# Overview of acoustic signal processing research

Prof. Dr. Simon Doclo
University of Oldenburg
Dept. of Medical Physics and Acoustics, Cluster of Excellence Hearing4All

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

# Hearing research in Oldenburg

Carl von Ossietzky Universität Oldenburg

since 1993

**Research Groups Medical Physics, Acoustics, Signal Processing, Machine Learning**

- Basic research
- Education

8 Professors
20+ Postdocs
50+ PhD students

Hörzentrum Oldenburg – Institut an der Universität Oldenburg und am Evangelischen Krankenhaus Oldenburg

since 1996

**Hörzentrum GmbH**

- Market and trend research
- Audiological consulting
- Evaluation studies

Jade Hochschule – Wilhelmshaven Oldenburg Elsfleth

since 2000

**Institute of Hearing technology and Audiology**

- Education
- Application-oriented research

HörTech

since 2001

**Centre of Compentence Hörtech gGmbH**

- Product development
- Application-oriented research (hearing devices)

Fraunhofer IDMT

since 2008

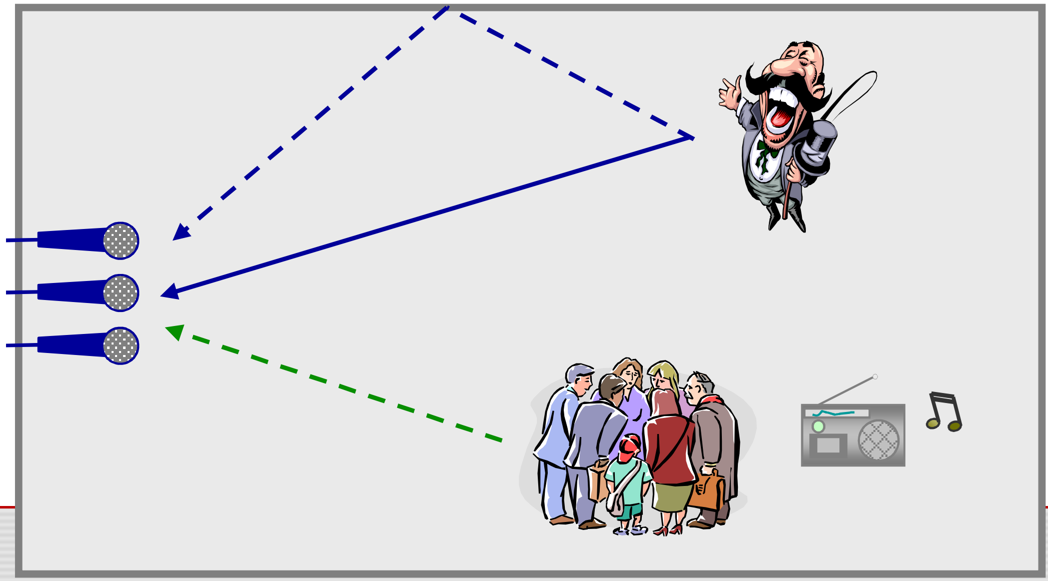
**Project group Hearing, Speech and Audio Technology**

- Application-oriented research (consumer electronics)

In total about 250 researchers in these institutes

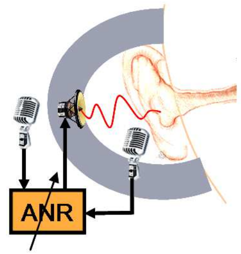
# Signal Processing Group

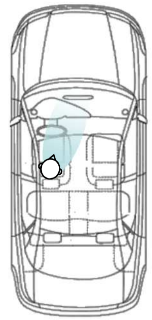
- Research, development and implementation of signal processing **algorithms** for acoustical and biomedical systems
- **Speech acquisition in adverse acoustic environments**
  - Signal enhancement
    - *noise reduction, dereverberation, blind source separation*
  - Microphone array processing
    - *adaptive beamforming, source localization*
  - Computational auditory scene analysis, sound classification
  - Acoustic echo cancellation and acoustic feedback cancellation



# Signal Processing Group

- Research, development and implementation of signal processing **algorithms** for acoustical and biomedical systems
- **Speech acquisition in adverse acoustic environments**
  - Signal enhancement
    - *noise reduction, dereverberation, blind source separation*
  - Microphone array processing
    - *adaptive beamforming, source localization*
  - Computational auditory scene analysis, sound classification
  - Acoustic echo cancellation and acoustic feedback cancellation

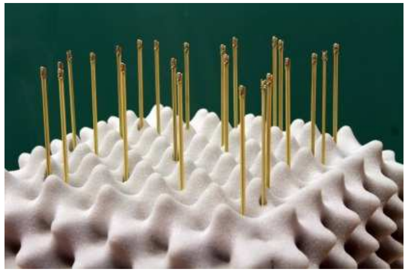


- **Sound reproduction**
  - Loudspeaker array processing
  - Active noise reduction
- **Applications:** hearing aids, cochlear implants, headsets, speech communication systems (mobile phone, voice-controlled systems)

# Current research topics

- **Speech enhancement for ear-mounted communication devices**
  - **Binaural noise reduction**, aiming to preserve spatial impression of acoustic scene (binaural cues)
  - Open-fitting hearing devices: **feedback cancellation** and **active noise control** (acoustically transparent earpiece)
  - EEG-based **auditory attention decoding** for steering beamformers



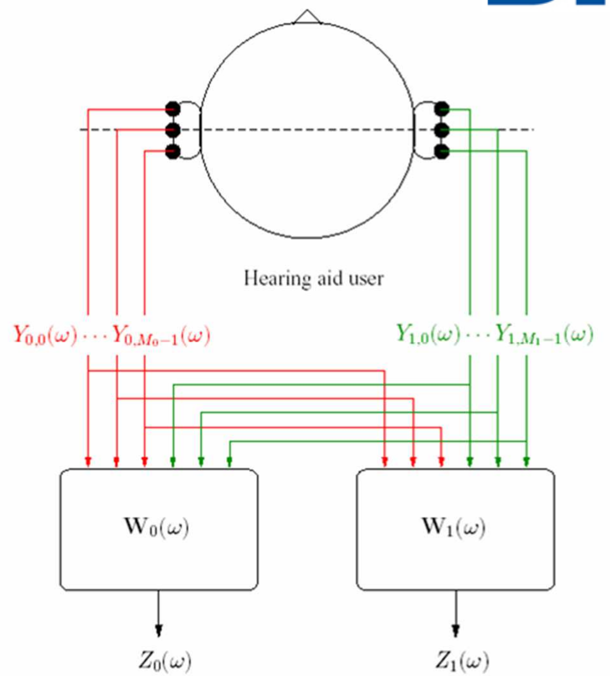
- **MIMO acoustics**
  - **Beamformer design** (e.g., virtual artificial head)
  - **Dereverberation and noise reduction** (spectral enhancement, multi-channel equalization, blind probabilistic model-based)
  - **Acoustic sensor networks** (spatially distributed microphones, sampling rate offset estimation, distributed processing)
  - **Computational acoustic scene analysis** (CASA)

# Binaural noise reduction

# Binaural noise reduction

VolkswagenStiftung
DFG

- **Problem**
  - Hearing impaired suffer from loss of speech understanding in noisy environments
  - Improvement of speech intelligibility by noise reduction algorithms
- **Objectives**
  - Develop binaural noise reduction algorithms, avoiding signal distortions and preserving spatial awareness
- **Approaches**
  - Novel binaural algorithms, merging advantages of *spectral post-filtering* (preservation of cues) and *spatial processing* (no artefacts)
  - Incorporate psychoacoustic properties of the human auditory system in binaural noise reduction algorithms
  - Integration with CASA (scene analysis)

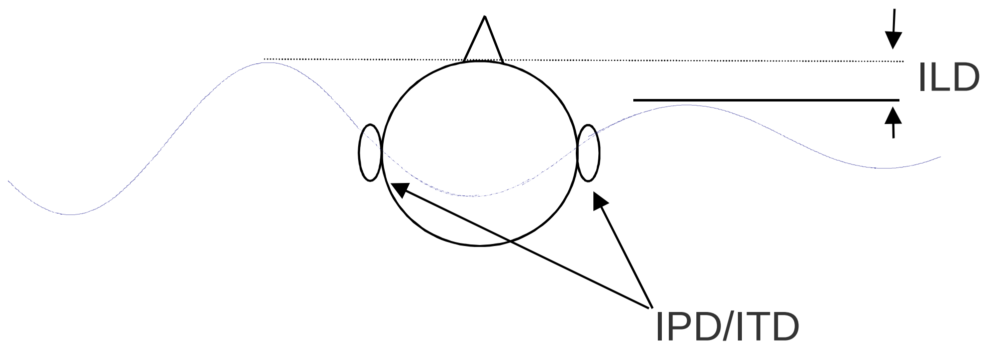



Daniel Marquardt

Dörte Fischer

# Binaural auditory cues

- **Interaural Time/Phase Difference (ITD/IPD)**
  **Interaural Level Difference (ILD)**
  **Interaural Coherence (IC)**
  - ITD: f < 1500 Hz, ILD: f > 2000 Hz
  - IC: describes spatial characteristics, e.g. perceived width, of diffuse noise, and determines when ITD/ILD cues are *reliable*
- Binaural cues, in addition to spectro-temporal cues, play an important role in auditory scene analysis (source segregation) and speech intelligibility



# Binaural auditory cues

- **Spatial release from masking (BMLD):**
  - *Localized noise source* : large effect for NH listeners (especially in free-field)
  - *Diffuse noise* : about 2-3 dB

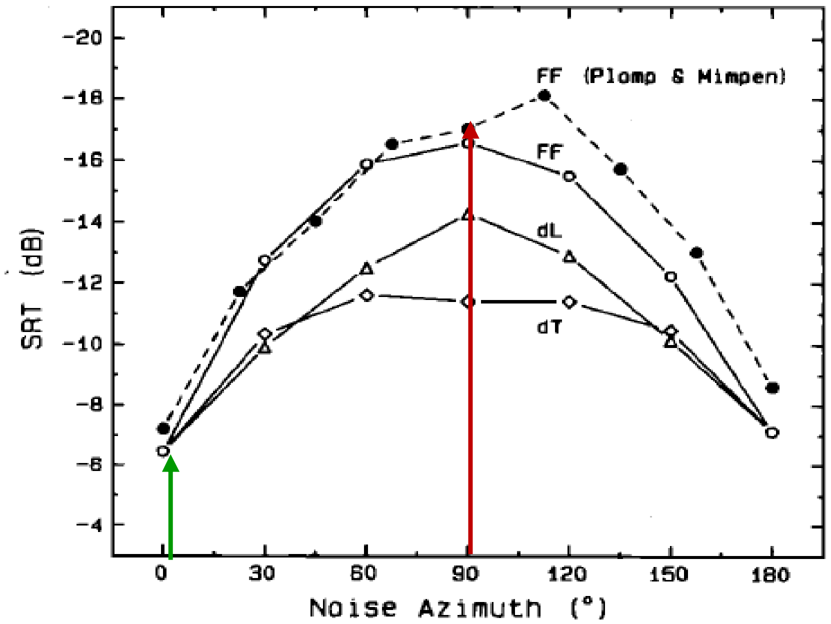
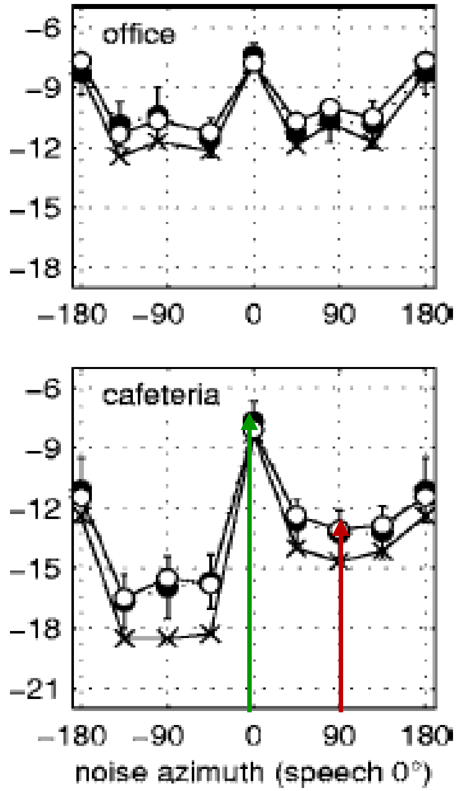



FIG. 5. Mean speech reception thresholds obtained in experiment I for three different noise types: FF (free field), dL (headshadow only), and dT (ITD only). The closed data points represent results of Plomp and Mimpen (1981) obtained in a free field.

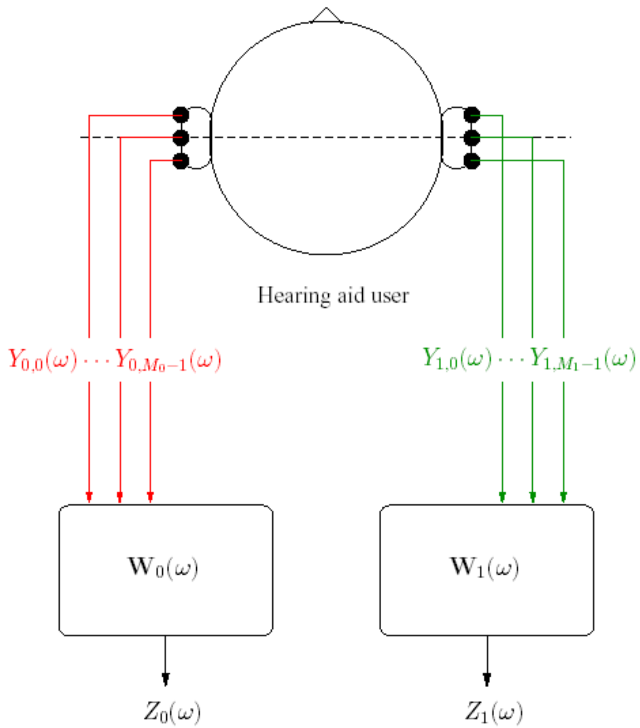
[Bronkhorst and Plomp, 1988]




[Beutelmann and Brand, 2006]

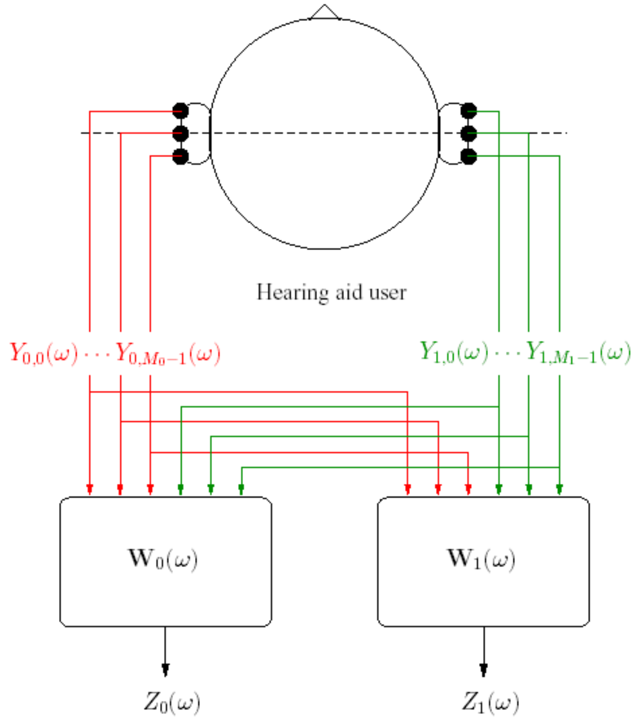
# Binaural noise reduction: Configuration

## Monaural/Bilateral system




⊖ **Independent** left/right processing:
- No cooperation (e.g. different environment classification)
- preservation of binaural cues ?

