Acoustically Transparent Earpiece: Equalization, Feedback cancellation, Active noise control and Own voice pickup

Prof. Dr. Simon Doclo

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

Webinar Danish Sound Cluster, May 9 2023

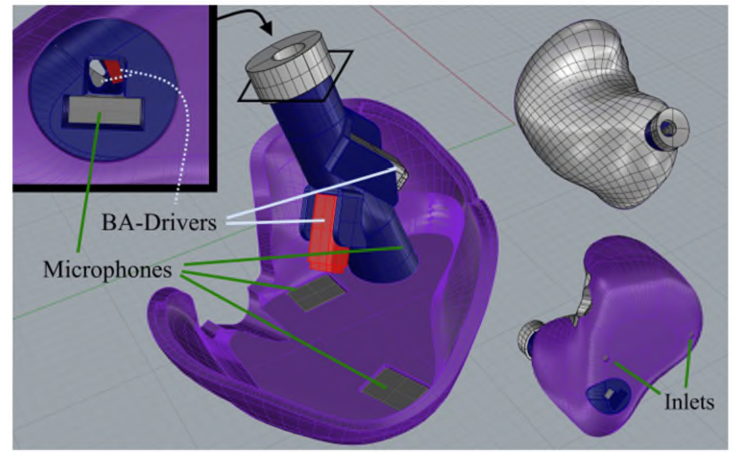
Acoustically Transparent Earpiece

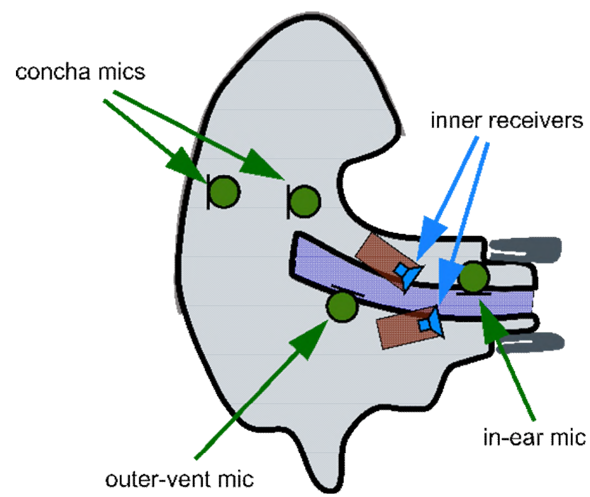
- **Long-term objective:**
  - Acoustically transparent speech communication using earpiece with multiple microphones and receivers/loudspeakers

- Develop, implement, and evaluate **individualized algorithms** for
  1. sound pressure equalization (transparency)
  2. acoustic feedback cancellation
  3. active noise/occlusion control
  4. own voice extraction

Acoustically Transparent Earpiece (Hearpiece)

- One-size-fits-all design: fits about 90% of human ears
- Vent: **2 microphones, 2 receivers**
- Concha: **2 microphones**
- Two versions: vented + closed
- ***Available at Hörzentrum / InEar GmbH:***
  https://www.hz-ol.de/en/hearpiece.html





[Denk et al., *AES Conference Headphone Technology*, 2019]

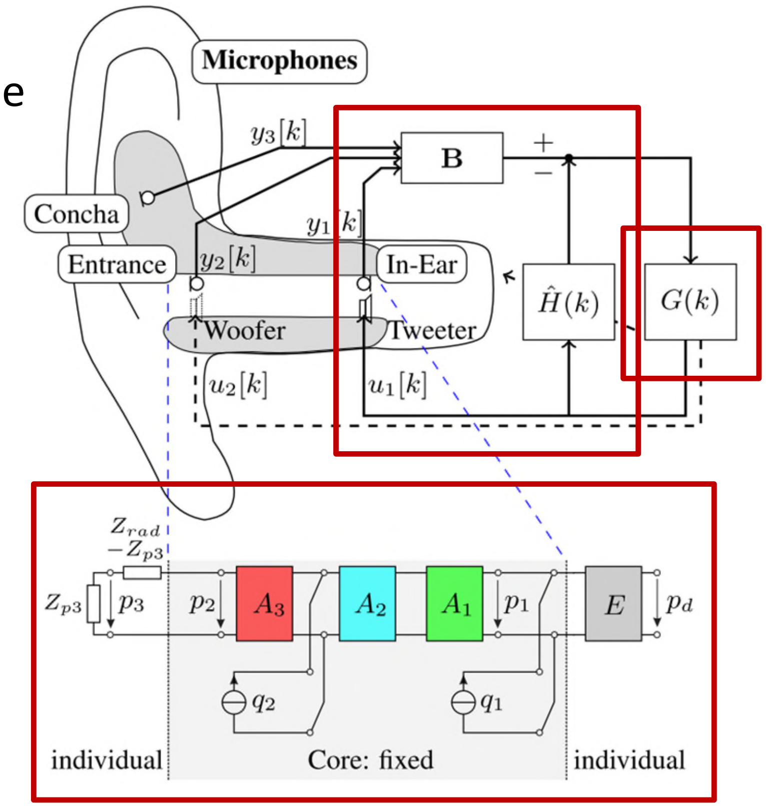
Acoustically Transparent Earpiece

1. Acoustically transparent sound presentation:

- Enable hearing comparable to open ear (equalization using single/multiple receivers)

2. Individualized Electro-Acoustic Model:

- Better understand acoustics
- Predict sound pressure and transfer functions (eardrum)



3. Acoustic Feedback cancellation

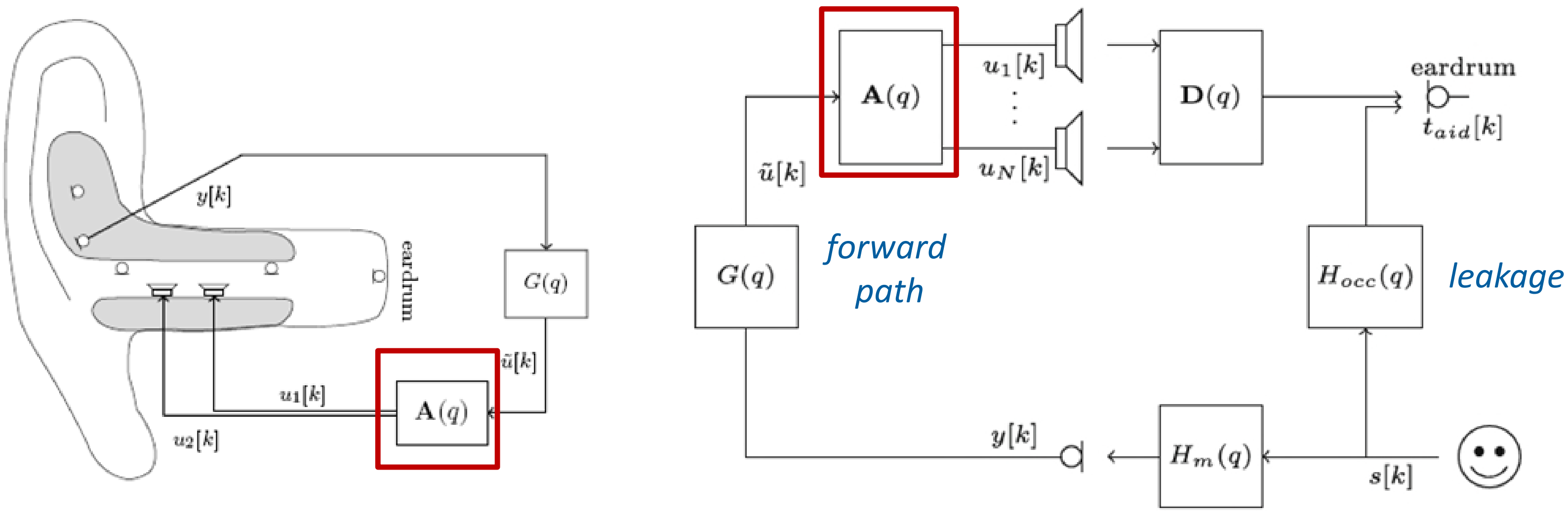
- Exploit multiple microphones to steer null towards position of receiver
- Exploit multiple receivers

4. Hearing support:

- Amplification and dynamic range compression
- Active noise and occlusion control

1. Acoustically transparent sound presentation

- **Single/Multi-Loudspeaker Equalization**
  - **Goal:** Achieve sound pressure at aided ear that is (physically or perceptually) equivalent to sound pressure at open ear
  - **Design and apply equalization filter(s) A*(q)*** to concha microphone signal, taking into account leakage and hearing device processing (forward path)



$$G(q)H_m(q)\mathbf{D}^T(q)\mathbf{A}(q) = G(q)H_{open}(q) - H_{occ}(q)$$

[Schepker et al., *EURASIP JASMP,* 2022]

