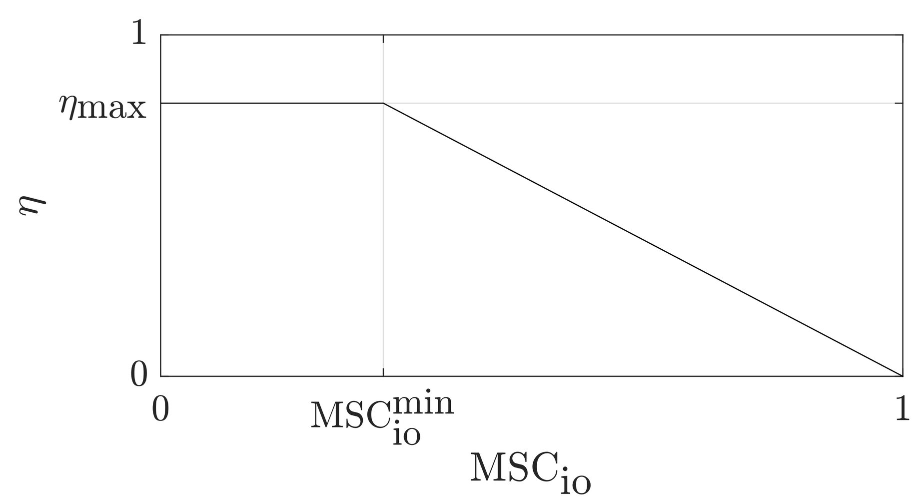
  $\eta_0 = \eta_1 = 1$: perfect binaural cue preservation, no noise reduction

  $\eta_0 = \eta_1 = 0$: maximum noise reduction, no binaural cue preservation

Trade-off Parameters

- **Fixed broadband** values (e.g. $\eta_0 = \eta_1 = 0.2 \ldots 0.3$)
- **Frequency-dependent** values, based on **IC discrimination ability** of human auditory system [3]
- **Frequency-dependent** values, based on **input/output SNR** [4]
  - large SNR: more important to keep maximum noise reduction (BMVDR)
  - low SNR: more important to preserve binaural cues (scaled input signals)
- **Frequency-dependent continuous function**, based on magnitude squared coherence (MSC) between noisy reference microphone signals and output signals of BMVDR beamformer

$$\text{MSC}_{\text{io},0} = \frac{|\mathcal{E}\{Y_0 Z_0^*\}|^2}{\mathcal{E}\{|Y_0|^2\}\,\mathcal{E}\{|Z_0|^2\}}, \quad \text{MSC}_{\text{io},1} = \frac{|\mathcal{E}\{Y_1 Z_1^*\}|^2}{\mathcal{E}\{|Y_1|^2\}\,\mathcal{E}\{|Z_1|^2\}},$$



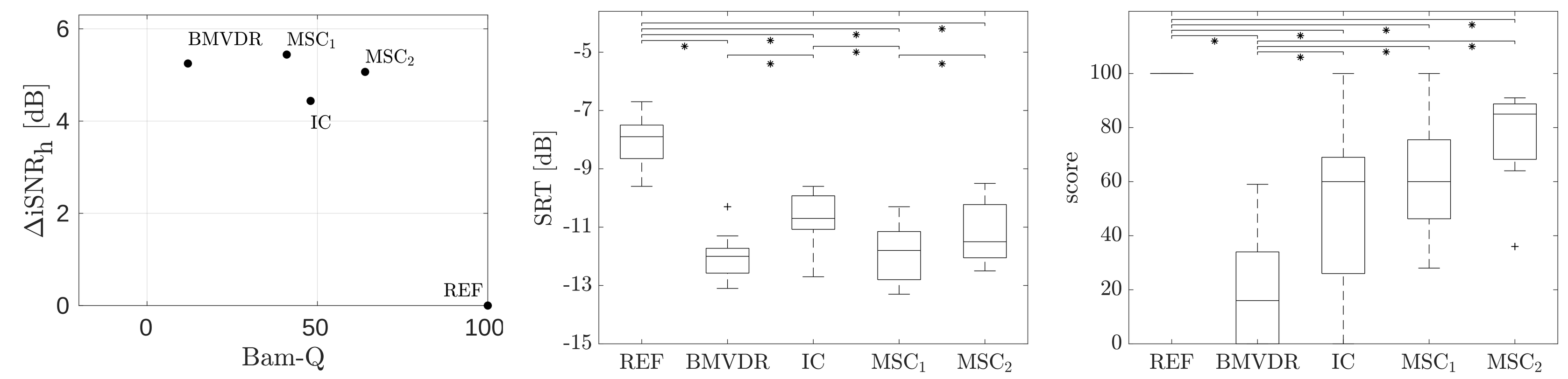
Characteristic curve between MSC_{io} and trade-off parameter η.