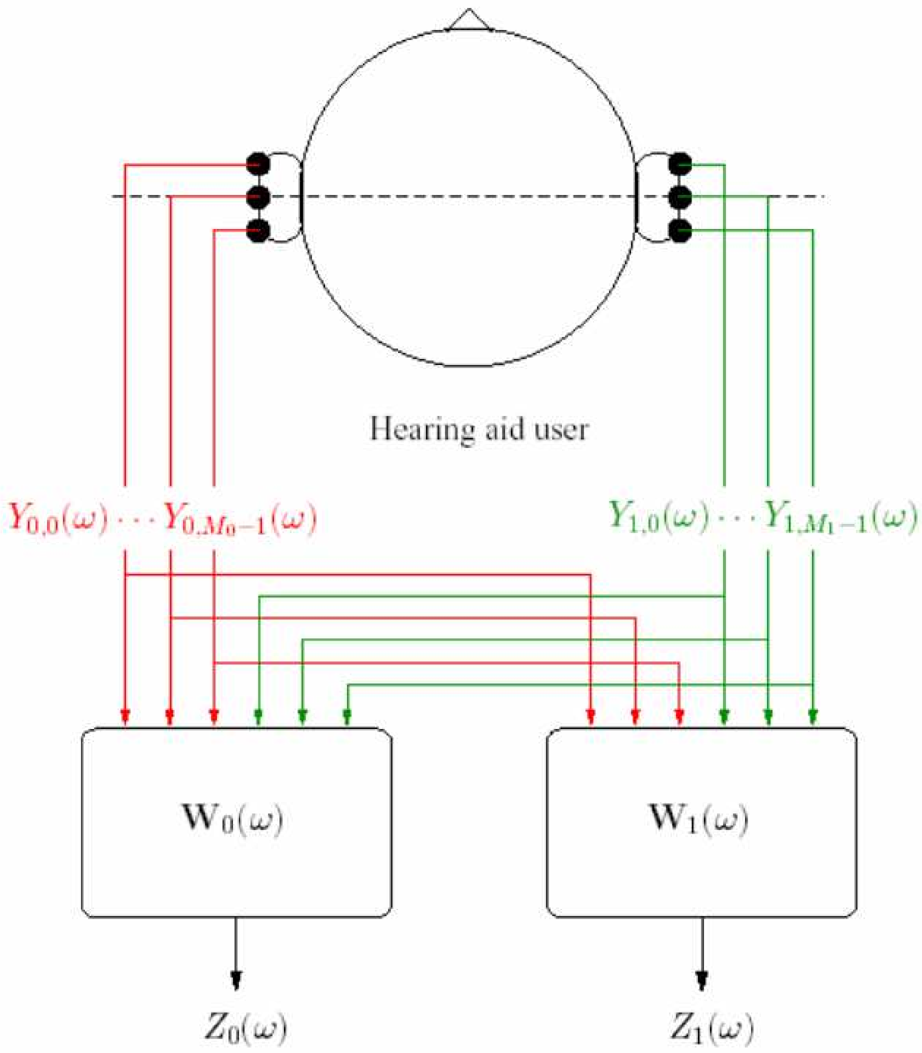
## Binaural system




⊕ Exchange of:
- **parameters** (volume, environment)
- **signals** (cooperative processing for noise reduction, feedback, ...)

⊖ Need for wireless binaural link

# Binaural noise reduction: Configuration




- Binaural hearing aid configuration:
  - Two hearing aids with in total *M* microphones
  - All microphone signals **Y** are assumed to be available at both hearing aids (perfect wireless link)
- Apply a filter $\mathbf{W}_0$ and $\mathbf{W}_1$ at the left and the right hearing aid, generating binaural output signals $Z_0$ and $Z_1$

$$\boxed{Z_0(\omega) = \mathbf{W}_0^H(\omega)\mathbf{Y}(\omega), \quad Z_1(\omega) = \mathbf{W}_1^H(\omega)\mathbf{Y}(\omega)}$$

# Binaural noise reduction: Acoustic scenario

- The microphone signals **Y** are composed of
  - (desired) speech component $\mathbf{X} = S_d\mathbf{A}$
  - (undesired) directional interference component $\mathbf{U} = S_u\mathbf{B}$
  - (undesired) background noise component **N**

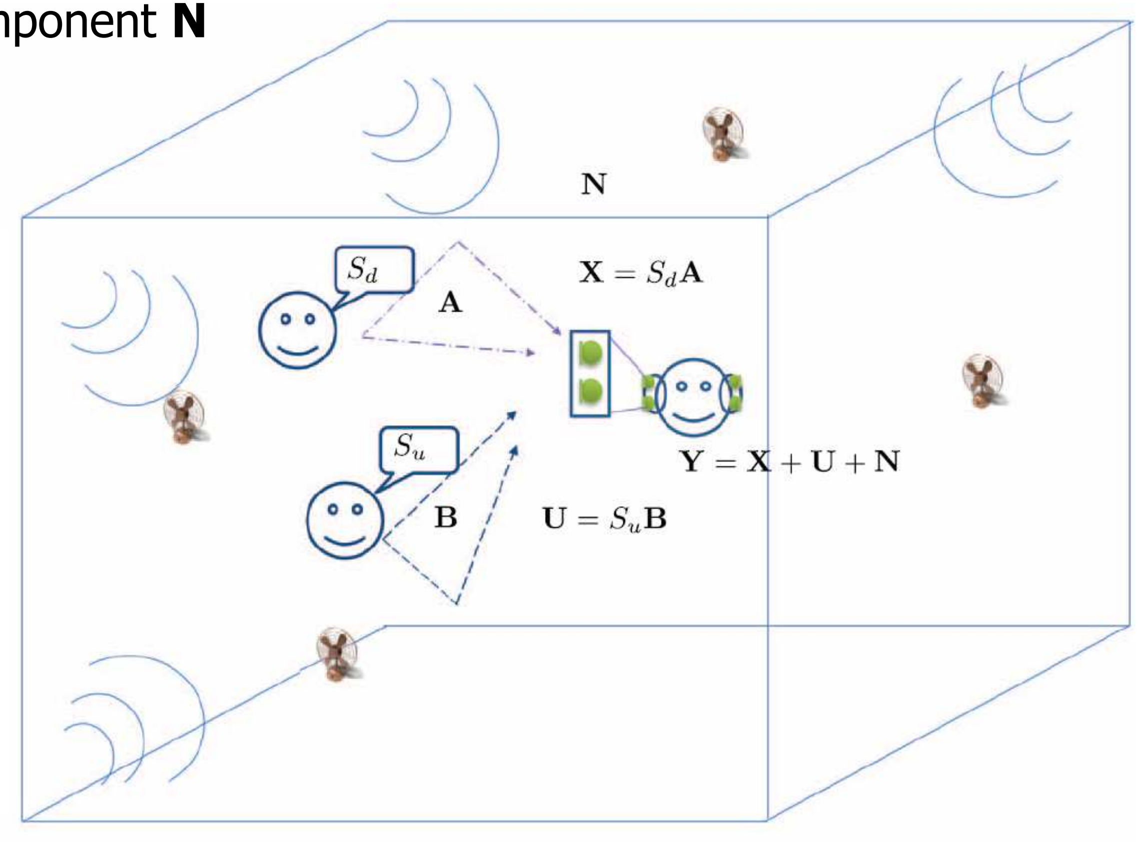
Acoustic Transfer Functions (ATFs)

- Correlation matrices:

$$\mathbf{R}_y = \mathbf{R}_x + \underbrace{\mathbf{R}_u + \mathbf{R}_n}_{\mathbf{R}_v}$$

$$\mathbf{R}_x = \mathcal{E}\left\{\mathbf{X}\mathbf{X}^H\right\} = P_s\mathbf{A}\mathbf{A}^H$$
$$\mathbf{R}_u = \mathcal{E}\left\{\mathbf{U}\mathbf{U}^H\right\} = P_u\mathbf{B}\mathbf{B}^H$$
$$\mathbf{R}_n = \mathcal{E}\left\{\mathbf{N}\mathbf{N}^H\right\},$$

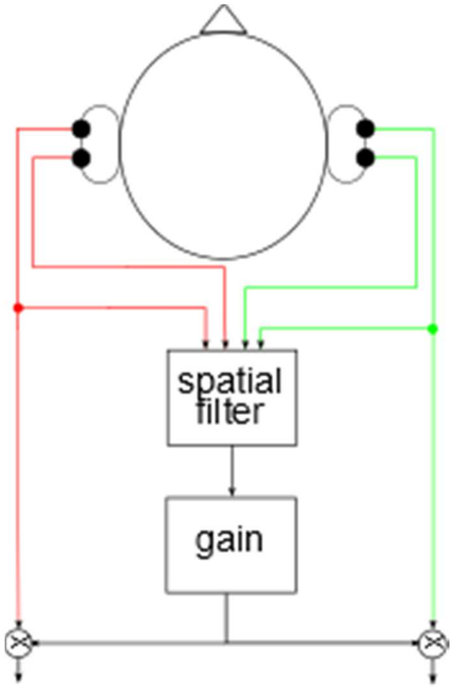
- All **binaural cues** can be written in terms of these matrices



# Binaural noise reduction: Two main paradigms

## Spectral post-filtering (based on multi-microphone noise reduction)

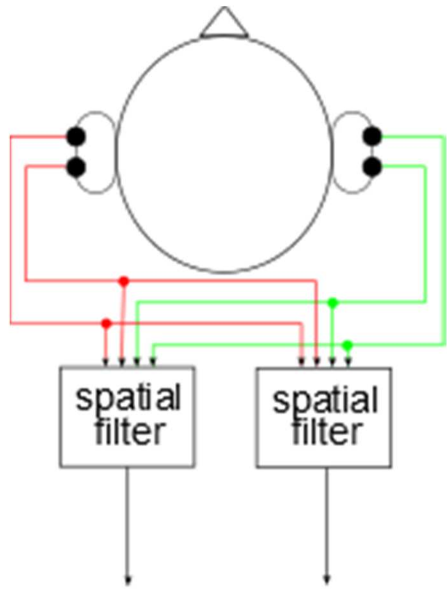
[Doerbecker 1996, Wittkop 2003, Lotter 2006, Rohdenburg 2007, Grimm 2009, Reindl 2012]




- ⊕ Binaural cue preservation
- ⊖ Possible single-channel artifacts

## Binaural multi-microphone noise reduction techniques

[Welker 1997, Doclo 2010, Cornelis 2012, Hadad 2014-2016, Marquardt 2014-2016]




- ⊕ Larger noise reduction performance
- ⊕ Merge spatial and spectral post-filtering
- ⊖ Binaural cue preservation not guaranteed

# Binaural MVDR and MWF

## Minimum-Variance-Distortionless-Response (MVDR) beamformer

**Goal:** minimize output noise power without distorting speech component in reference microphone signals

$$\min_{\mathbf{W}_0} \underbrace{\mathbf{W}_0^H \mathbf{R}_\mathrm{v} \mathbf{W}_0}_{\text{noise reduction}} \quad \text{subject to} \quad \underbrace{\mathbf{W}_0^H \mathbf{A} = A_0}_{\text{distortionless constraint}}$$

$$\min_{\mathbf{W}_1} \mathbf{W}_1^H \mathbf{R}_\mathrm{v} \mathbf{W}_1 \quad \text{subject to} \quad \mathbf{W}_1^H \mathbf{A} = A_1$$

$$\mathbf{W}_{\mathrm{MVDR},0} = \frac{\mathbf{R}_\mathrm{v}^{-1}\mathbf{A}}{\mathbf{A}^H \mathbf{R}_\mathrm{v}^{-1}\mathbf{A}} A_0^*$$

$$\mathbf{W}_{\mathrm{MVDR},1} = \frac{\mathbf{R}_\mathrm{v}^{-1}\mathbf{A}}{\mathbf{A}^H \mathbf{R}_\mathrm{v}^{-1}\mathbf{A}} A_1^*$$

## Multi-channel Wiener Filter (MWF)

**Goal:** estimate speech component in reference microphone signals + trade off noise reduction and speech distortion

$$J_{\mathrm{MWF}}(\mathbf{W}) = \mathcal{E}\left\{ \underbrace{\left\| \begin{bmatrix} X_0 - \mathbf{W}_0^H \mathbf{X} \\ X_1 - \mathbf{W}_1^H \mathbf{X} \end{bmatrix} \right\|^2}_{\text{speech distortion}} + \mu \underbrace{\left\| \begin{bmatrix} \mathbf{W}_0^H \mathbf{V} \\ \mathbf{W}_1^H \mathbf{V} \end{bmatrix} \right\|^2}_{\text{noise reduction}} \right\}$$

$$\mathbf{W}_{\mathrm{MWF},0} = (\mathbf{R}_x + \mu \mathbf{R}_v)^{-1} \mathbf{r}_{\mathrm{x},0}$$

$$\mathbf{W}_{\mathrm{MWF},1} = (\mathbf{R}_x + \mu \mathbf{R}_v)^{-1} \mathbf{r}_{\mathrm{x},1}$$

# Binaural MVDR and MWF

## Minimum-Variance-Distortionless-Response (MVDR) beamformer

**Goal:** minimize output noise power without distorting speech component in reference microphone signals

$$\min_{\mathbf{W}_0} \mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 \quad \text{subject to} \quad \mathbf{W}_0^H \mathbf{A} = A_0$$

$$\min_{\mathbf{W}_1} \underbrace{\mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}_{\text{noise reduction}} \quad \text{subject to} \quad \underbrace{\mathbf{W}_1^H \mathbf{A} = A_1}_{\text{distortionless constraint}}$$

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

## Multi-channel Wiener Filter (MWF)

**Goal:** estimate speech component in reference microphone signals + trade off noise reduction and speech distortion

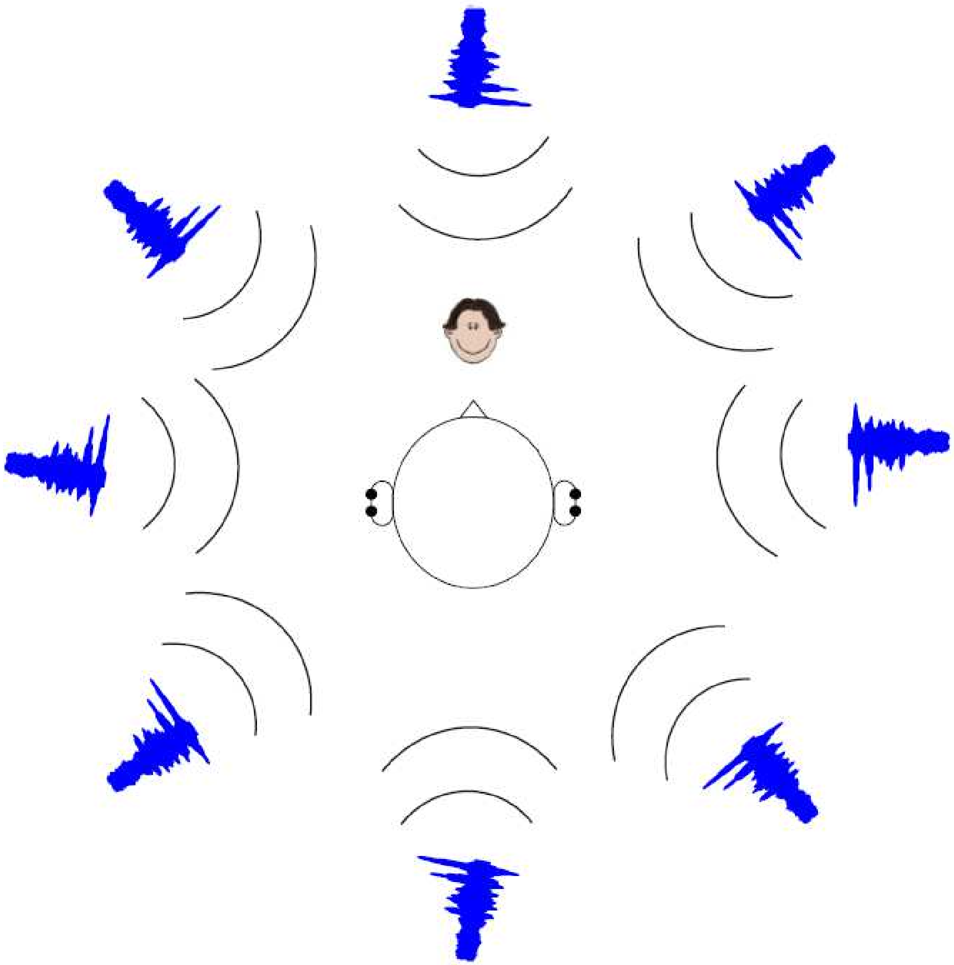
$$J_{\text{MWF}}(\mathbf{W}) = \mathcal{E}\left\{ \underbrace{\left\| \begin{bmatrix} X_0 - \mathbf{W}_0^H \mathbf{X} \\ X_1 - \mathbf{W}_1^H \mathbf{X} \end{bmatrix} \right\|^2}_{\text{speech distortion}} + \mu \underbrace{\left\| \begin{bmatrix} \mathbf{W}_0^H \mathbf{V} \\ \mathbf{W}_1^H \mathbf{V} \end{bmatrix} \right\|^2}_{\text{noise reduction}} \right\}$$