1. Acoustically transparent sound presentation

- **Single/Multi-Loudspeaker Equalization**
  - **Goal:** Achieve sound pressure at aided ear that is (physically or perceptually) equivalent to sound pressure at open ear
  - **Design and apply equalization filter(s) A***(q)* to concha microphone signal, taking into account leakage and hearing device processing (forward path)
  - **Robust least-squares-based design procedure** (with group delay compensation and frequency-dependent regularization)
  - **Requires** (one or multiple) measurements of all transfer functions

multiple measurements — *regularization*

$$J_{mfr\Delta LS}(\mathbf{a}) = \sum_{i=1}^{I} \|\mathbf{D}_{\Delta,i}\mathbf{a} - \tilde{\mathbf{v}}_{\Delta,i}\|_2^2 + \lambda\|\mathbf{W}\mathbf{F}\mathbf{a}\|_2^2$$

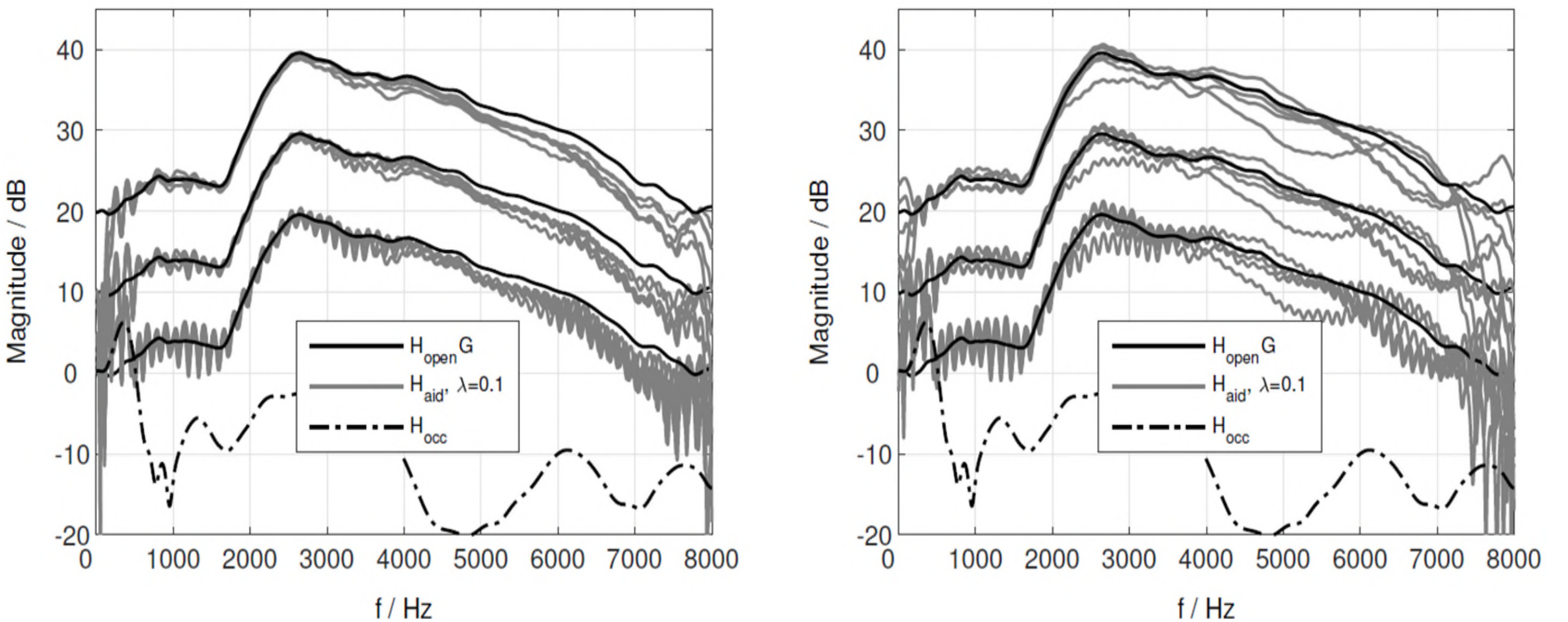
$$\mathbf{a}_{mfr\Delta LS} = \left(\bar{\mathbf{D}}_\Delta^T\bar{\mathbf{D}}_\Delta + \lambda\mathbf{F}^H\mathbf{W}^H\mathbf{W}\mathbf{F}\right)^{-1}\bar{\mathbf{D}}_\Delta^T\bar{\mathbf{v}}_\Delta$$

$$\tilde{\mathbf{v}}_\Delta = \left(\mathbf{H}_{m,\Delta}^T\mathbf{G}_\Delta^T\mathbf{G}_\Delta\mathbf{H}_{m,\Delta}\right)^{-1}\mathbf{H}_{m,\Delta}^T\mathbf{G}_\Delta^T\left(\mathbf{G}_\Delta\tilde{\mathbf{h}}_{open,\Delta} - \tilde{\mathbf{h}}_{occ,\Delta}\right)$$

[Schepker et al., *EURASIP JASMP,* 2022]

1. Acoustically transparent sound presentation

- **Single/Multi-Loudspeaker Equalization**



(a) $N = 1.$ (b) $N = 2.$

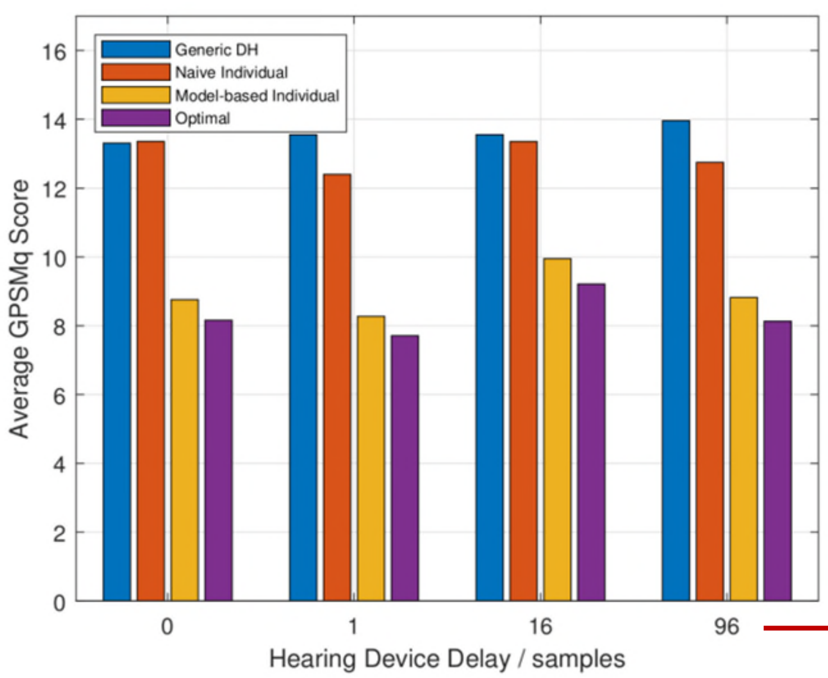
Robust equalization possible both using 1 and 2 loudspeakers

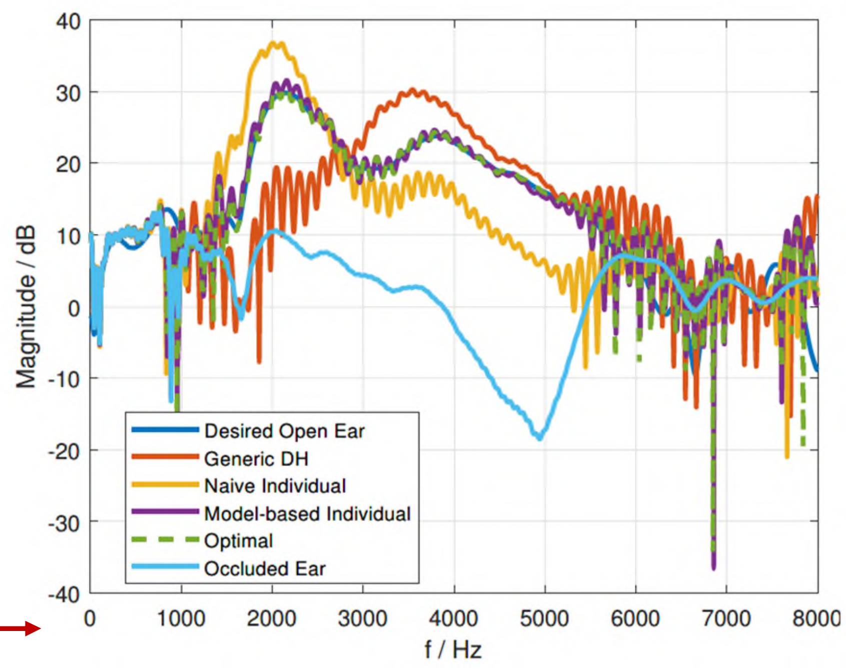
f_s = 16 kHz, τ = 6 ms,
L_A = 99, d_H = 32, λ = 0.1,
I = 4, G_0 = [0, 10, 20] dB