MSC-based algorithms:

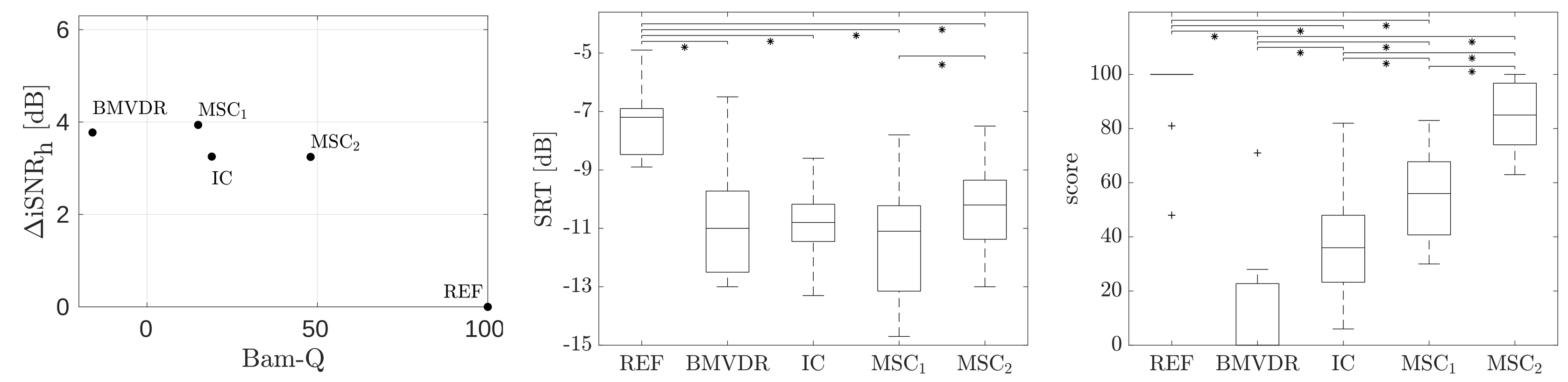
- MSC_1 with $\eta_{max} = 0.7$ and $\text{MSC}_{\text{io}}^{\text{min}} = 0$
- MSC_2 with $\eta_{max} = 1$ and $\text{MSC}_{\text{io}}^{\text{min}} = 0.1$

Experimental Results

- **Acoustic scenario**
  - Measured impulse responses from a binaural hearing aid (2 microphones each) in reverberant cafeteria ($T_{60} \approx 1250$ ms) [5]
  - Target speaker at 0° (scenario 1) and at -35° (scenario 2)
  - Ambient noise (babble noise, clacking plates) recorded in the same cafeteria
- **Algorithm implementation**
  - STFT framework: frame length 30 ms, frame shift 15 ms, f_s =16 kHz
  - Noise covariance matrix $\mathbf{R}_v$: diffuse noise assumption
  - RTF vectors: calculated from anechoic HRIRs, no DOA estimation errors
  - Recursive smoothing: time constant for MSC 20 ms
- **Objective performance measures**
  - Intelligibility-weighted hybrid SNR, taking into acount better ear glimpsing
  - Bam-Q: spatial impression of acoustic scene [6]
- **Subjective listening test**
  - N = 11 normal-hearing subjects
  - Speech reception threshold (SRT): Oldenburg sentence test
  - Spatial quality: MUSHRA using reference microphone signals as hidden reference



(a) scenario 1 (0°) (b) scenario 1 (0°) (c) scenario 1 (0°)



(d) scenario 2 (-35°) (e) scenario 2 (-35°) (f) scenario 2 (-35°)

- hybrid SNR and Bam-Q predict subjective results rather well
- **BMVDR**: SRT improvement of about 4 dB, poor spatial quality
- **BMVDR-IC**: similar SRT improvement as BMVDR, significantly improved spatial quality compared to BMVDR
- **BMVDR-MSC: similar SRT improvement as BMVDR, significantly improved spatial quality compared to BMVDR and BMVDR-IC (scenario 2)**

References

[1] B. Cornelis, S. Doclo, T. Van den Bogaert, J. Wouters, and M. Moonen. Theoretical analysis of binaural multi-microphone noise reduction techniques. *IEEE Trans. on Audio, Speech and Language Processing*, 18(2):342–355, Feb. 2010.

[2] S. Doclo, S. Gannot, D. Marquardt, and E. Hadad. Binaural Speech Processing with Application to Hearing Devices. In *Audio Source Separation and Speech Enhancement*, chapter 18. Wiley, 2018.

[3] D. Marquardt and 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*, 26(7):1257–1270, Jul 2018.

[4] J. Thiemann, M. Müller, D. Marquardt, S. Doclo, and 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(1):12, Feb. 2016.

[5] H. Kayser, S. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing*, 2009:10, 2009.

[6] J. Fleßner, S. Ewert, B. Kollmeier, and R. Huber. Quality assessment of multi-channel audio processing schemes based on a binaural auditory model. *in Proc., IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2014.