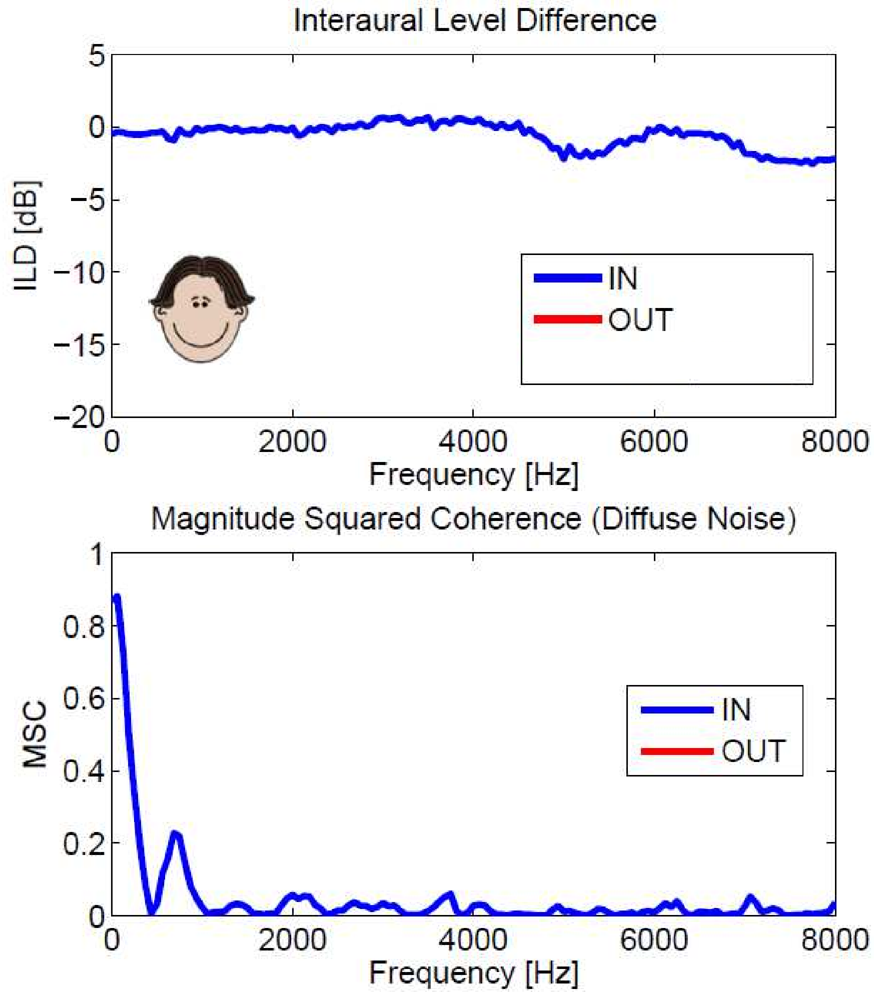
**Requires** estimate of speech and noise covariance matrices, e.g. based on VAD

Can be decomposed as binaural MVDR beamformer and spectral postfilter

**Good noise reduction performance, what about binaural cues ?**

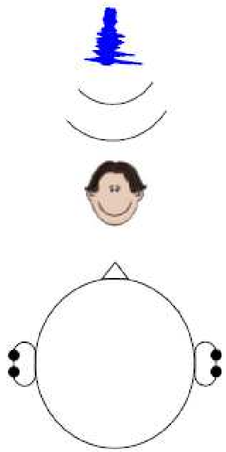
# Binaural MVDR/MWF: binaural cues

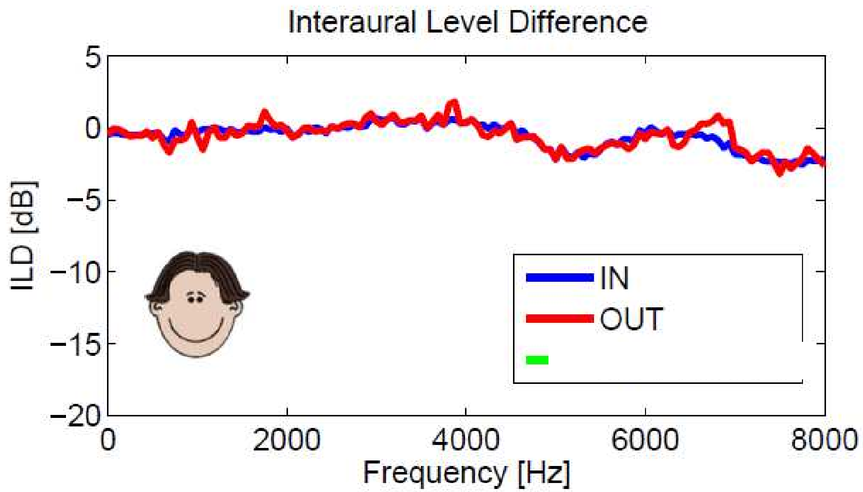


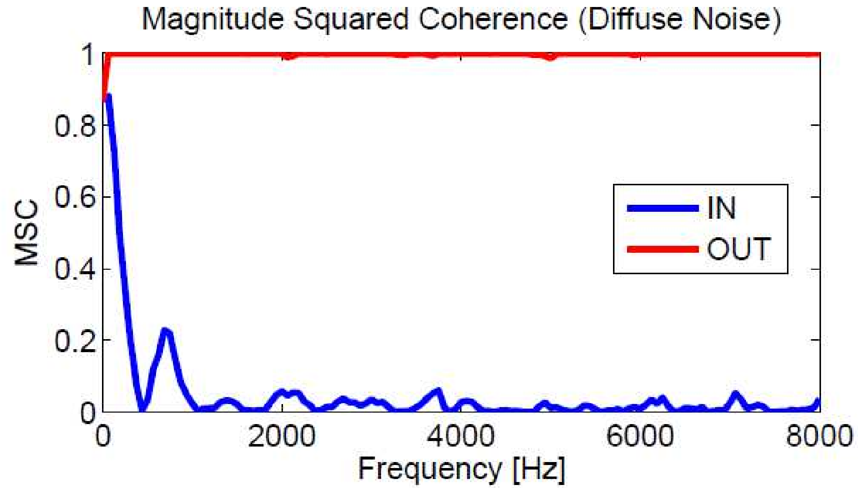



*Note: MSC = Magnitude Squared Coherence*

# Binaural MVDR/MWF: binaural cues






**Binaural cues for residual noise/interference in binaural MVDR/MWF not preserved**

# Binaural MWF: Extensions for diffuse noise

Binaural MWF

- ⊕ SNR improvement
- ⊕ Binaural cues of speech source
- ⊖ Binaural cues of noise

Interaural coherence preservation (MWF-IC)

$$J_{MWF-IC}(\mathbf{W}) = J_{MWF}(\mathbf{W}) + \lambda \left| \frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_1}{\sqrt{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}} - IC_v^{des} \right|^2$$

⊖ No closed-form solution, iterative optimization procedures required

Partial noise estimation (MWF-N)

$$J_{\mathrm{MWF-N}}(\mathbf{W}) = \mathcal{E}\left\{ \left\| \begin{bmatrix} X_0 - \mathbf{W}_0^H \mathbf{X} \\ X_1 - \mathbf{W}_1^H \mathbf{X} \end{bmatrix} \right\|^2 + \mu \left\| \begin{bmatrix} \eta V_0 - \mathbf{W}_0^H \mathbf{V} \\ \eta V_1 - \mathbf{W}_1^H \mathbf{V} \end{bmatrix} \right\|^2 \right\}$$

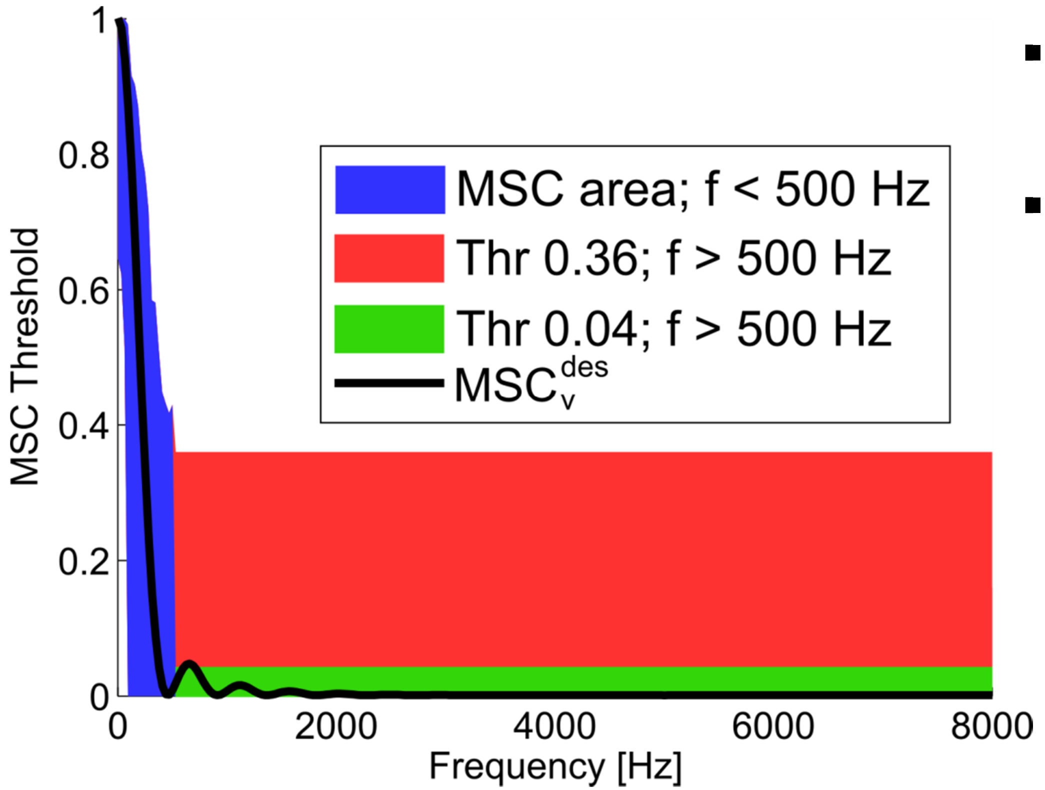
⊕ Closed-form solution (mixing with reference microphone signals)

⊜ **Trade-off** between SNR improvement and binaural cue preservation, depending on **parameters** (η and λ)

# Binaural MWF: Extensions for diffuse noise

## Determine (frequency-dependent) trade-off parameters based on psycho-acoustic criteria

- Amount of IC preservation based on subjective listening experiments evaluating the IC discrimination abilities of the human auditory system

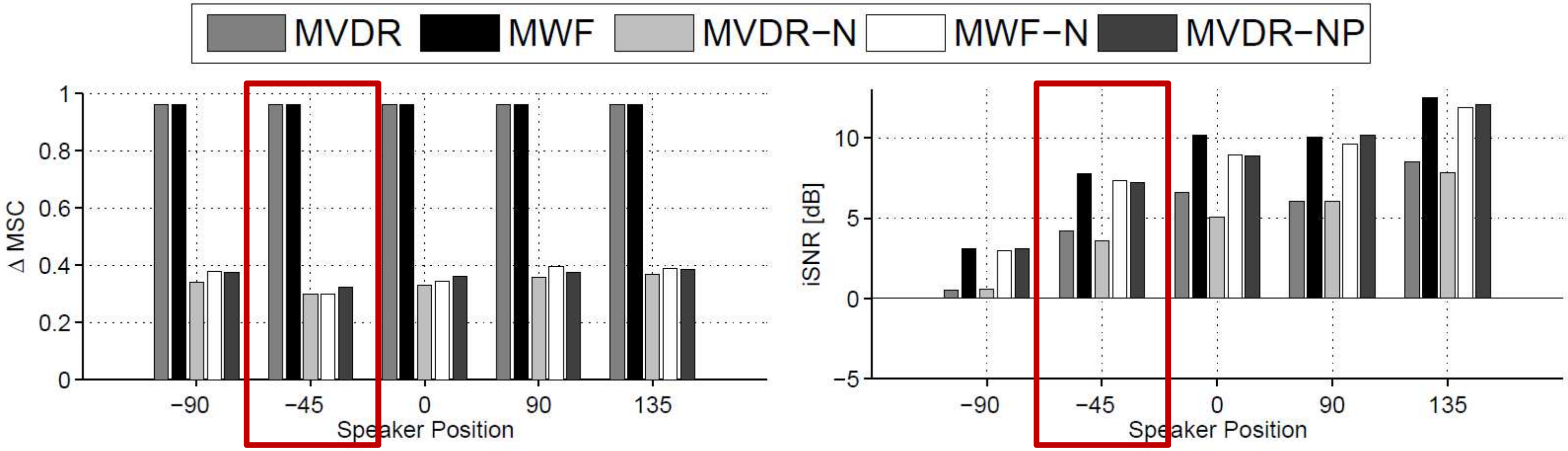



- IC discrimination ability depends on magnitude of reference IC
- **Boundaries on Magnitude Squared Coherence** ($MSC=|IC|^2$) :
  - For f < 500 Hz ("large" IC): frequency-dependent MSC boundaries (**blue**)
  - For f > 500 Hz ("small" IC): fixed MSC boundary, e.g. 0.36 (**red**) or 0.04 (**green**)

[Marquardt 2014/2015]

# Binaural MWF: Extensions for diffuse noise

## ❑ Instrumental evaluation / sound samples




(a) $\Delta\mathrm{MSC_v}$ (b) $\mathrm{iSNR_{be}^{out}}$

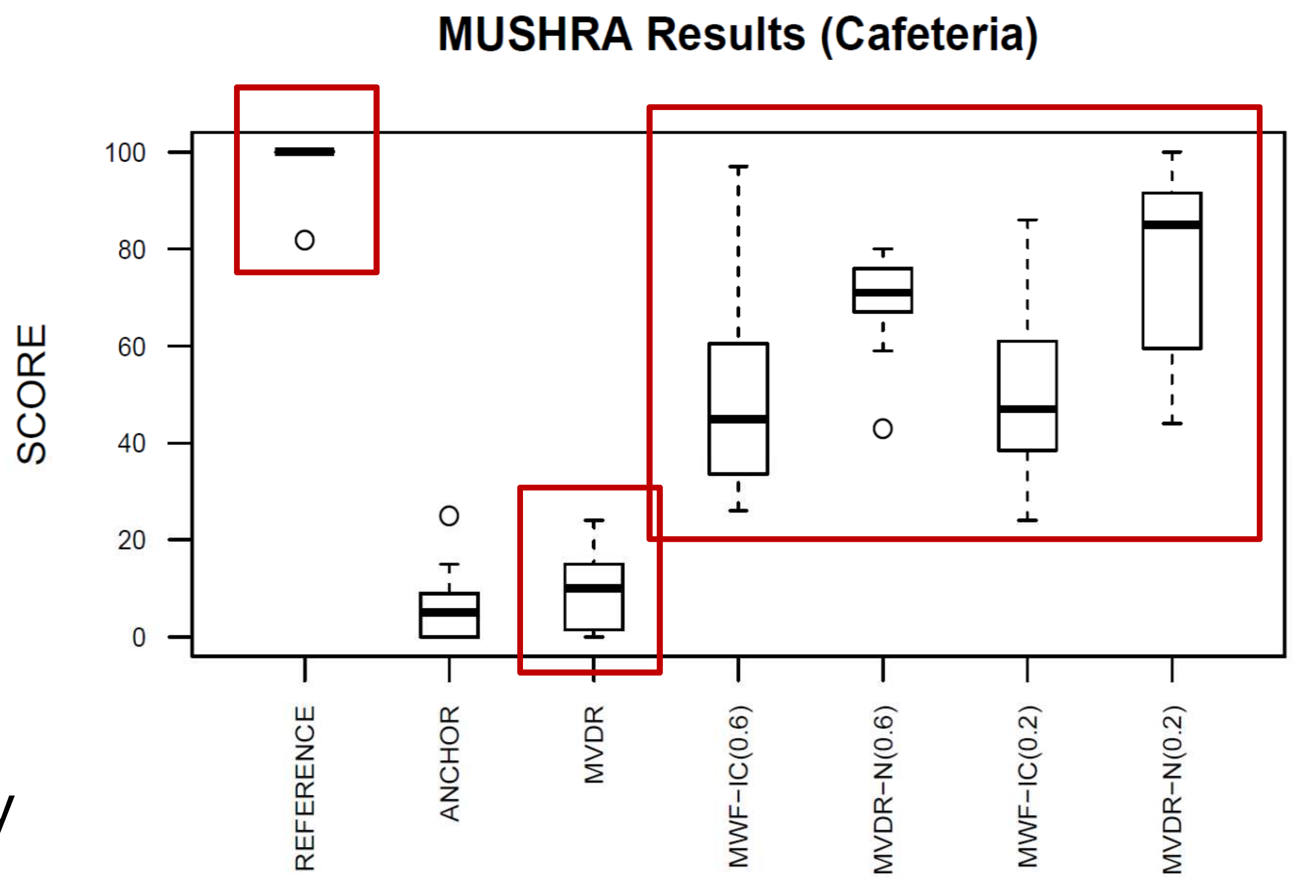
| Input | MVDR | MWF | MVDR-N | MWF-N | MVDR-NP |
|---|---|---|---|---|---|
| | | | | | |

Cafeteria with recorded ambient noise, speaker at -45°, 0 dB input iSNR (left hearing aid)
MVDR: anechoic ATF, DOA known, spatial coherence matrix calculated from anechoic ATFs / MWF = MVDR + postfilter (SPP-based)