[Schepker et al., *EURASIP JASMP*, 2022]

1. Acoustically transparent sound presentation

- **Single/Multi-Loudspeaker Equalization**
  - **Individualized equalization filters** (based on individualized electro-acoustic model) outperform equalization filters based on dummy-head measurements or based on in-ear microphone signal

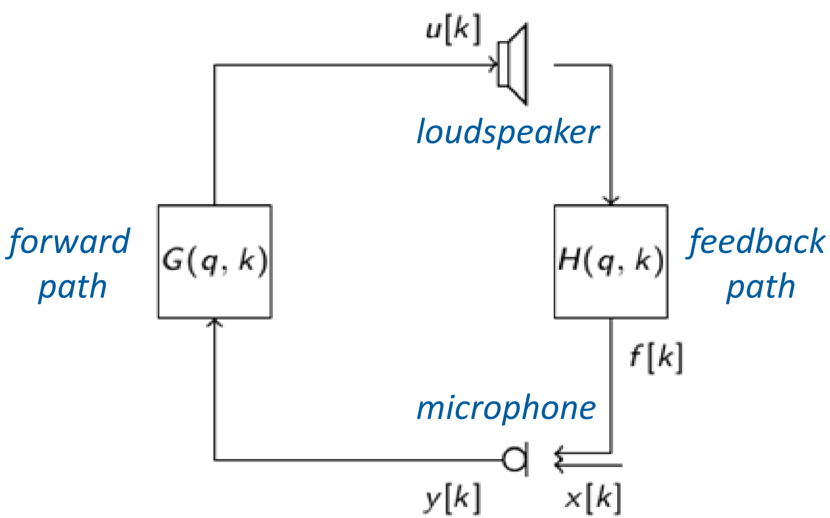


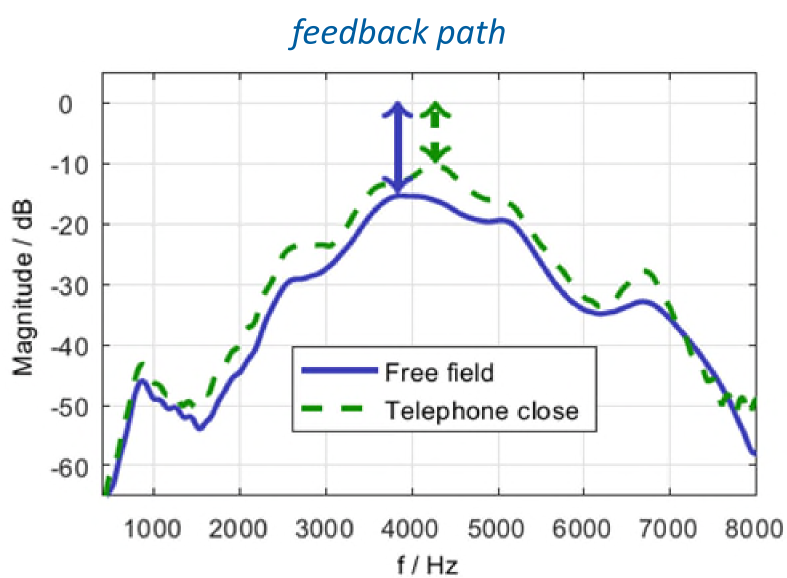


[Schepker et al., *JASA-EL*, submitted]

2. Acoustic feedback cancellation

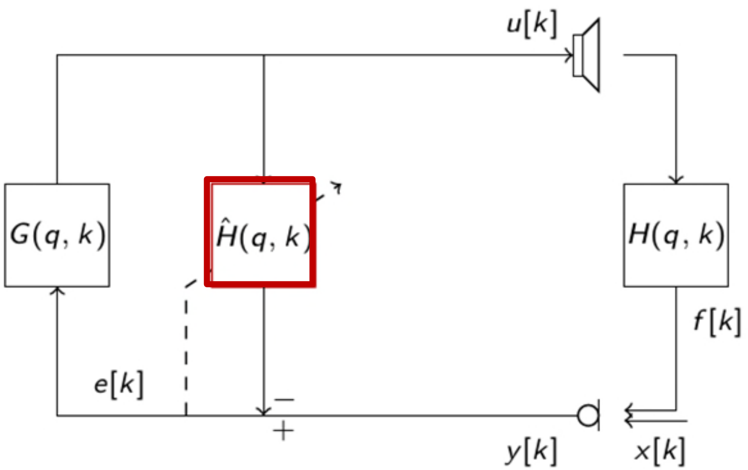
- Feedback arises due to **acoustic coupling** between loudspeaker(s) and microphone(s)

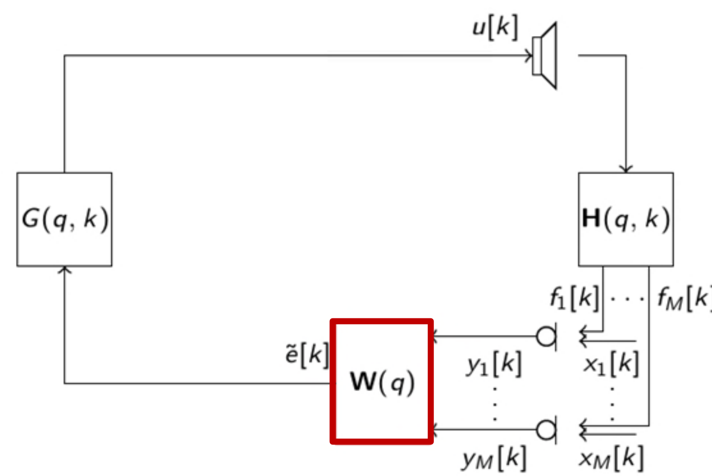




2. Acoustic feedback cancellation

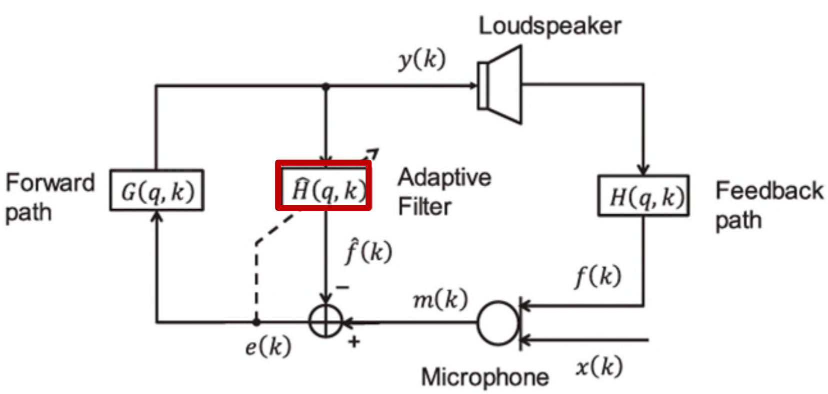
- Feedback arises due to **acoustic coupling** between loudspeaker(s) and microphone(s)
- **Feedback suppression approaches:**
  1. Feedforward suppression → distortion
  2. Adaptive feedback cancellation → decorrelation between loudspeaker and incoming signal
  3. Spatial filtering → requires multiple microphones

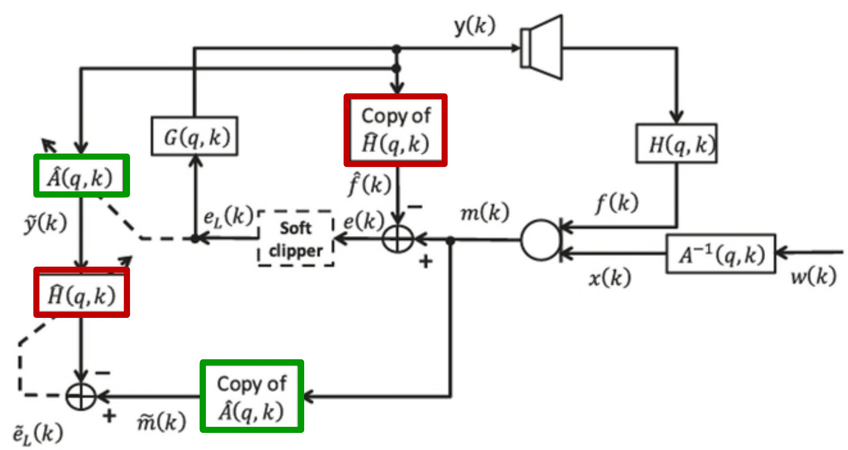




2. Acoustic feedback cancellation

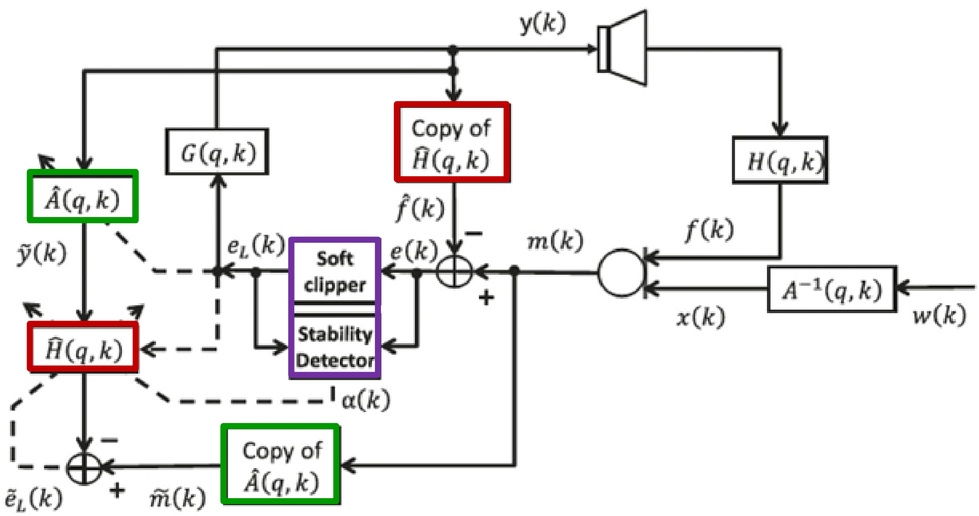
- **Adaptive feedback cancellation**
  1. Normalized least mean squares (NLMS): bias due to correlation, fast re-convergence from howling
  2. Prediction error method (PEM-NLMS): pre-whitening based on auto-regressive model → reduced bias, but slow re-convergence from howling





2. Acoustic feedback cancellation

- **Adaptive feedback cancellation**
  1. Normalized least mean squares (NLMS): bias due to correlation, fast re-convergence from howling
  2. Prediction error method (PEM-NLMS): pre-whitening based on auto-regressive model → reduced bias, but slow re-convergence from howling
  3. **Hybrid algorithm (H-NLMS):** switched combination of NLMS and PEM-NLMS update, controlled by soft-clipping-based stability detector



$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \underbrace{[1-\alpha(k)]\overbrace{\mu_1(k)\tilde{\mathbf{y}}(k)\tilde{e}_L(k)}^{\textit{PEM-NLMS}}}_{} + \alpha(k)\underbrace{\mu_2(k)\mathbf{y}(k)e_L(k)}_{\textit{NLMS}},$$

Binary control signal:

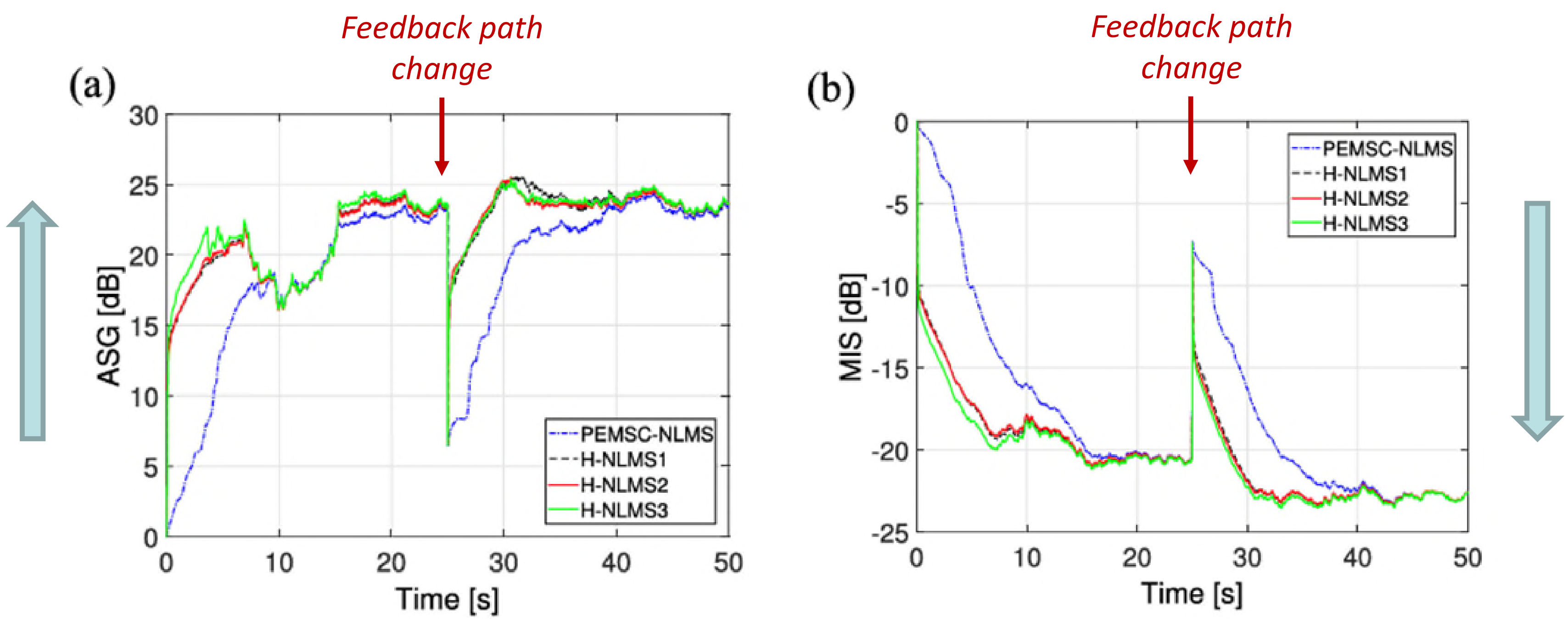
$$\alpha(k) = I\{|e(k) - e_L(k)| < \gamma\}$$

$$e_L(k) = \beta \tanh\left(\frac{e(k)}{\beta}\right)$$

[Nordholm et al., *JASA,* 2018]

2. Acoustic feedback cancellation

- **Simulation results:** added stable gain (ASG), misalignment (MIS)



H-NLMS algorithm converges much faster than PEM-NLMS while maintaining similar misalignment

f_s = 16 kHz, L_h = 64, L_a = 20 (Levinson-Durbin), forward path: τ = 6 ms, G_0 = 45 dB, β = 2, γ = 0.15, μ_1 = 0.001, μ_2 = [0.1, 0.2, 0.5]

[Nordholm et al., *JASA*, 2018]

2. Acoustic feedback cancellation

- **Spatial filtering:** reduce acoustic feedback in the vent microphone by steering a (robust) spatial null towards the hearing aid receiver
- Perfect feedback cancellation:

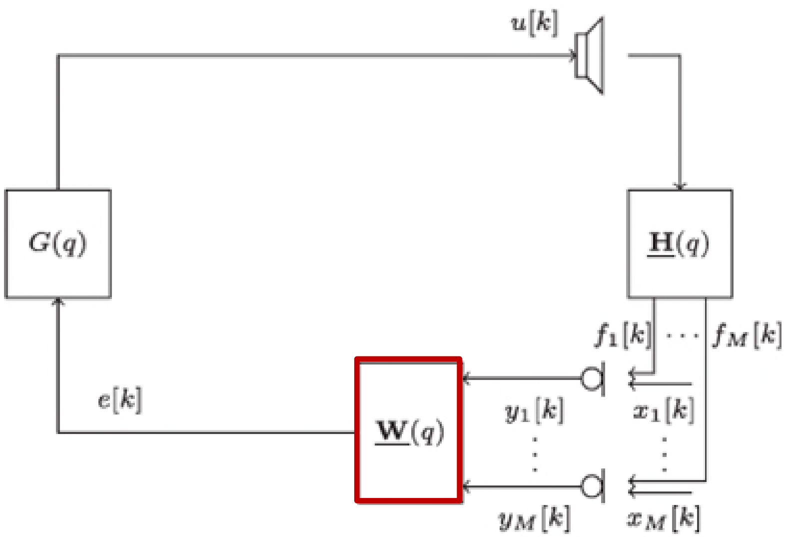
$$\underline{\mathbf{W}}^T(q)\underline{\mathbf{H}}(q) = 0$$

feedback paths

- Incoming signal preservation:

$$\underline{\mathbf{W}}^T(q)\underline{\tilde{\mathbf{D}}}(q) = q^{-L_d}$$

relative transfer functions for incoming source



[Schepker et al., *IEEE/ACM TASLP,* 2020]