**Does binaural unmasking compensate for SNR decrease ?**

# Evaluation: Spatial quality (MUSHRA)

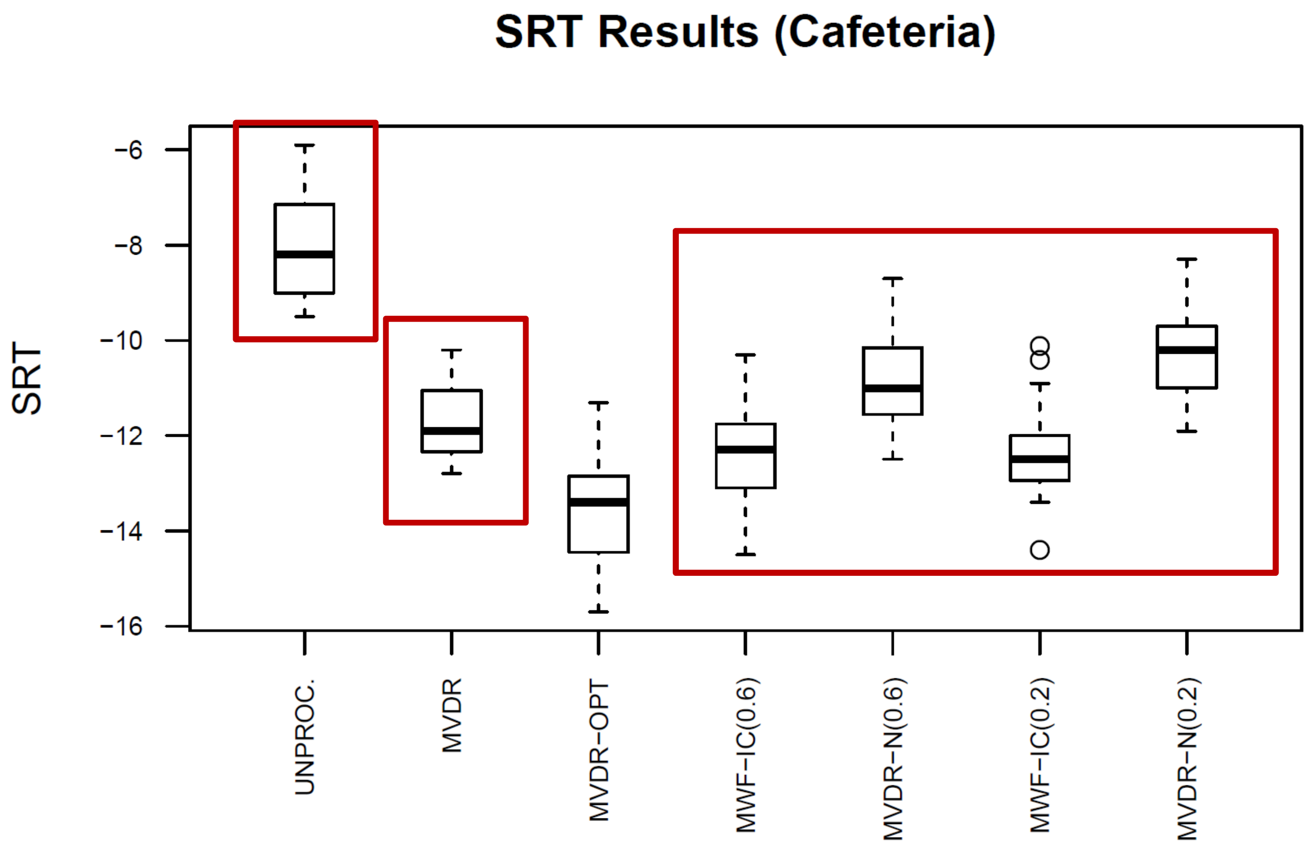
- Evaluate spatial difference between reference and output signal
- **MWF-IC and MVDR-N outperform MVDR**
  - MVDR-N shows better results than MWF-IC
  - Decreasing the MSC threshold slightly improves spatial quality




**Binaural cue preservation for diffuse noise improves spatial quality**

# Evaluation: Speech intelligibility (SRT)

- All algorithms show a highly significant SRT improvement
- The SRT results mainly reflect the SNR differences between algorithms: MWF-IC outperforms MVDR-N
- **No significant SRT difference between MVDR and MWF-IC**




**Binaural cue preservation for diffuse noise does not/hardly affect speech intelligibility**

# Binaural MVDR: Extensions for interfering source

Binaural MVDR

- ⊕ SNR improvement
- ⊕ Binaural cues of speech source
- ⊖ Binaural cues of interferer

Relative transfer function (BMVDR-RTF)

$$\min_{\mathbf{W}_0,\mathbf{W}_1} \left\{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1\right\}$$
$$\text{s.t. } \mathbf{W}_0^H \mathbf{A} = A_0, \mathbf{W}_1^H \mathbf{A} = A_1, \frac{\mathbf{W}_0^H \mathbf{B}}{\mathbf{W}_1^H \mathbf{B}} = \frac{B_0}{B_1}.$$

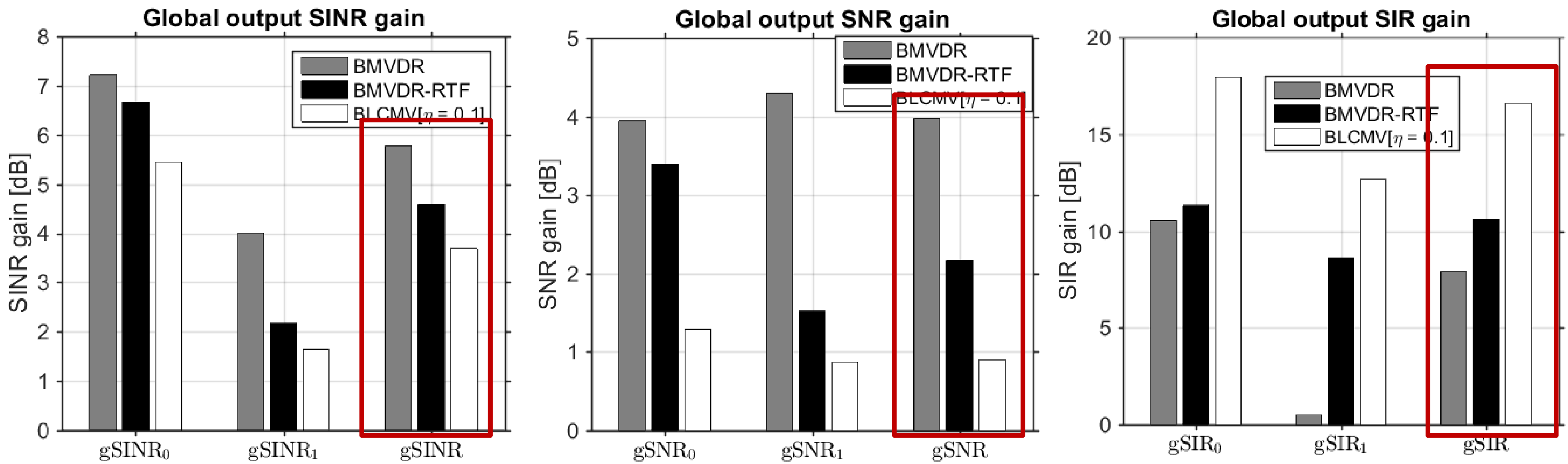
Interference rejection (BMVDR-IR)

$$\min_{\mathbf{W}_0} \left\{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0\right\} \text{ s.t. } \mathbf{W}_0^H \mathbf{A} = A_0, \mathbf{W}_0^H \mathbf{B} = \eta B_0$$
$$\min_{\mathbf{W}_1} \left\{\mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1\right\} \text{ s.t. } \mathbf{W}_1^H \mathbf{A} = A_1, \mathbf{W}_1^H \mathbf{B} = \eta B_1$$

- ⊕ Binaural cues of speech source **and** interfering source preserved
- ⊕ Also binaural MWF-based versions (incl. spectral filtering) can be derived
- ⊖ Background noise: MSC not exactly preserved, possible noise amplification

# Binaural MVDR: Extensions for interfering source

## ❑ Instrumental evaluation / sound samples




| Input | BMVDR | BMVDR-RTF | BMVDR-IR (η = 0.1) |
|---|---|---|---|
| | | | |

Cafeteria with recorded ambient noise, speaker at 0°, Interference at -45°, 0 dB input SIR and SNR (left hearing aid)
RTF calculated from correlation matrix (Rx and Ru), 3 microphones (2 left, 1 right)

[Hadad 2014/2015/2016, Marquardt 2014/2015]

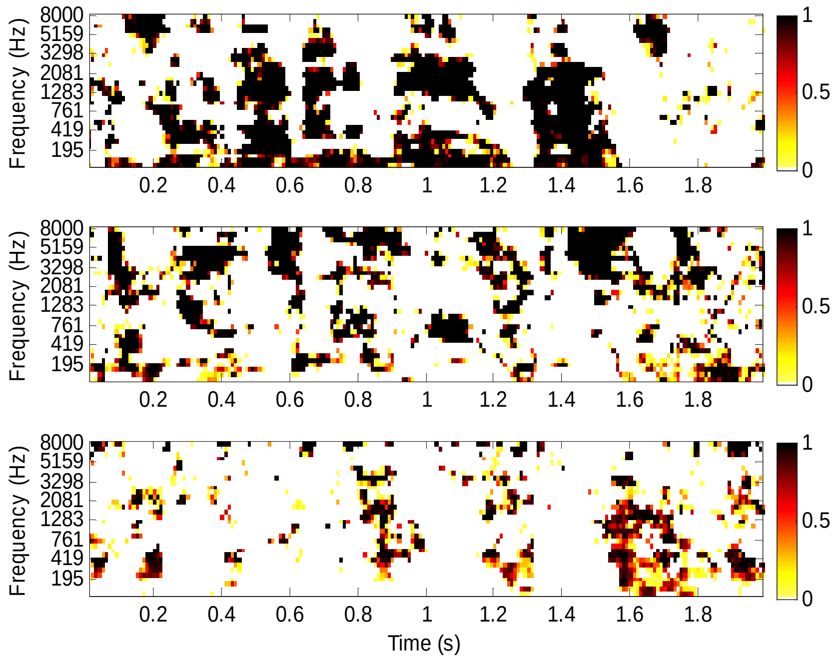
# Current/Future work

- Binaural noise reduction algorithms for **interfering sources** (BMVDR-IR, BMVDR-RTF):
  - Subjective evaluation (incl. binaural cue preservation) for HA/CI users
  - Robustness against RTF estimation errors

- **Mixed noise fields and time-varying scenarios**: incorporate computational acoustic scene analysis (CASA) into developed algorithms

- Extend algorithms to include **external microphones (acoustic sensor networks)**

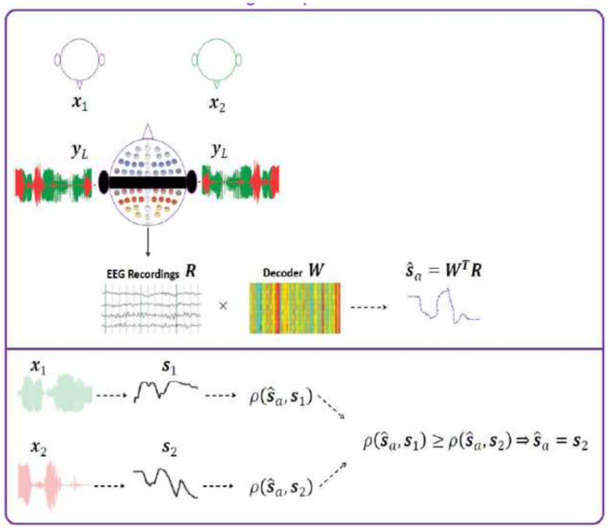


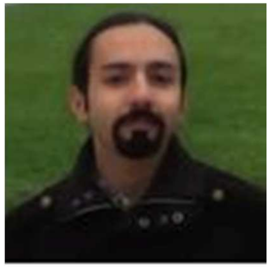



# Auditory attention decoding

Niedersächsisches Ministerium für Wissenschaft und Kultur

- **Problem**
  - Multi-microphone noise reduction in complex acoustic scenarios with interfering speaker(s)
  - Many algorithms rely on **pre-defined assumptions about target speaker** (e.g. direction / energy)
- **Objectives**
  - Use brain computer interface to **control multi-microphone noise reduction techniques**, to enhance target speaker to which user is attending
- **Approach**
  - Control of binaural noise reduction techniques through BCI (e.g. correlation of EEG and acoustical signals / features)
  - Investigate feedback/reinforcement mechanism by presenting enhanced source



Ali Aroudi

# Recent publications

- 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.
- E. Hadad, S. Doclo, S. Gannot, *The Binaural LCMV Beamformer and its Performance Analysis*, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 24, no. 3, pp. 543-558, Mar. 2016.
- E. Hadad, D. Marquardt, S. Doclo, S. Gannot, *Theoretical Analysis of Binaural Transfer Function MVDR Beamformers with Interference Cue Preservation Constraints*, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 23, no. 12, pp. 2449-2464, Dec. 2015.
- D. Marquardt, E. Hadad, S. Gannot, S. Doclo, *Theoretical Analysis of Linearly Constrained Multi-channel Wiener Filtering Algorithms for Combined Noise Reduction and Binaural Cue Preservation in Binaural Hearing Aids*, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 23, no. 12, pp. 2384-2397, Dec. 2015.
- R. Baumgärtel, M. Krawczyk-Becker, D. Marquardt, C. Völker, H. Hu, T. Herzke, G. Coleman, K. Adiloglu, S. Ernst, T. Gerkmann, S. Doclo, B. Kollmeier, V. Hohmann, M. Dietz, *Comparing binaural pre-processing strategies I: Instrumental evaluation*, *Trends in Hearing*, vol. 19, pp. 1-16, 2015.
- R. Baumgärtel, H. Hu, M. Krawczyk-Becker, D. Marquardt, T. Herzke, G. Coleman, K. Adiloglu, K. Bomke, K. Plotz, T. Gerkmann, S. Doclo, B. Kollmeier, V. Hohmann, M. Dietz, *Comparing binaural pre-processing strategies II: Speech intelligibility of bilateral cochlear implant users*, *Trends in Hearing*, vol. 19, pp. 1-18, 2015.

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

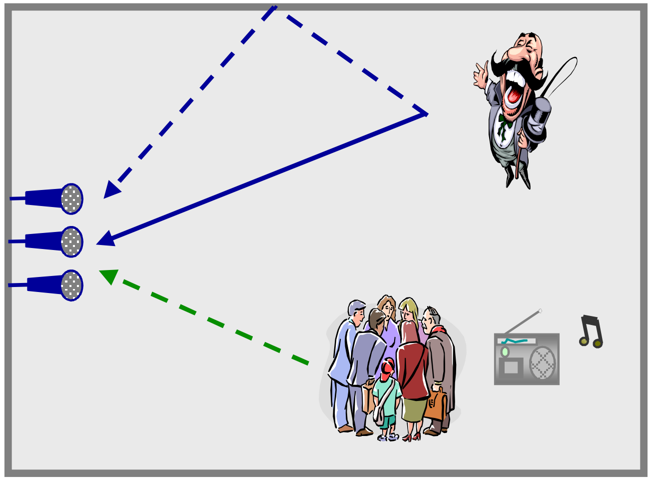
# Joint dereverberation and noise reduction

# Dereverberation and noise reduction

German-Israeli Foundation for Scientific Research and Development

DREAMS Dereverberation and Reverberation of Audio, Music, and Speech

- **Problem**
  - Noise and reverberation jointly present in typical acoustic environments
  - Speech quality and intelligibility degradation
  - Performance degradation of ASR systems
- **Objectives**
  - Develop single- and multi-channel joint dereverberation and noise reduction algorithms
  - Exploit knowledge or statistical models of room acoustics
- **Approaches**
  1. Single-microphone spectral enhancement (estimation of LRSV, inverse filtering)
  2. Robust multi-channel equalization
  3. Probabilistic estimation using statistical models of desired signal and reverberation



Ina Kodrasi

Ante Jukić

Benjamin Cauchi

# Signal model

- **Scenario:** speech source in noisy and reverberant environment, *M* microphones
- **Time-domain model:** “perfect” model

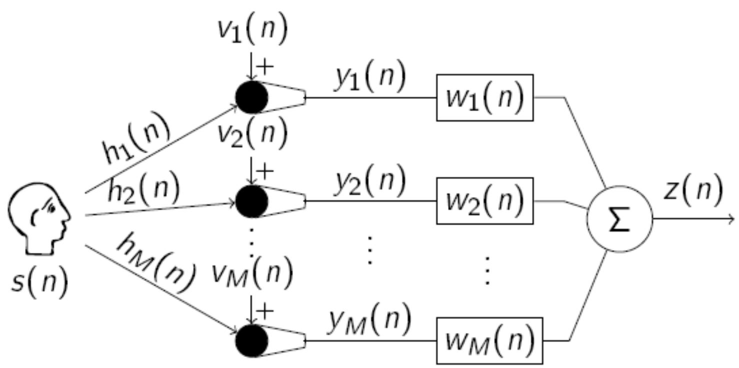
$$y_m(n) = x_m(n) + v_m(n) = s(n) * h_m(n) + v_m(n)$$

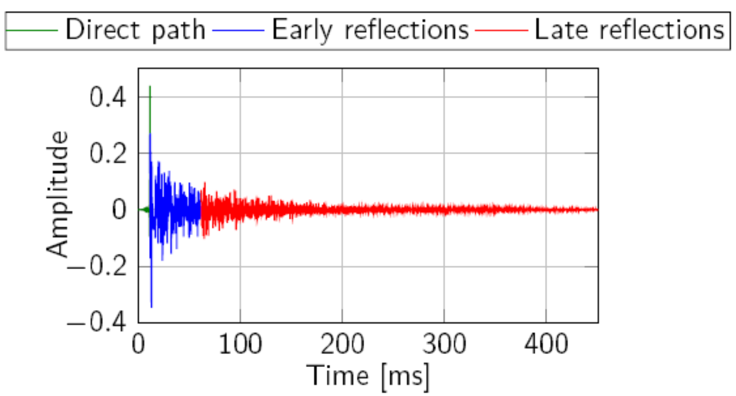
$h_m(n)$ = room impulse response (RIR), typically long and difficult to blindly estimate

- **STFT-domain model:** approximation of time-domain model

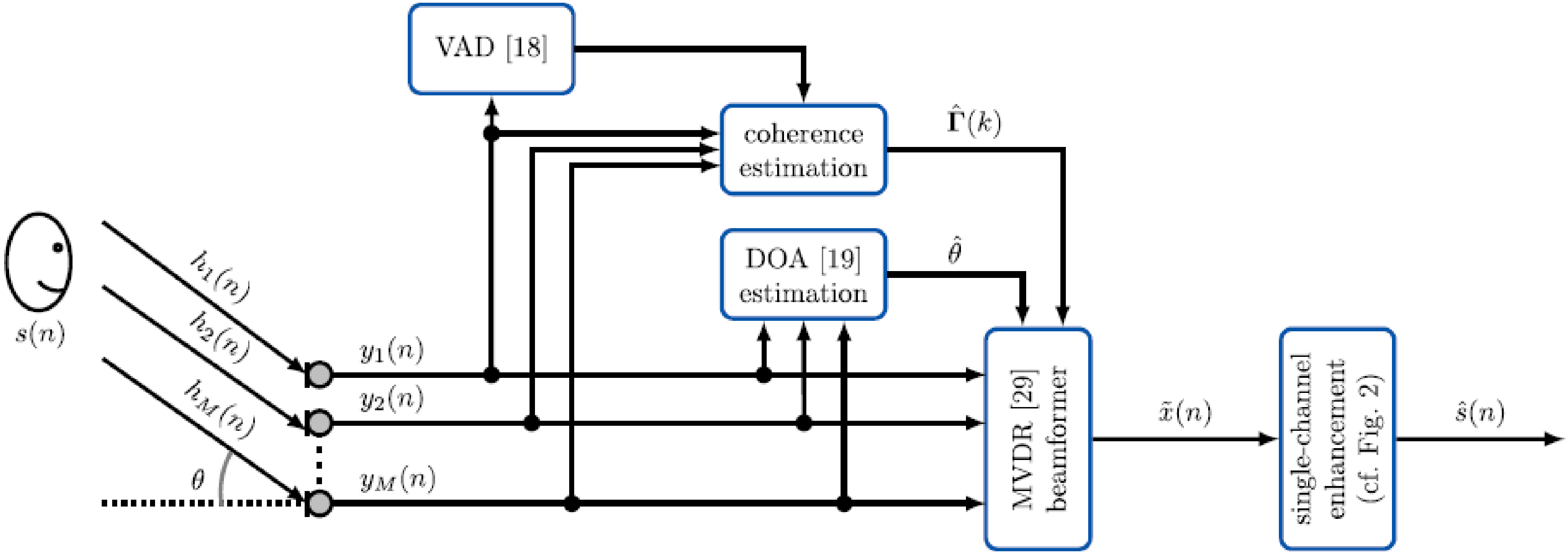
$$y_m(k,\ell) = \underbrace{h_m(k,\ell) * s(k,\ell)}_{x_m(k,\ell)} + v_m(k,\ell)$$

$h_m(k,l)$ = convolutive transfer function (CTF) in frequency bin $k$ and time frame $l$



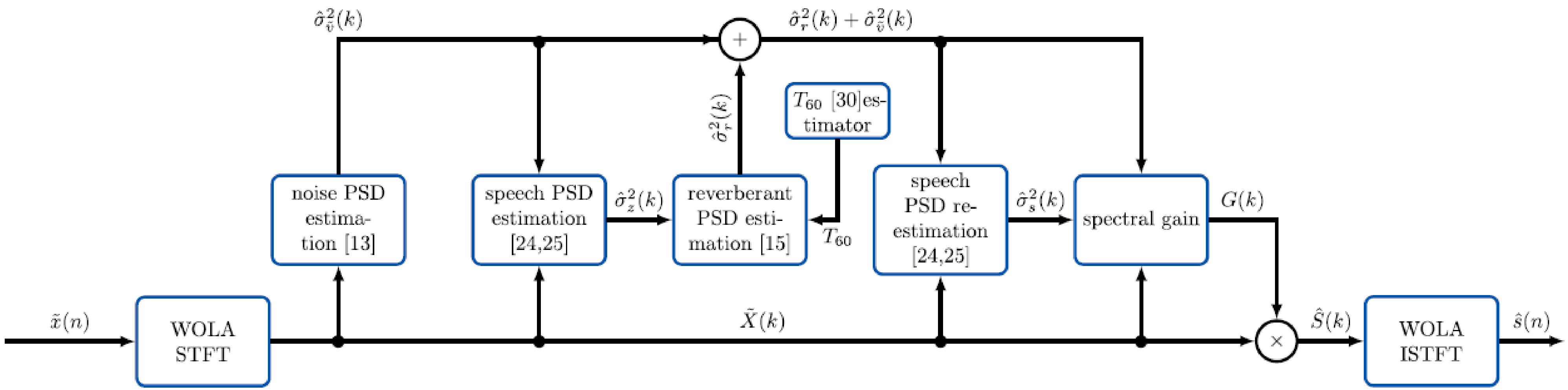



# 1. Beamforming + spectral post-filtering




- **MVDR beamformer:** $\mathbf{W}_\theta(k) = \dfrac{\Gamma^{-1}(k)\mathbf{d}_\theta(k)}{\mathbf{d}_\theta^H(k)\Gamma^{-1}(k)\mathbf{d}_\theta(k)}$
  - Anechoic **steering vector** based on DOA estimate (MUSIC): $\mathbf{d}_\theta(k) = \left[e^{-j2\pi f_k \tau_1(\theta)}\ e^{-j2\pi f_k \tau_2(\theta)}\ \ldots e^{-j2\pi f_k \tau_M(\theta)}\right]$
  - **Coherence matrix** adaptively estimated based on VAD (or assuming diffuse noise and reverberation): $\hat{\Gamma}(k) = \frac{1}{\mathbb{L}_\nu} \sum_{\ell \in \mathbb{L}_\nu} \mathbf{V}(k,\ell)\mathbf{V}^H(k,\ell)$ $\quad \Gamma_{i,i'}(k) = \dfrac{\sin(2\pi f_k l_{i,i'}/c)}{2\pi f_k l_{i,i'}/c}$

# 1. Beamforming + spectral post-filtering

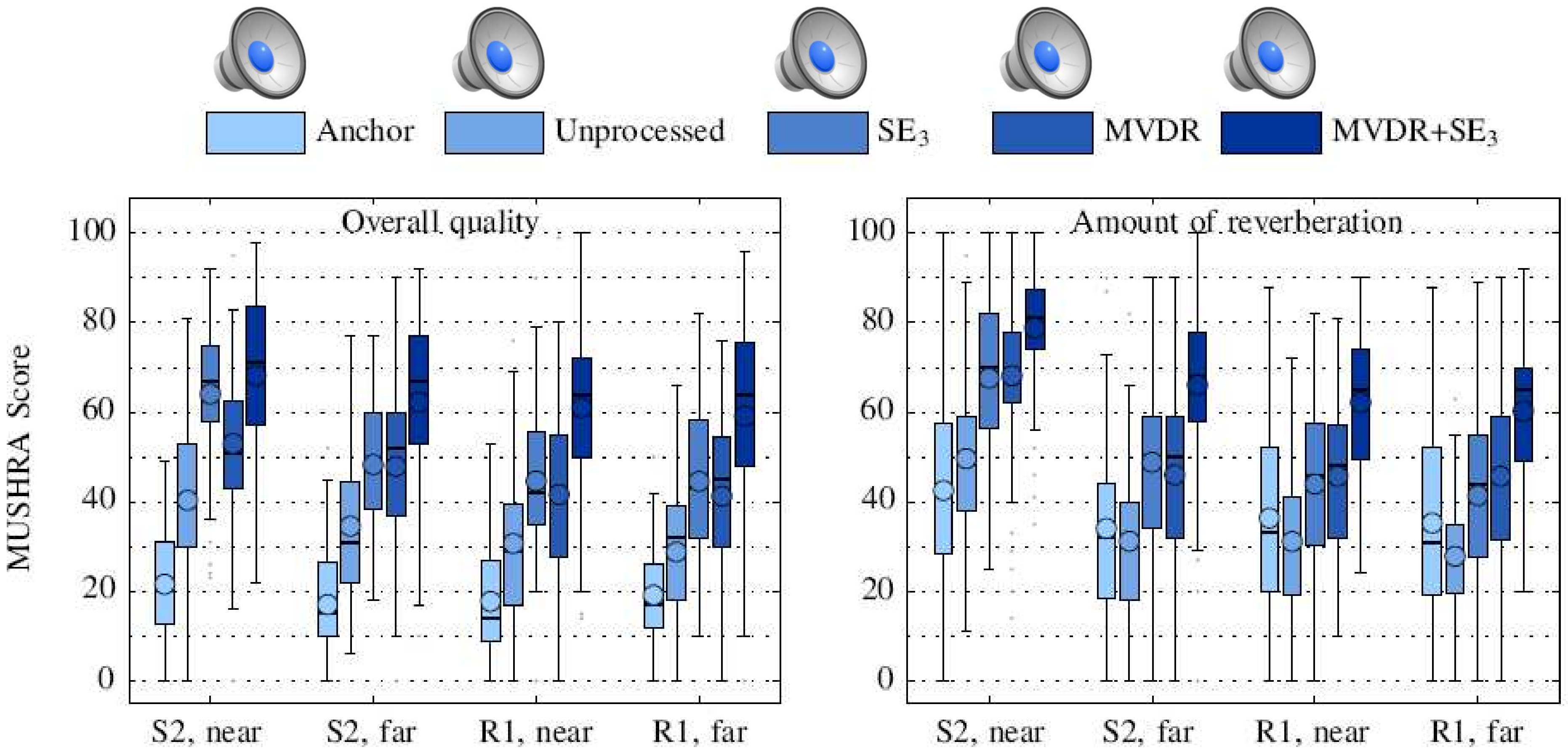



- **Spectral post-filter:** $\hat{S}(k,\ell) = G(k,\ell)\tilde{X}(k,\ell)$; $\xi(k,\ell) = \dfrac{\sigma_s^2(k,\ell)}{\sigma_r^2(k,\ell) + \sigma_{\tilde{v}}^2(k,\ell)}$

  1. **Noise PSD**: minimum statistics approach (longer window as usual)
  2. **Reverberant speech PSD**: ML estimate + cepstro-temporal smoothing
  3. **Late reverberant PSD**: assuming exponential decay (requiring T60 estimate) $\hat{\sigma}_r^2(k,\ell) = e^{-2\Delta T_d f_s} \hat{\sigma}_z^2(k,\ell - T_d/T_s)$
  4. **Clean speech PSD**: ML estimate + cepstro-temporal smoothing

# 1. Beamforming + spectral post-filtering

- **Subjective evaluation (evaluation set of REVERB challenge)**

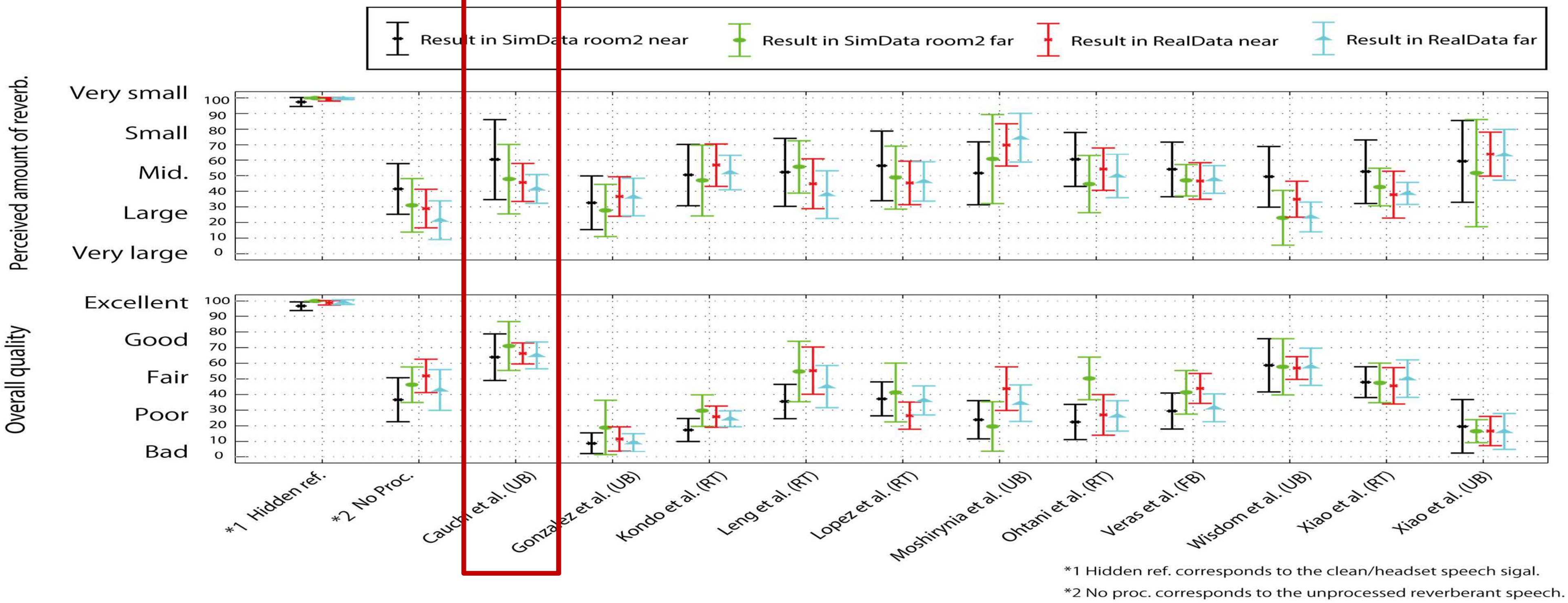



Circular array (M=8, d = 20 cm), fs = 16 kHz, SNR = 20 dB; S2: T60 = 500 ms (0.5m, 2m), R1: T60 = 700 ms (1m, 2.5m)
STFT: 32 ms, 50% overlap, Hann; MVDR: WNGmax = -10 dB; Postfilter: $\beta$=0.5, $\mu$=0.5, Gmin = -10dB, Td = 80 ms, MS window = 3s

[Cauchi 2015]

# 1. Beamforming + spectral post-filtering

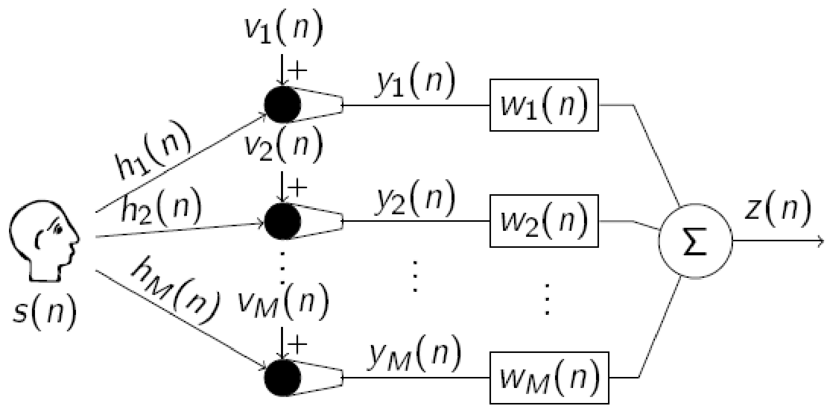
- **Subjective evaluation (evaluation set of REVERB challenge)**




[Cauchi 2015]