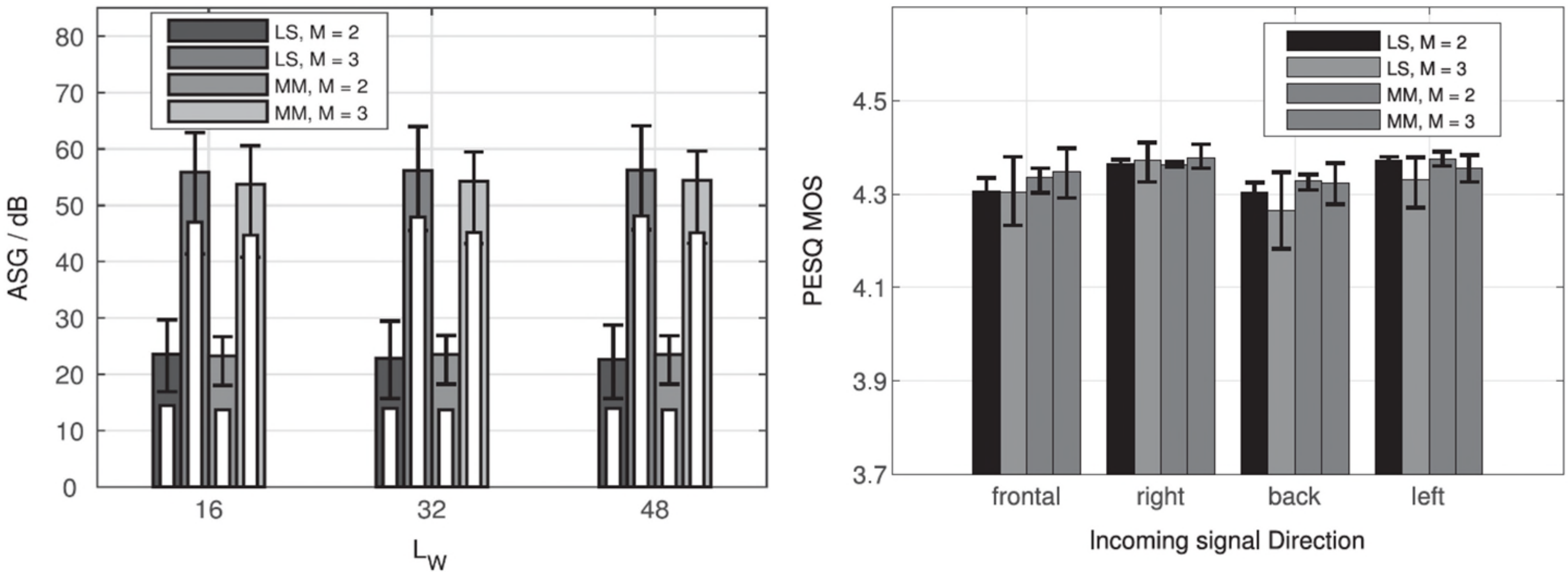
2. Acoustic feedback cancellation

- **Robust design procedures for fixed beamformer** based on (multiple) measurements of acoustic feedback paths

	Minimize residual feedback power	Maximize maximum stable gain
Hard constraint	$\min_{\mathbf{w}} \sum_{i=1}^{I} \|(\mathbf{H}^{(i)})\mathbf{w}\|_2^2$ subject to $\tilde{\mathbf{D}}^{(j)}\mathbf{w} = \tilde{\mathbf{e}}_{L_d} \quad \forall j = 1, \ldots, J$ *Constrained least-squares problem*	$\min_{\mathbf{w}} \max_{\omega_n, i} \|(\underline{\mathbf{H}}^{(i)})^H(\omega_n)\underline{\mathbf{W}}(\omega_n)\|^2$ subject to $\tilde{\mathbf{D}}^{(j)}\mathbf{w} = \tilde{\mathbf{e}}_{L_d} \quad j = 1, \ldots, J$ *Linear programming problem*
Soft constraint	$\min_{\mathbf{w}} \|\bar{\mathbf{H}}\mathbf{w}\|_2^2 + \lambda\|\bar{\mathbf{D}}\mathbf{w} - \bar{\mathbf{e}}_{L_d}\|_2^2$ *Unconstrained least-squares problem*	$\min_{\mathbf{w}} \max_{\omega_n, i} \|(\underline{\mathbf{H}}^{(i)})^H(\omega_n)\underline{\mathbf{W}}(\omega_n)\|^2 + \lambda\|\bar{\mathbf{D}}\mathbf{w} - \bar{\mathbf{e}}_{L_d}\|_2^2$ *Quadratically constrained quadratic programming (QCQP) problem*

[Schepker et al., *IEEE/ACM TASLP*, 2020]

2. Acoustic feedback cancellation



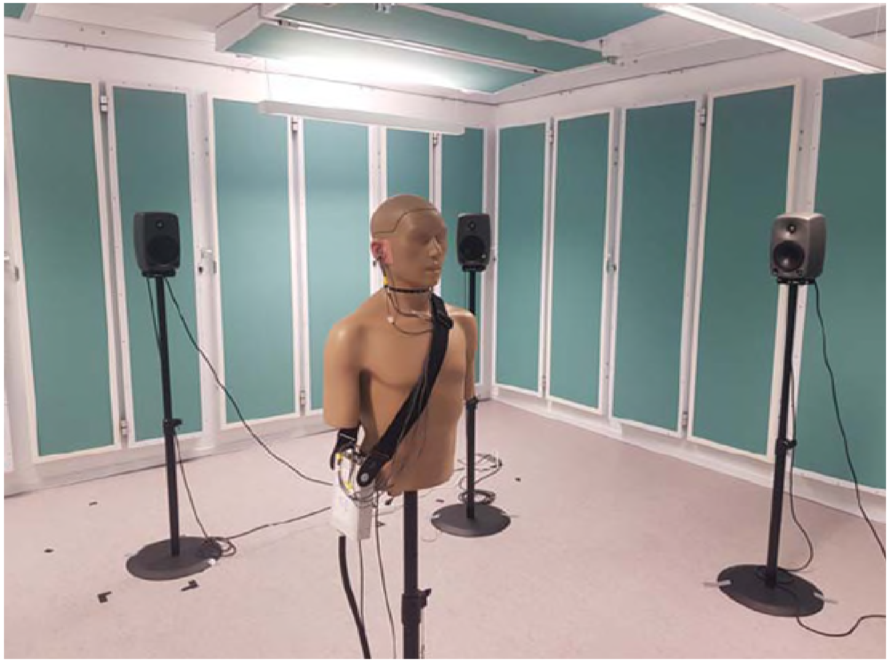
Robust reduction of acoustic feedback of up to 50dB while hardly distorting incoming signal

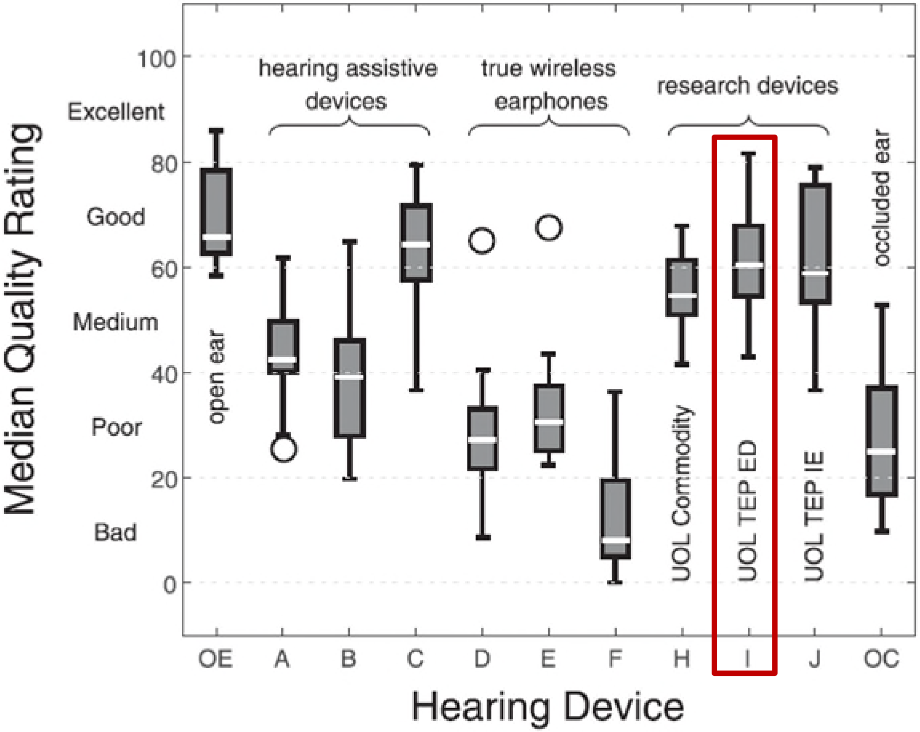
f_s = 16 kHz, T_{60} = 300 ms,
I = 20, 4 source directions,
forward path: τ = 6 ms,
G_0 = 45 dB, ε_{MSG} = 10 dB