# 2. Acoustic multi-channel equalization

- **Time-domain approach** *(although frequency-domain versions possible)*
- **Indirect approach*:***
  1. estimate/measure RIRs
  2. Estimate the clean speech signal by inverting/equalizing the acoustic system + suppressing noise



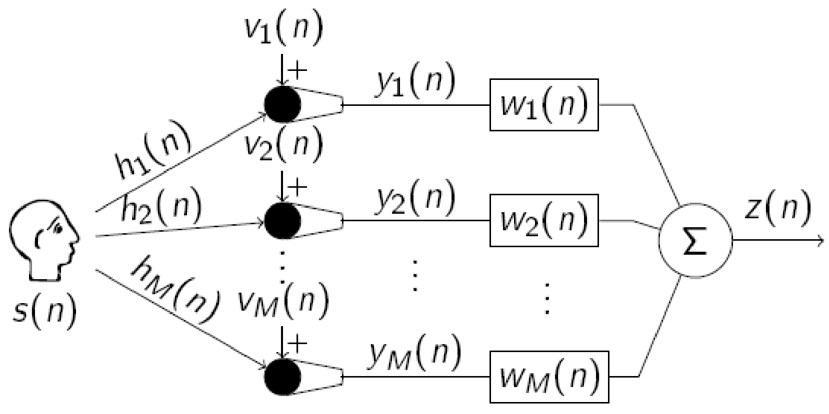

$$z(n) = \underbrace{\mathbf{w}^T\mathbf{H}^T}_{\mathbf{c}^T}\mathbf{s}(n) + \mathbf{w}^T\mathbf{v}(n)$$

**Speech enhancement objectives**

- Dereverberation: Optimize $\mathbf{c}$
- Noise reduction: Minimize the noise output power while controlling the speech distortion
- Joint dereverberation and noise reduction: Optimize $\mathbf{c}$ and minimize the noise output power

# 2. Acoustic multi-channel equalization

- **Time-domain approach** *(although frequency-domain versions possible)*
- **Indirect approach*:***
  1. estimate/measure RIRs
  2. Estimate the clean speech signal by inverting/equalizing the acoustic system + suppressing noise




$$z(n) = \underbrace{\mathbf{w}^T \mathbf{H}^T}_{\mathbf{c}^T} \mathbf{s}(n)$$

- If RIRs do not share common zeros and length of equalization filter is well chosen: **perfect dereverberation possible** (MINT theorem)

$$\mathbf{H}\mathbf{w} = \mathbf{c}_t$$

$\mathbf{c}_t$ = user-defined dereverberated target response (delayed impulse, early reflections, …)

- In practice: **large distortions** due to RIR perturbations (estimation errors, spatial errors, …)

$$\hat{\mathbf{H}}\mathbf{w} = \mathbf{c}_t$$

# 2. Robust acoustic multi-channel equalization

- **Framework for least-squares dereverberation**

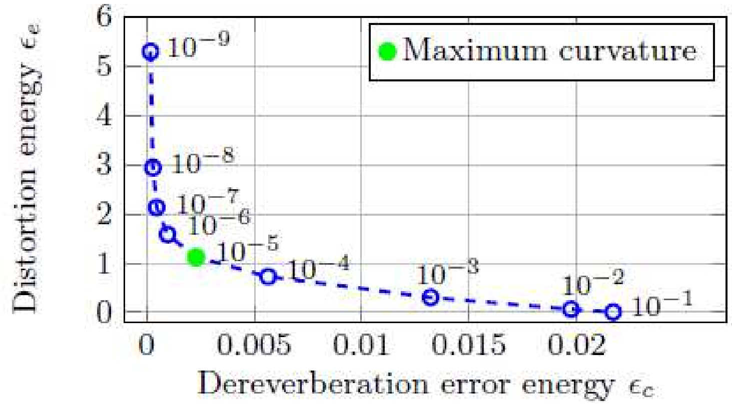
$$\|\mathbf{W}(\hat{\mathbf{H}}\mathbf{w} - \mathbf{c}_t)\|_2^2 \qquad \mathbf{w} = (\mathbf{W}\hat{\mathbf{H}})^+(\mathbf{W}\mathbf{c}_t)$$

  – Perceptually motivated target response $\mathbf{c}_t$ (**P-MINT**): suppress only late reflections while constraining early reflections

- **Increase robustness by:**
  1. *Decreasing filter length*
  2. *Signal-independent regularization*: control distortion energy due to RIR perturbations

$$J = \underbrace{\|\mathbf{W}(\hat{\mathbf{H}}\mathbf{w} - \mathbf{c}_t)\|_2^2}_{\epsilon_c} + \delta \underbrace{\mathbf{w}^T \mathbf{R}_e \mathbf{w}}_{\epsilon_e}$$

- **Closed-form solution**
- **Automatic procedure for selecting the regularization parameter** δ (based on L-curve), yielding both low dereverberation error energy and distortion energy




[Kodrasi 2013]

# 2. Robust acoustic multi-channel equalization

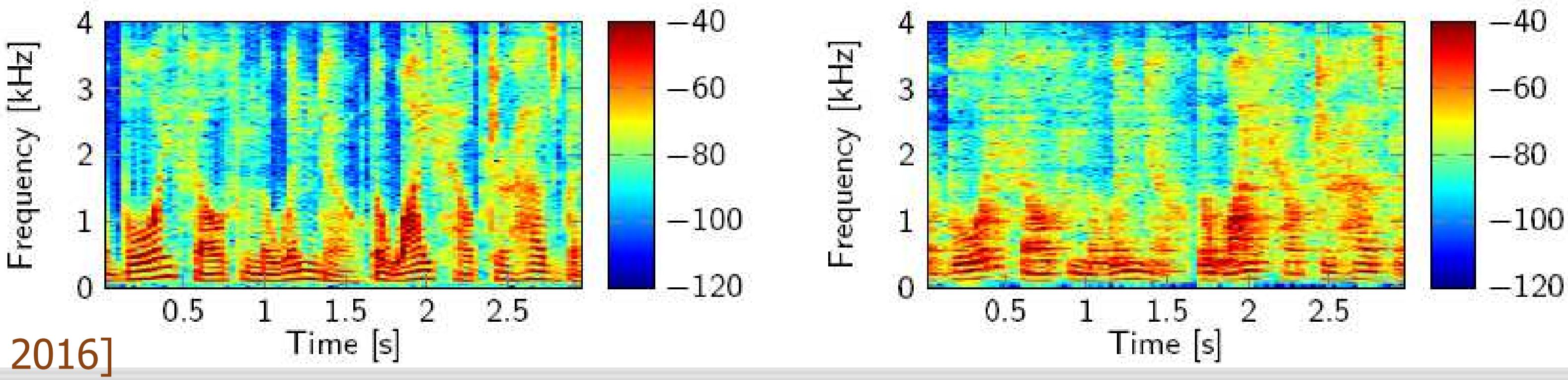
- **Framework for least-squares dereverberation**

$$\|\mathbf{W}(\hat{\mathbf{H}}\mathbf{w} - \mathbf{c}_t)\|_2^2 \qquad \mathbf{w} = (\mathbf{W}\hat{\mathbf{H}})^+(\mathbf{W}\mathbf{c}_t)$$

  – Perceptually motivated target response $\mathbf{c}_t$ (**P-MINT**): suppress only late reflections while constraining early reflections

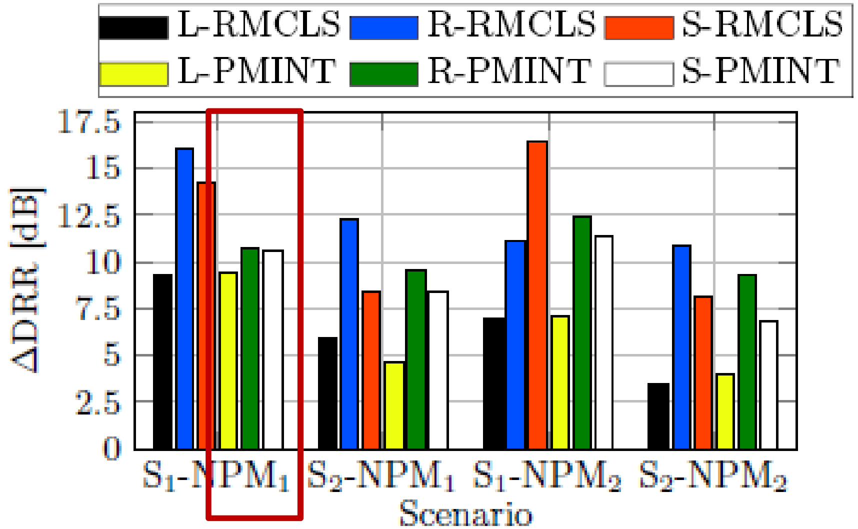
- **Increase robustness by:**
  1. *Decreasing filter length*
  2. *Signal-independent regularization*: control distortion energy due to RIR perturbations
  3. *Signal-dependent regularization*: enforce output signal to exhibit characteristics of clean signal, e.g. **promote sparsity of STFT coefficients** (weighted $l_1$-norm)

$$\min_{\mathbf{w}} \left[\|\mathbf{W}(\hat{\mathbf{H}}\mathbf{w} - \mathbf{c}_t)\|_2^2 + \eta f_{\text{sp}}(\mathbf{z}(n))\right]$$



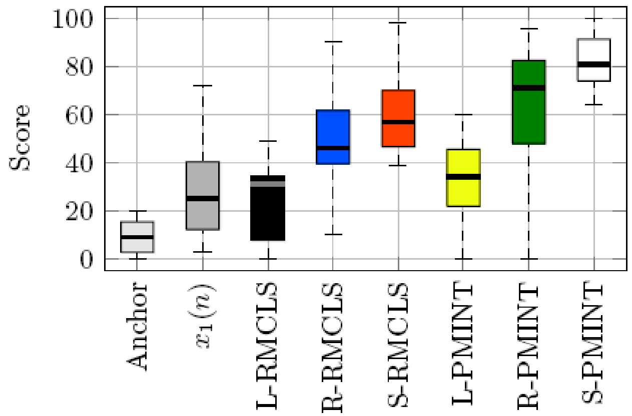
# 2. Robust acoustic multi-channel equalization

- **Instrumental validation**

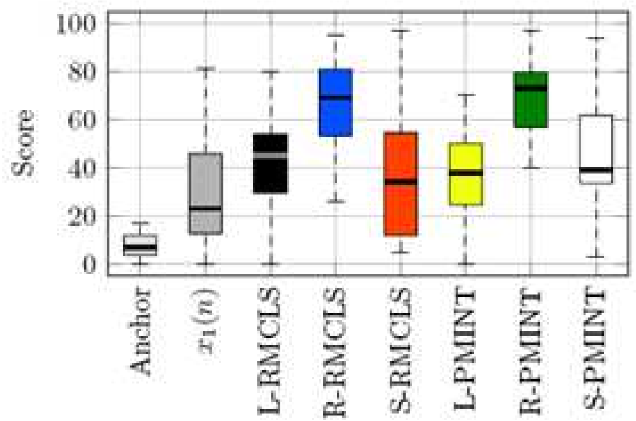



| s(n) | $x_1$(n) | PMINT | L-PMINT (intrusive) | R-PMINT (intrusive) | R-PMINT (auto) | S-PMINT (intrusive) |
|---|---|---|---|---|---|---|
| | | | | | | |

- **Subjective listening test**




($S_1$-$NPM_1$)




($S_1$-$NPM_2$)

M = 4, $S_1$: T60 = 450 msec, DRR = 0 dB, $S_2$: T60 = 610 msec, DRR = -2 dB, fs = 8 kHz; RIR estimation errors: $NPM_1$ = -33 dB, $NPM_2$ = -15 dB, L-RMCLS/L-PMINT: intrusively chosen filter length, R-RMCLS/R-PMINT: intrusively regularized, S-RMCLS/S-PMINT: intrusively regularized, $\tau$ = 90, $L_d$ = 10msec

# 2. Robust acoustic multi-channel equalization

- Equalization techniques for dereverberation lead to **noise amplification**
- **Cost functions for joint dereverberation and noise reduction:**
  1. Incorporate **noise statistics** into regularized P-MINT (RPM-DNR)

$$J = \underbrace{\|\hat{\mathbf{H}}\mathbf{w} - \hat{\mathbf{h}}_1^d\|_2^2}_{\epsilon_c} + \delta \underbrace{\mathbf{w}^T \mathbf{R}_e \mathbf{w}}_{\epsilon_e} + \mu \underbrace{\mathbf{w}^T \mathbf{R}_v \mathbf{w}}_{\epsilon_v}$$

  2. Incorporate **speech statistics** → Multi-channel Wiener Filter, using dereverberated output signal of regularized P-MINT as reference signal (MWF-DNR)

$$J = \mathcal{E}\{(\mathbf{w}^T \mathbf{x}(n) - \mathbf{w}_{RP}^T \mathbf{x}(n))^2\} + \mu \mathcal{E}\{(\mathbf{w}^T \mathbf{v}(n))^2\}$$

- Automatic selection of trade-off parameter(s)

| $y_1(n)$ | PMINT | R-PMINT | RPM-DNR | MWF-DNR |
|---|---|---|---|---|
| | | | | |

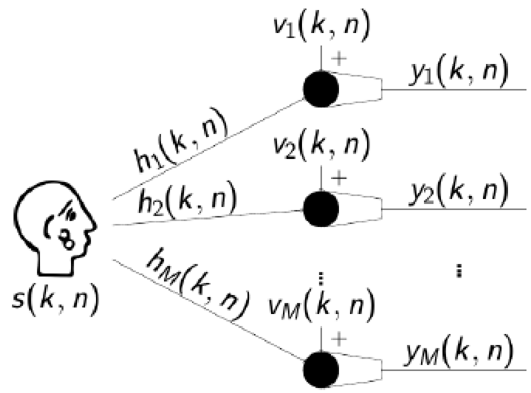
| Measure | PMINT | RPMINT | RPM-DNR | MWF-DNR |
|---|---|---|---|---|
| ΔDRR [dB] | −3.3 | **9.9** | 9.8 | 9.1 |
| ΔPESQ | −0.4 | **0.7** | **0.7** | 0.6 |
| $\psi_{\mathrm{NR}}$ [dB] | −26.8 | 1.9 | 3.2 | **13.0** |
| ΔfwSSNR [dB] | −3.0 | 0.9 | 1.1 | **3.2** |

M=4, T60=610 msec, DRR=-2 dB, fs=8 kHz, NPM=-33 dB, SIR=0 dB, SNR=10 dB (diffuse noise), no estimation errors in correlation matrices

[Kodrasi 2016]

# 3. Blind probabilistic model-based approach

- **STFT-domain approach** *(although time-domain versions possible)*
  - Low computational complexity (independent frequency bin processing)
  - Speech properties (e.g. sparsity) can be modelled more naturally in STFT-domain
- **Direct approach:** directly estimate clean speech STFT coefficients *s(k,n)* from reverberant (and noisy) STFT coefficients $y_m(k,n)$




$$y_m(k,n) = \underbrace{h_m(k,n) * s(k,n)}_{x_m(k,n)} + v_m(k,n)$$

1. Directly using CTF model → sparse Bayesian deconvolution based on variational Bayesian inference

2. Transform to equivalent AR model → sparse **multi-channel linear prediction (MCLP)**

$$x_1(k,n) = d(k,n) + \sum_{m=1}^{M} \sum_{l=0}^{L_g-1} g_m(k,l) x_m(k,n-\tau-l)$$

**clean signal (incl. early reflections)** **prediction filters** **delay (early reflections)**

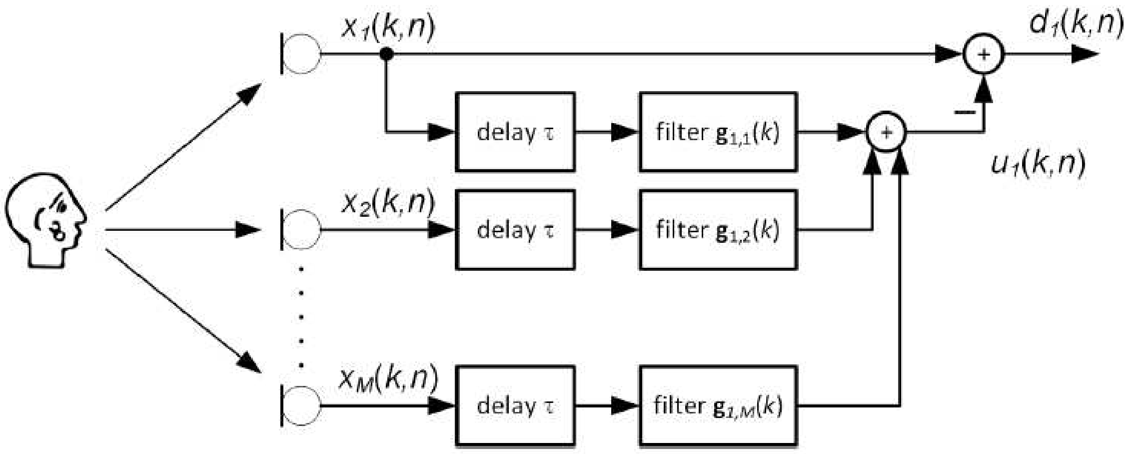
# 3. Multi-channel linear prediction

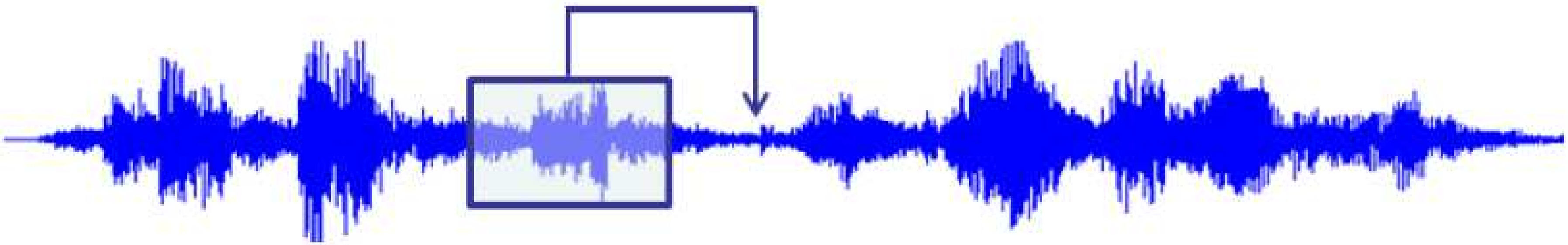
- **AR model of reverberant speech**

$$\mathbf{x}_1(k) = \mathbf{d}(k) + \mathbf{X}_\tau(k)\mathbf{g}(k)$$

$$\hat{\mathbf{d}}(k) = \mathbf{x}_1(k) - \mathbf{X}_\tau(k)\hat{\mathbf{g}}(k)$$

**predicted reverberation**



**How to select suitable cost function for prediction filters ?**

# 3. Multi-channel linear prediction

- Model clean speech STFT coefficients using **circular sparse prior**

$$\rho(d(n)) = \max_{\lambda(n)>0} \mathcal{N}_{\mathbb{C}}(d(n); 0, \lambda(n))\psi(\lambda(n))$$

  - **Time-varying variance** $\lambda(n)$
  - Hyper-prior on variance determined by scaling function $\psi(.)$

- **Maximum-Likelihood Estimation**

$$\mathcal{L}(\mathbf{g}) = \prod_{n=1}^{N} \rho(d(n)) \qquad \min_{\boldsymbol{\lambda}>0,\mathbf{g}} \sum_{n=1}^{N} \left( \frac{|d(n)|^2}{\lambda(n)} + \log \pi\lambda(n) - \log \psi(\lambda(n)) \right)$$

- **Alternating optimization procedure**
  1. Estimate **prediction vector** (assuming fixed variances)

$$\hat{\mathbf{g}}^{(i+1)} = \left( \mathbf{X}_\tau^H \mathcal{D}_{\hat{\boldsymbol{\lambda}}^{(i)}}^{-1} \mathbf{X}_\tau \right)^{-1} \mathbf{X}_\tau^H \mathcal{D}_{\hat{\boldsymbol{\lambda}}^{(i)}}^{-1} \mathbf{x}_1$$

  2. Estimate **variances** (assuming fixed prediction vector)

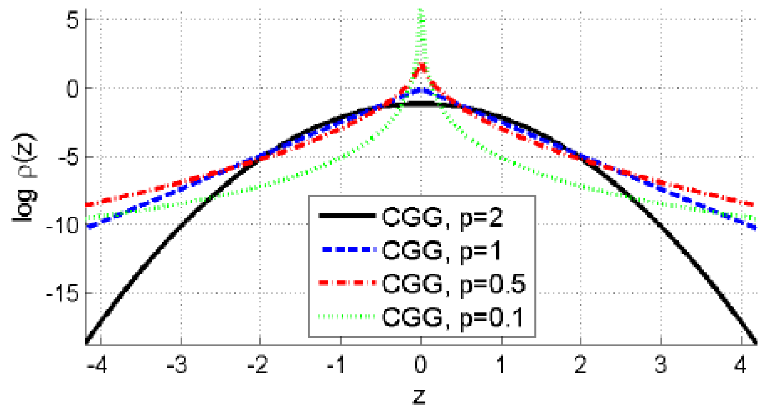
$$\hat{\lambda}^{(i+1)}(n) = \arg\min_{\lambda(n)>0} \frac{|\hat{d}^{(i+1)}(n)|^2}{\lambda(n)} + \log \pi\lambda(n) - \log \psi(\lambda(n))$$

# 3. Multi-channel linear prediction

- **Example:** complex generalized Gaussian (CGG) prior with shape parameter *p*

$$\rho(z) = \frac{p}{2\pi\gamma\Gamma(2/p)} e^{-\frac{|z|^p}{\gamma^{p/2}}}$$

$$\boxed{\hat{\lambda}^{(i+1)}(n) = |\hat{d}^{(i+1)}(n)|^{2-p},}$$

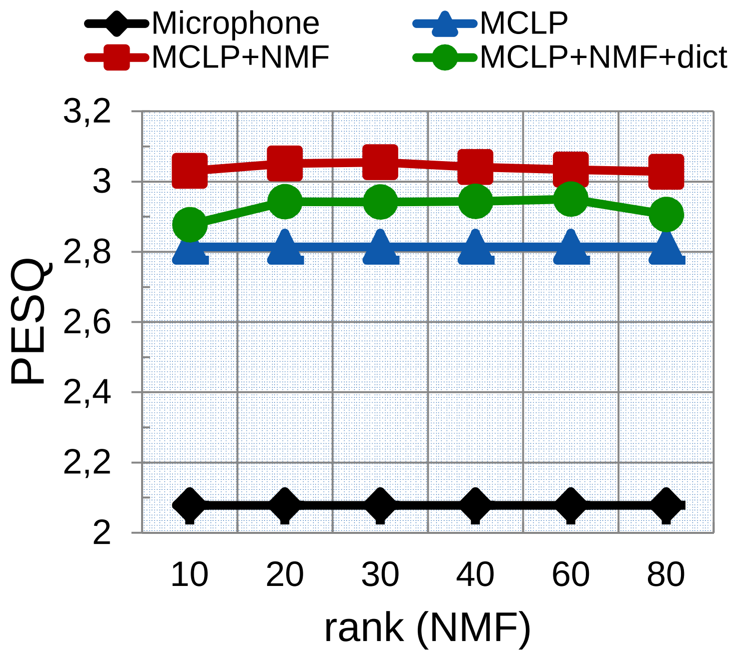


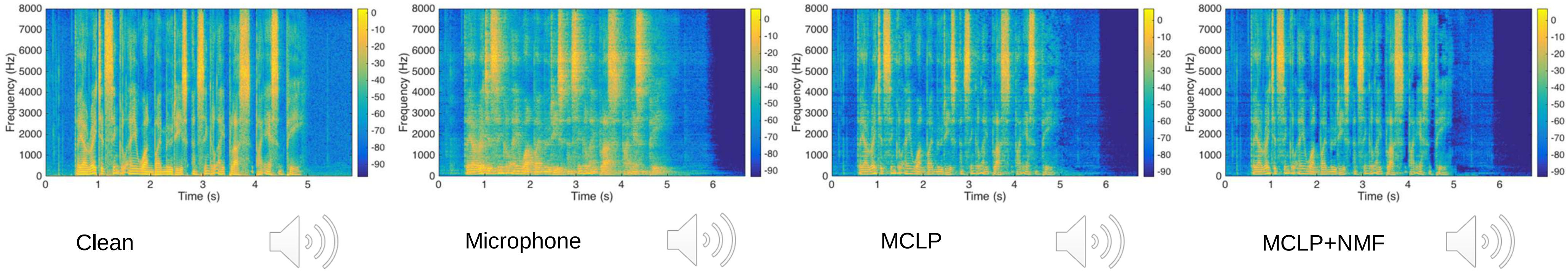

- **Remarks:**
  - Conventional method (TVG): p = 0
  - ML estimation using CGG prior is equivalent to ***$l_p$*-norm minimization** → iterative reweighted least-squares (IRLS) procedure $\min_{\mathbf{g}} \|\mathbf{d}\|_p^p$
  - Incorporate additional knowledge of speech signal, e.g. **low-rank structure** (NMF) $|\mathbf{D}|^2 \approx \underbrace{\mathbf{W}}_{\text{spectral dictionary}} \mathbf{H}$
  - **Group sparsity** for MIMO speech dereverberation → mixed norms

$$\|\mathbf{D}\|_{\Phi;2,p} = \left(\sum_{n=1}^{N} \|\mathbf{d}_{n,:}\|_{\Phi;2}^{p}\right)^{1/p}$$

# 3. Multi-channel linear prediction

- **Instrumental validation (noiseless, batch)**
  - MCLP exploits sparsity
  - NMF introduces speech structure (unsupervised vs. supervised NMF)



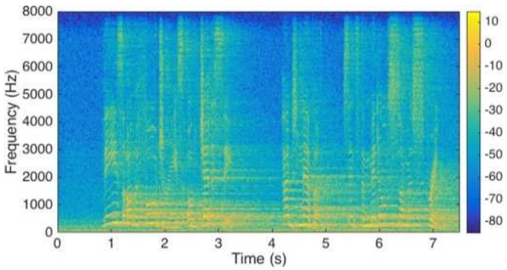

$T_{60}$ ~ 700ms, M=4, fs=16 kHz; STFT: 64ms (overlap 16ms); MCLP: $L_g$=8, $\tau$=2, p=0

[Jukic 2015]

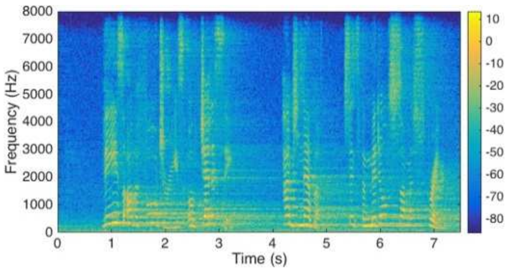
# 3. Multi-channel linear prediction

- **Instrumental validation (high reverberation + noisy, recursive)**

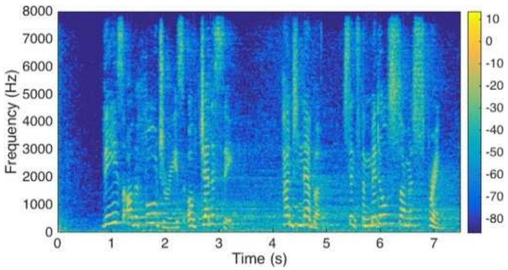
d ~ 2m



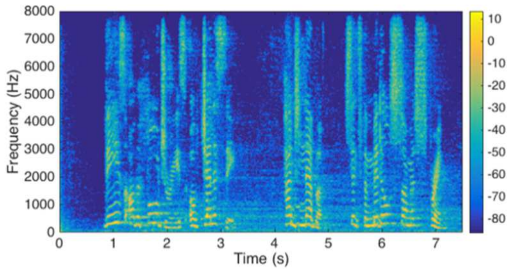

Microphone




1ch SE [REVERB]




Adaptive MCLP



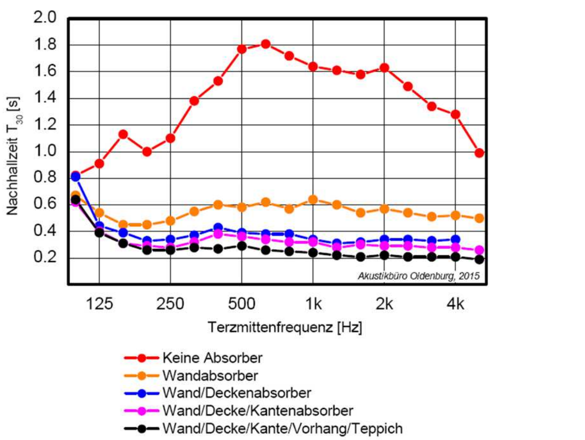

Adaptive MCLP + SE

T60 ~ 6s (St Alban The Martyr Church, London), M=2 (spacing~1m), fs=16 kHz, real recordings
STFT: 64ms (overlap 16ms); MCLP: $L_g$=30, $\tau$=2, p=0, recursive version ($\lambda$=0.96)

# Current / future work

- **Blind probabilistic model-based approach**
  - Comparison of CTF model vs. AR model
  - Recursive/adaptive versions of MCLP
- **Distributed MCLP** for acoustic sensor networks
- **Instrumental measures**: prediction of perceived level of reverberation, by optimizing/redesigning SRMR measure (joint project with Tiago Falk)
- Inaugurate new **varechoic lab**

Abbildung 1: In Raum E10 in den in Tabelle 1 angegebenen Raumzuständen gemessenen Nachhallzeiten in Terzbändern im Vergleich




# Recent publications

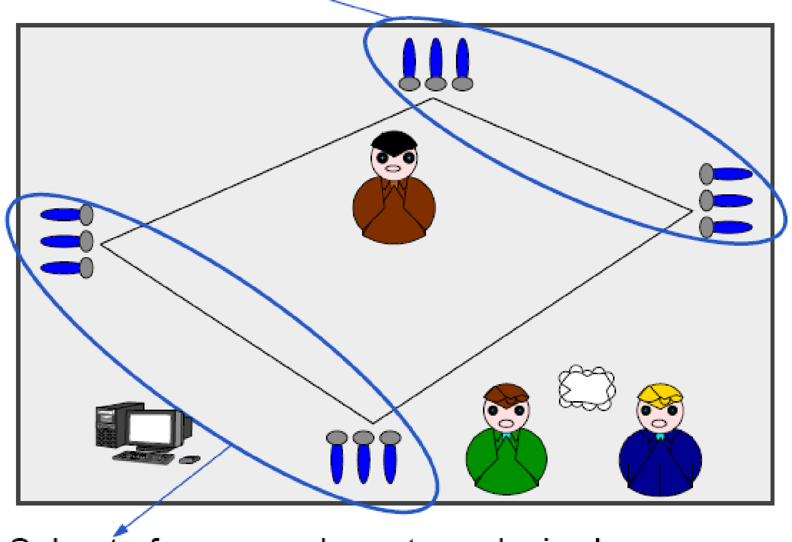

- B. Cauchi, I. Kodrasi, R. Rehr, S. Gerlach, A. Jukić, T. Gerkmann, S. Doclo, S. Goetze, Combination of MVDR beamforming and single-channel spectral processing for enhancing noisy and reverberant speech, *EURASIP Journal on Advances in Signal Processing*, 2015:61, pp. 1-12.
- I. Kodrasi, S. Doclo, Joint Dereverberation and Noise Reduction Based on Acoustic Multichannel Equalization, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 24, no. 4, pp. 680-693, Apr. 2016.
- I. Kodrasi, A. Jukic, S. Doclo, *Robust sparsity-promoting acoustic multi-channel equalization for speech dereverberation*, in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, Mar. 2016.
- I. Kodrasi, S. Goetze, S. Doclo, Regularization for Partial Multichannel Equalization for Speech Dereverberation, *IEEE Trans. Audio, Speech and Language Processing*, vol. 21, no. 9, pp. 1879-1890, Sep. 2013.
- A. Jukić, T. van Waterschoot, T. Gerkmann, S. Doclo, Multi-channel linear prediction-based speech dereverberation with sparse priors, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 23, no. 9, pp. 1509-1520, Sep. 2015.
- A. Jukić, T. van Waterschoot, T. Gerkmann, S. Doclo, Group sparsity for MIMO speech dereverberation, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, USA, Oct. 2015, pp. 1-5.
- A. Jukić, N. Mohammadiha, T. van Waterschoot, T. Gerkmann, S. Doclo, Multi-channel linear prediction-based speech dereverberation with low-rank power spectrogram approximation, in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 96-100.
- A. Jukić, T. van Waterschoot, T. Gerkmann, S. Doclo, Speech Dereverberation with Convolutive Transfer Function Approximation Using MAP and Variational Deconvolution Approaches, in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan les Pins, France, Sep. 2014, pp. 51-55.
- N. Mohammadiha, S. Doclo, Speech Dereverberation Using Non-negative Convolutive Transfer Function and Spectro-temporal Modeling, *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 24, no. 2, pp. 276-289, Feb. 2016.