[Schepker et al., *IEEE/ACM TASLP*, 2020]

Acoustically transparent sound presentation

- **Acoustic transparency feature compared to six commercial hearables**

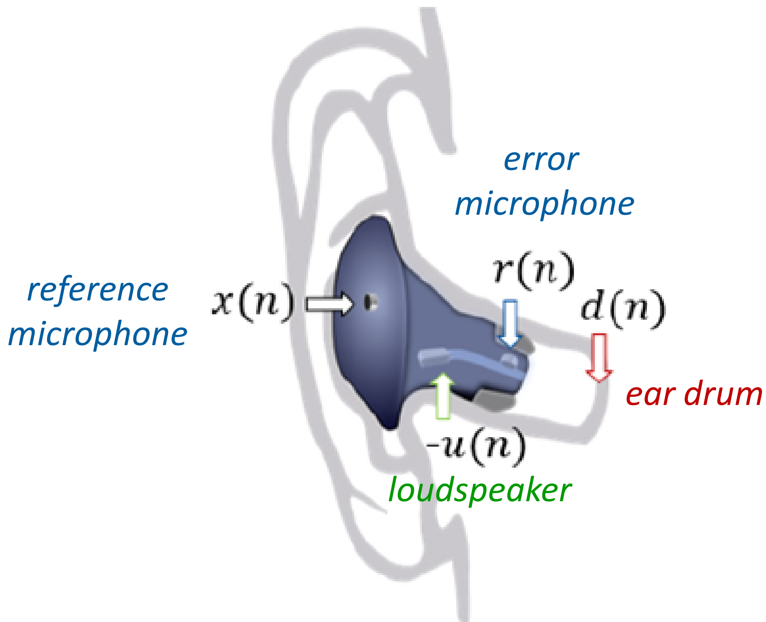




[Schepker et al., *JAES,* 2020]

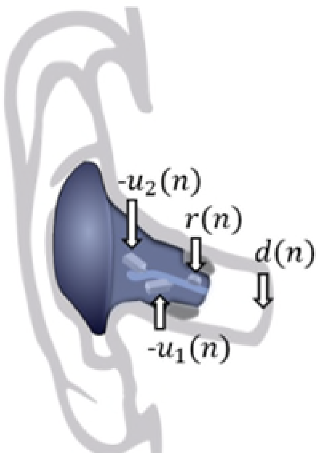
3. Active noise control

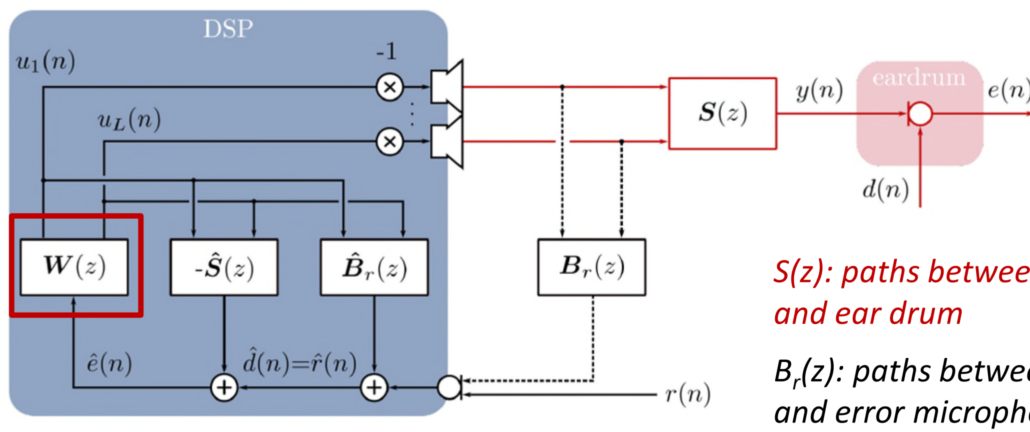
- **Aim:** play back anti-noise by one or more loudspeakers aiming at generating “zone of quiet” at certain position (e.g. ear drum)
- **Approaches:**
  - **Feedforward ANC:** filter reference microphone signal
  - **Feedback ANC:** filter error microphone signal



3. Active noise control

- Virtual sensing fixed feedback ANC exploiting **multiple loudspeakers**, aiming at minimizing sound pressure at **ear drum**





S(z): paths between loudspeakers and ear drum

$B_r(z)$: paths between loudspeakers and error microphone

- Minimize power spectral density of sound pressure at ear drum

$$\Phi_{ee}(f) = \left(1 - \frac{|\Phi_{dr}(f)|^2}{\Phi_{dd}(f)\Phi_{rr}(f)}\right)\Phi_{dd}(f) + \left|\frac{\Phi_{dr}(f)}{\Phi_{rr}(f)} - \frac{\boldsymbol{W}^{\mathrm{T}}(f)\boldsymbol{S}(f)}{1 + \boldsymbol{W}^{\mathrm{T}}(f)\left(\hat{\boldsymbol{S}}(f) + \boldsymbol{B}_r(f) - \hat{\boldsymbol{B}}_r(f)\right)}\right|^2 \Phi_{rr}(f)$$

Virtual microphone arrangement → $d(n) = r(n)$

$$\Phi_{ee}^{\mathrm{vma}}(f) = \left|\frac{1}{1 + \boldsymbol{W}^{\mathrm{T}}(f)\hat{\boldsymbol{S}}(f)}\right|^2 \Phi_{rr}(f)$$

[Rivera Benois et al., *ICASSP*, 2022]

3. Active noise control

- Turn non-convex optimization problem into **convex optimization problem**

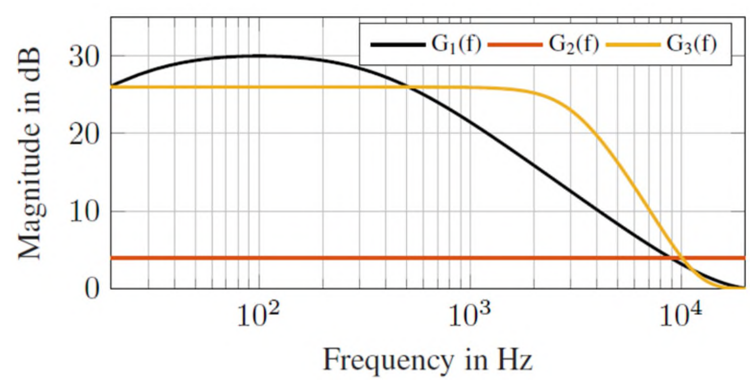
$$\min_{\boldsymbol{w}} \left| \frac{1}{1+\boldsymbol{W}^{\mathrm{T}}(f)\widehat{\boldsymbol{S}}(f)} \right|^2 \phi_{rr}(f) \quad \Longrightarrow \quad \max_{\mathbf{w}} \sum_{k=0}^{L_{\mathrm{DFT}}/2-1} \left|1+\boldsymbol{W}^{\mathrm{T}}(\Omega_k)\hat{\boldsymbol{S}}(\Omega_k)\right|^2 G_1^2(\Omega_k)$$

subject to constraints (stability, amplification, gain)

$$|\varrho - \boldsymbol{W}^{\mathrm{T}}(\Omega_k)\hat{\boldsymbol{S}}(\Omega_k)| \leq |\varrho + \boldsymbol{W}^{\mathrm{T}}(\Omega_k)\hat{\boldsymbol{S}}(\Omega_k)| + 2\rho, \forall \Omega_k$$

$$1 \leq G_2(\Omega_k)|1 + \boldsymbol{W}^{\mathrm{T}}(\Omega_k)\hat{\boldsymbol{S}}(\Omega_k)|, \forall \Omega_k$$