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

# Acoustic Sensor Networks

# Acoustic Sensor Networks

- **Problem**
  - Traditional microphone arrays located at fixed and distant position
  - Poor performance of signal enhancement algorithms due to low SNR and/or low DRR
- **Objectives**
  - Develop centralized and distributed noise reduction and dereverberation algorithms
  - Optimise positions of distributed microphones
  - Impact of wireless link capacity, sampling rate offset estimation and compensation
- **Approaches**
  - Using statistical room acoustics model, compute spatially averaged performance → selection of optimal microphone configuration
  - Exploit diversity of room impulse responses → generalized/alternative versions of MWF





Toby Lawin-Ore


Nico Gößling

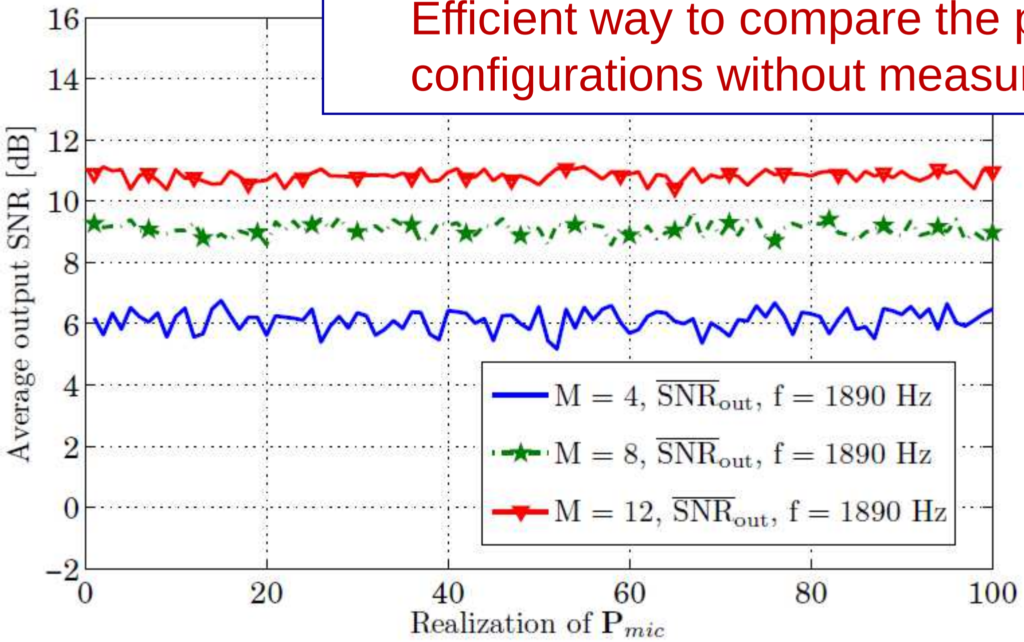
# Spatially averaged performance of MWF

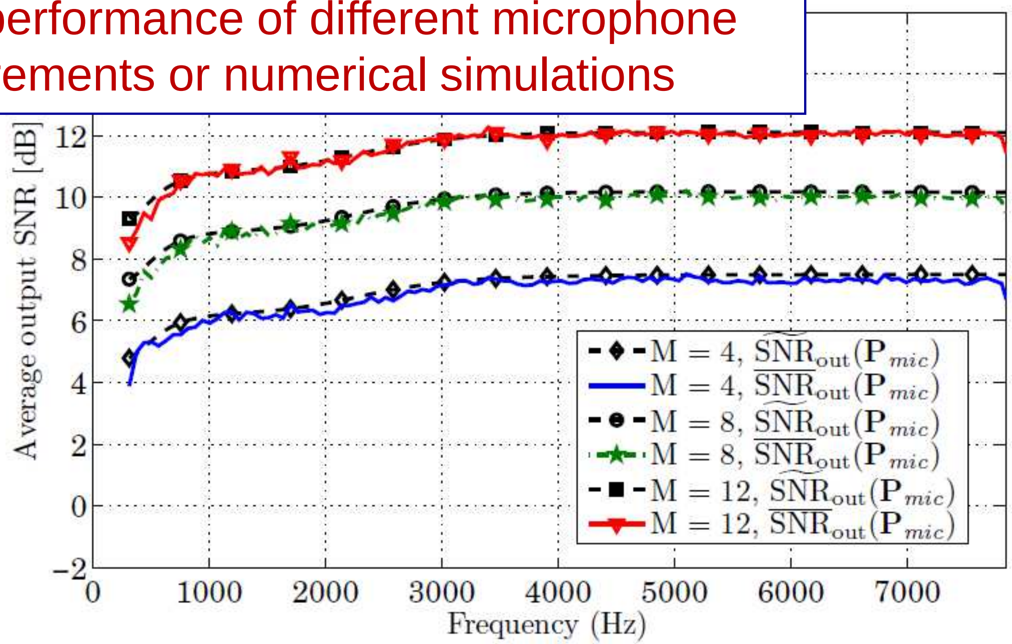
- **Goal :** using statistical room acoustic (**Schroeder's model**), compute the average output SNR of the multichannel Wiener filter → comparison of different microphone configurations
  - Specific microphone array position
  - Average over all source positions

$$\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}^j_{mic}) = \mathcal{E}_{\mathbf{d}}\{\mathcal{E}_{\mathbf{P}|\mathbf{d}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}\},\ \forall j$$

$$\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i) = \mathcal{E}_{\mathbf{P}|\mathbf{d}^i}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\} = \frac{\phi_s}{\phi_v}\sum_{m=1}^{M}\sum_{n=1}^{M}\breve{\gamma}_{mn}\left(\frac{e^{j\frac{\omega}{c}(d^i_n - d^i_m)}}{(4\pi)^2 d^i_m d^i_n} + \frac{1-\bar{\alpha}}{\pi\bar{\alpha}A}\frac{\sin\left(\frac{\omega}{c}r_{mn}\right)}{\frac{\omega}{c}r_{mn}}\right)$$

Efficient way to compare the performance of different microphone configurations without measurements or numerical simulations




[Lawin-Ore 2015]

# Generalized multi-channel Wiener filter

- **Multichannel Wiener filter (MWF) in acoustic sensor networks**
  - **Objective:** estimate filtered version of speech signal + trade-off noise reduction and speech distortion

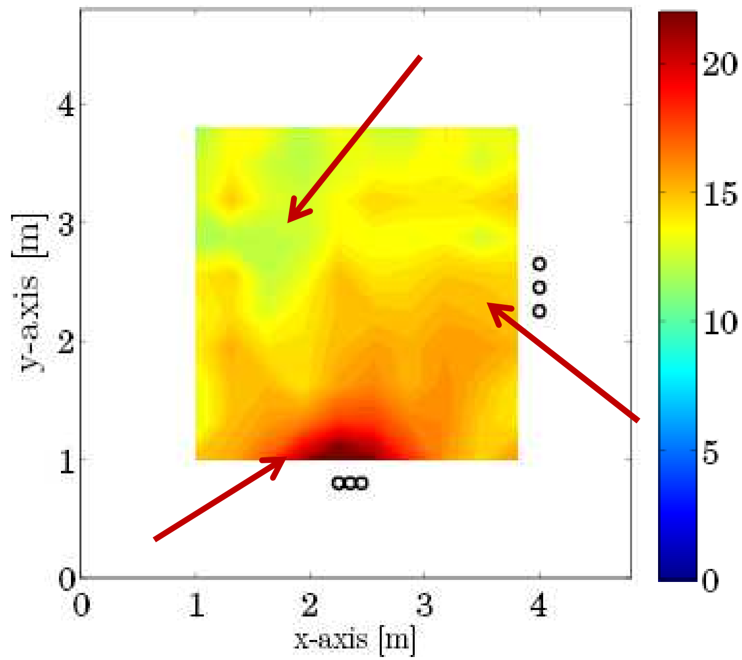
$$\xi(\mathbf{W}) = \mathscr{E}\{|A_d S - \mathbf{W}^H \mathbf{X}|^2\} + \mu \mathscr{E}\{|\mathbf{W}^H \mathbf{V}|^2\}$$

**Desired overall transfer function**

  - **"Standard" MWF (S-MWF):** speech component in reference microphone signal $m_0$

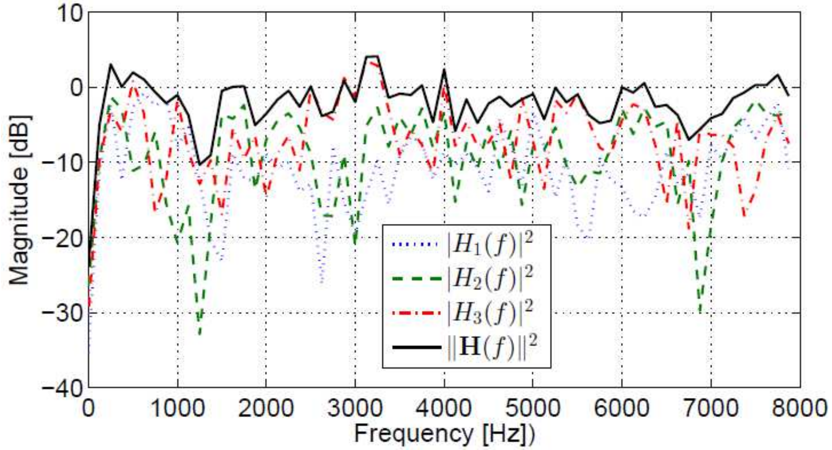
$$A_d = A_{m_0} \;\rightarrow\; \mathbf{W}_{\text{S-MWF}} = (\mathbf{\Phi}_x + \mu \mathbf{\Phi}_v)^{-1} \mathbf{\Phi}_x \mathbf{e}_{m_0}$$

  - For **spatially distributed microphones,** selection of reference microphone can have a major impact on output SNR (estimation errors depend on input SNR)



# Generalized multi-channel Wiener filter

- **Multichannel Wiener filter (MWF) in acoustic sensor networks**
  - **Generalized MWF:** define desired overall transfer function using envelope of ATFs




$$A_d = \sqrt{\sum_{m=1}^{M} \alpha_m |A_m|^2}\; e^{j\psi_{m_0}}$$

$$\downarrow$$

$$\boxed{\mathbf{W}_{\text{G-MWF}} = (\mathbf{\Phi}_x + \mu \mathbf{\Phi}_v)^{-1} \mathbf{\Phi}_x \mathbf{g}}$$

$$g_m = \frac{\sqrt{\Phi_x(m,m)}\sqrt{\sum \alpha_m \Phi_x(m,m)}}{\mathrm{tr}(\mathbf{\Phi}_x)} \frac{\Phi_x(m,m_0)}{|\Phi_x(m,m_0)|}$$

  - *Note: phase of reference microphone* $\psi_{m_0} = \arg(A_{m_0})$ *does not have influence on narrowband/broadband output SNR*

# Generalized multi-channel Wiener filter

- **Multichannel Wiener filter (MWF) in acoustic sensor networks**

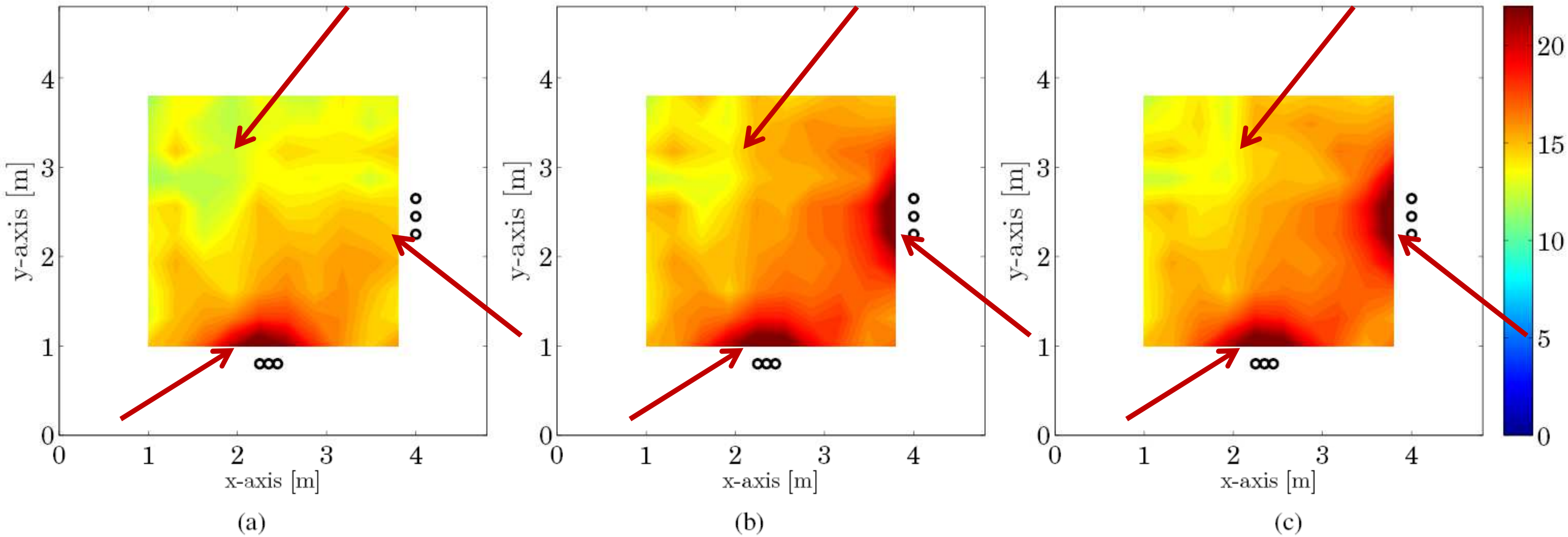



Figure 3: Position-dependent broadband output SNR of the different MWF filters: (a) S-MWF with $A_d = A_1$, (b) G-MWF with $A_d = A_1$, (c) G-MWF with $A_d = ||\mathbf{A}||e^{j\psi_1}$.

| S-MWF | | | | | | G-MWF | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_d{=}A_{m_0}$ | | | | | | $A_d{=}A_{m_0}$ | | | | | | $A_d{=}\|\|\mathbf{A}\|\|e^{j\psi_{m_0}}$ |
| $m_0{=}1$ | $m_0{=}2$ | $m_0{=}3$ | $m_0{=}4$ | $m_0{=}5$ | $m_0{=}6$ | $m_0{=}1$ | $m_0{=}2$ | $m_0{=}3$ | $m_0{=}4$ | $m_0{=}5$ | $m_0{=}6$ | $m_0{=}1\ldots6$ |
| **14.18** | 13.73 | 13.42 | 13.75 | 14.14 | 13.55 | **16.08** | 15.68 | 15.62 | 15.70 | 15.87 | 15.71 | **15.90** |

Table 1: Output SNR ($\text{oSNR}_{\text{avg}}$ [dB]) of the S-MWF and G-MWF filters, averaged over all considered source positions.

[Lawin-Ore 2014]

# Current/Future work

- Combination of acoustic sensor networks and **(binaural) hearing aids**
- Distributed algorithms for **dereverberation** (e.g. MCLP)

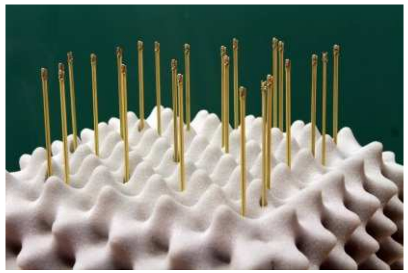
# Recent publications


- T. C. Lawin-Ore, S. Doclo, *Analysis of the average performance of the multichannel Wiener filter based noise reduction using statistical room acoustics*, *Signal Processing, special issue on on wireless acoustic sensor networks and ad hoc microphone arrays*, vol. 107, pp. 96-108, Feb. 2015.
- S. Stenzel, T. C. Lawin-Ore, J. Freudenberger, S. Doclo, *A Multichannel Wiener Filter with Partial Equalization for Distributed Microphones*, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz NY, USA, Oct. 2013.
- T. C. Lawin-Ore, S. Stenzel, J. Freudenberger, S. Doclo, *Generalized Multichannel Wiener Filter for Spatially Distributed Microphones*, in *Proc. ITG Conference on Speech Communication*, Erlangen, Germany, Sep. 2014.
- T. C. Lawin-Ore, S. Stenzel, J. Freudenberger, S. Doclo, *Alternative Formulation and Robustness Analysis of the Multichannel Wiener Filter for Spatially Distributed Microphones*, in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan les Pins, France, Sep. 2014, pp. 208-212.
- L. Wang, S. Doclo, *Correlation Maximization Based Sampling Rate Offset Estimation for Distributed Microphone Arrays*, *IEEE/ACM Trans. Audio, Speech and Language Processing*,vol. 24, no. 3, pp. 571-582, Mar. 2016.


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

# Current research topics

- **Speech enhancement for ear-mounted communication devices**
  - **Binaural noise reduction**, aiming to preserve spatial impression of acoustic scene (binaural cues)
  - Open-fitting hearing aids: **feedback cancellation** and **active noise control** (acoustically transparent earpiece)
  - EEG-based **auditory attention decoding** for steering beamformers



- **MIMO acoustics**
  - **Beamformer design** (e.g., virtual artificial head)
  - **Dereverberation and noise reduction** (spectral enhancement, multi-channel equalization, blind probabilistic model)
  - **Acoustic sensor networks** (spatially distributed microphones, sampling rate offset estimation, distributed processing)
  - **Computational acoustic scene analysis** (CASA)

# Questions ?

***House of Hearing, Oldenburg***