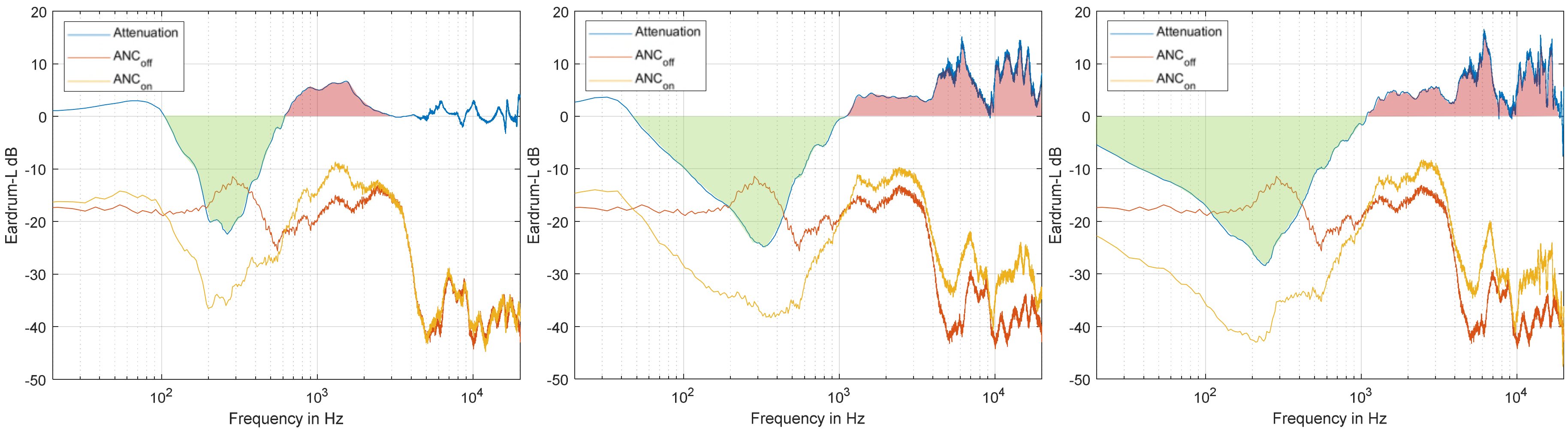
$$W_l(\Omega_k) \leq G_3(\Omega_k), \forall \Omega_k$$



[Rivera Benois et al., *ICASSP*, 2022]

3. Active noise control

	Classic	Virtual Microphone	Virtual Microphone
Loudspeakers	Single	Single	Multi
Available acoustic paths	$\hat{B}_r(z)$	$\hat{B}_r(z)$ $\hat{S}(z)$	$\hat{\boldsymbol{B}}_r(z)$ $\hat{\boldsymbol{S}}(z)$

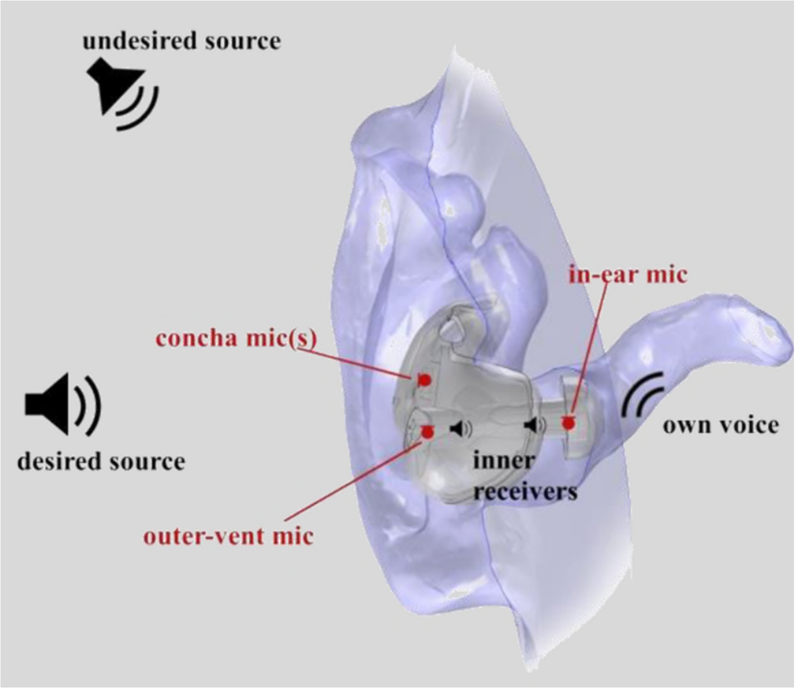


Multi-loudspeaker virtual sensing ANC approach improves attenuation magnitude and bandwidth

[Rivera Benois et al., *ICASSP*, 2022]

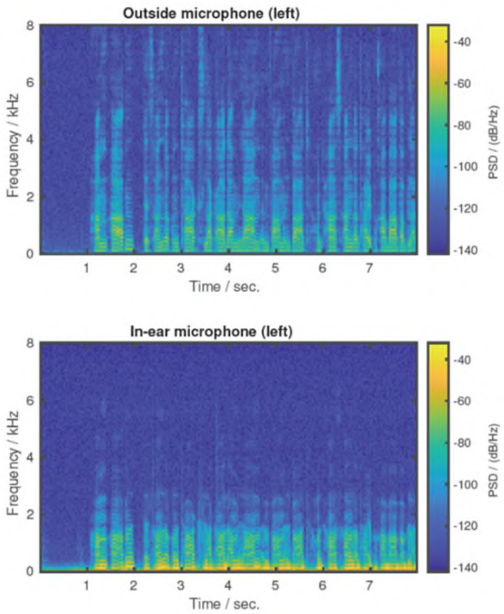
4. Own voice extraction

- **Aim:** enhance own voice of user wearing earpiece in noisy acoustic environment (e.g. industrial workplace)
- **Approach:** exploit in-ear microphone, possibly in combination with outer microphone(s)

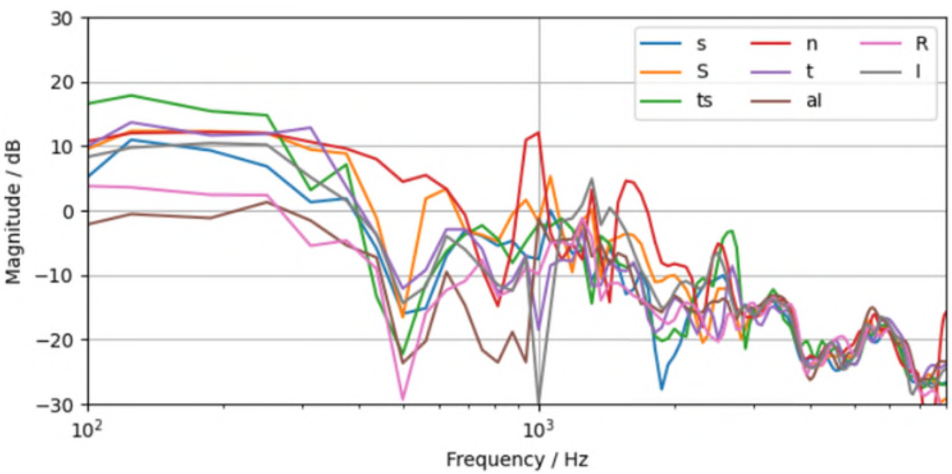


4. Own voice extraction

- **Different characteristics** for own voice and external noise at in-ear and outer microphones
  - **Outer microphone:** full bandwidth own voice, possibly low SNR (external noise)
  - **In-ear microphone:** bandlimited own voice (up to about 2 kHz), high SNR (external noise), body noise

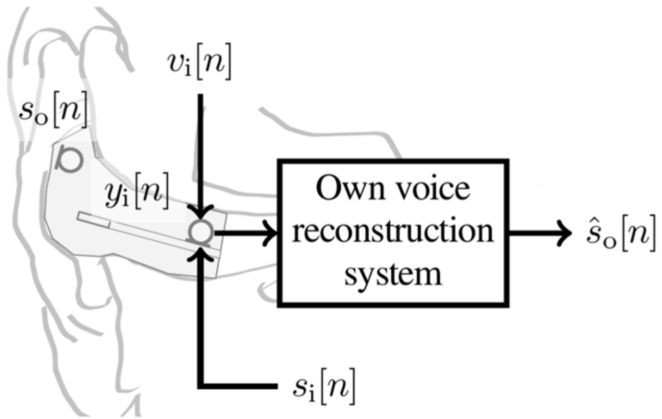


- **Relative transfer characteristics for own voice**
  - Time-varying (speech-dependent)
  - User- and device-dependent

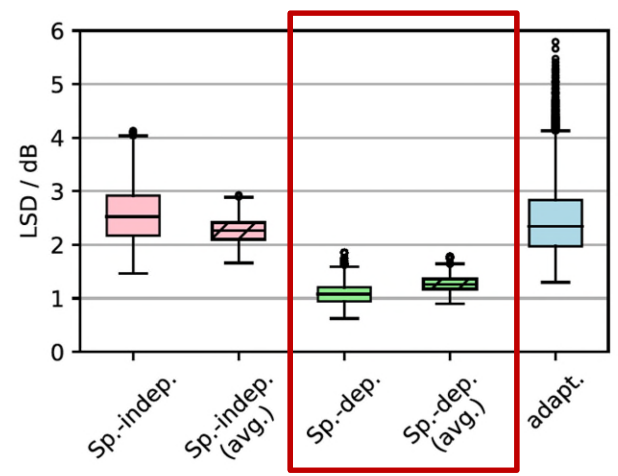


4. Own voice extraction

- **Objectives of algorithm:** estimate clean speech signal at outer microphone from
  - **in-ear microphone:** combined bandwidth extension, equalization and noise reduction (body + external noise)
  - **in-ear and outer microphone**



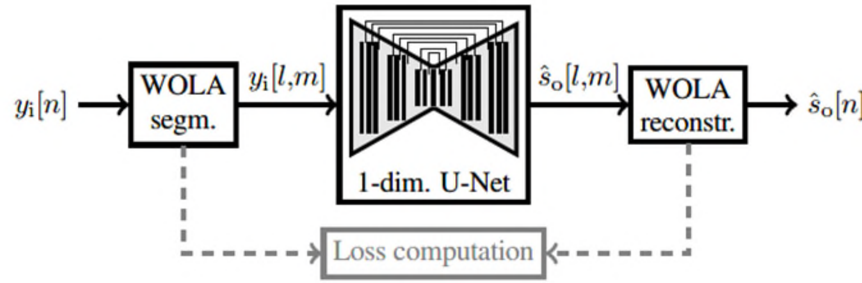
- **Limited training data available**:
  - use **models** to generate simulated data (data augmentation):
    - Fixed relative transfer function (sp.-indep.)
    - Phoneme-dependent relative transfer function (sp.dep.)
  - **domain transfer** (train with simulated data, fine-tune with real recordings)

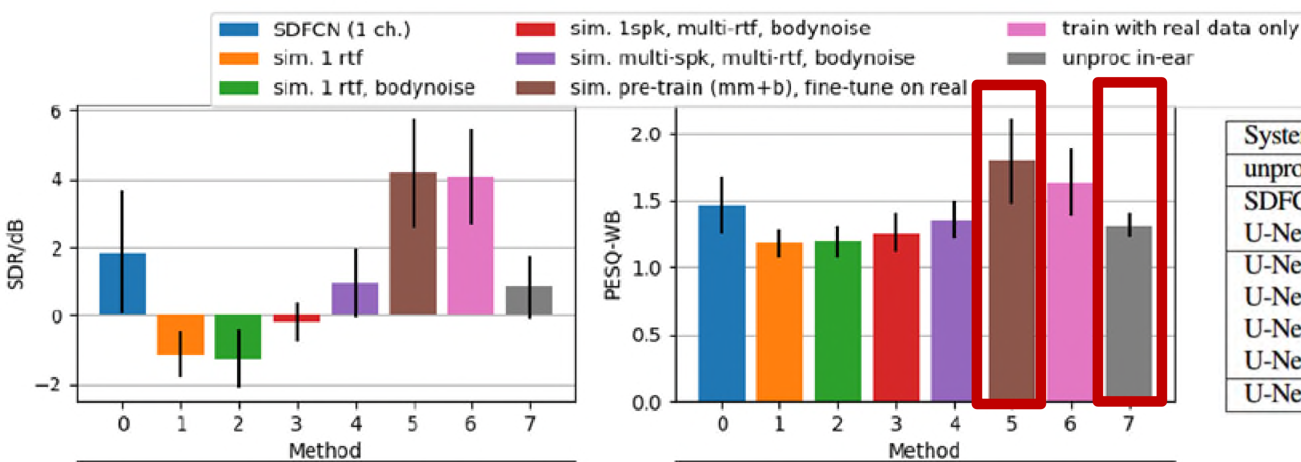


[Ohlenbusch et al., *Forum Acusticum,* 2023]

4. Own voice extraction

- **Results:**
  - Based on bandwidth extension system using U-Net architecture [Wang 2021]
  - Only exploiting in-ear microphone signal, only body noise (no external noise)
  - **Different training procedures** (real data, simulated data with one/multiple RTFs and one/multiple talkers, simulated data + fine-tuning with real data)



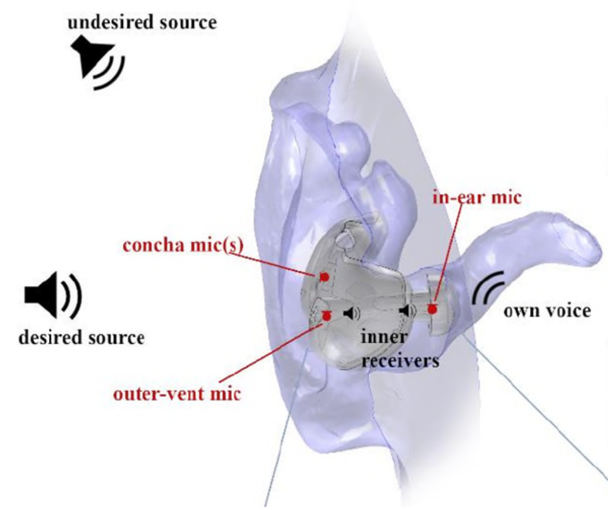


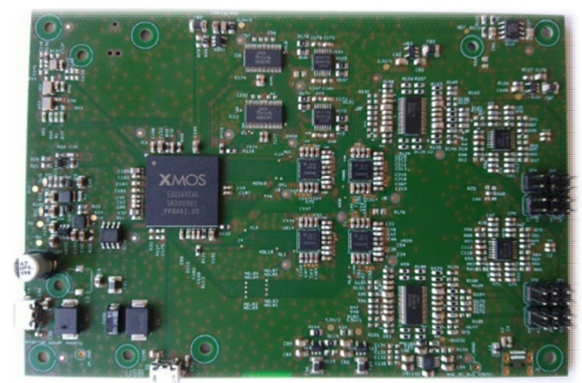
System	Data	RTFs used	LSD / dB	PESQ
unproc.	-	-	2.51	1.31
SDFCN	[R]	-	1.53	1.47
U-Net	[R]	-	1.48	1.64
U-Net	[S]	1T, s-RTF	1.35	1.18
U-Net	[S+]	1T, s-RTF	1.54	1.19
U-Net	[S+]	1T, m-RTF	1.51	1.26
U-Net	[S+]	14T, m-RTF	1.24	1.36
U-Net	[S+R]	14T, m-RTF	**1.05**	**1.80**

[Ohlenbusch et al., *IWAENC,* 2022]

Current / Future work

- **Direction-selective** acoustic transparency and active noise control, steered by CASA
- **Deep learning-based active noise control**
- **Speech communication exploiting in-ear microphone** (phoneme-dependent own voice models, DNN-based algorithms)
- Individualized and phoneme-dependent occlusion models and **active occlusion reduction**
- Implementation on **low-latency processor** (cooperation with Fraunhofer HSA)
- **Integration** with speech enhancement algorithms and self-adjusted hearing support





Acknowledgments / references

- S. Vogl and M. Blau, “Individualized prediction of the sound pressure at the eardrum for an earpiece with integrated receivers and microphones,” *The Journal of the Acoustical Society of America*, vol. 145, no. 2, pp. 917-930, Feb. 2019.
- H. Schepker et al., “Null-steering beamformer-based feedback cancellation for multi-microphone hearing aids with incoming signal preservation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 4, pp. 679-691, Apr. 2019.
- F. Denk et al., “A one-size-fits-all earpiece with multiple microphones and drivers for hearing device research,” in *Proc. AES Conference on Headphone Technology*, San Francisco, USA, Aug. 2019.
- H. Schepker, S. Nordholm, S. Doclo, “Acoustic feedback suppression for multi-microphone hearing devices using a soft-constrained null-steeering beamformer,” *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 28, pp. 929-940, 2020.
- F. Denk, H. Schepker, S. Doclo, B. Kollmeier, “Acoustic Transparency in Hearables - Technical Evaluation,” *Journal of the Audio Engineering Society*, vol. 68, no. 7/8, pp. 508–521, Jul./Aug. 2020.
- H. Schepker, F. Denk, B. Kollmeier, S. Doclo, “Acoustic Transparency in Hearables - Perceptual Sound Quality Evaluations,” *Journal of the Audio Engineering Society*, vol. 68, no. 7/8, pp. 495–507, Jul./Aug. 2020.
- F. Denk and B. Kollmeier (2021). The Hearpiece database of individual transfer functions of an in-the-ear earpiece for hearing device research.
- H. Schepker, F. Denk, B. Kollmeier, S. Doclo, “Robust single- and multi-loudspeaker least-squares-based equalization for hearing devices,” *EURASIP Journal of Audio, Speech and Music Processing,* 2022.
- P. Rivera Benois, R. Roden, M. Blau, S. Doclo, “Optimization of a Fixed Virtual Sensing Feedback ANC Controller for In-Ear Headphones with Multiple Loudspeakers,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* Singapore, May 2022.
- M. Ohlenbusch, C. Rollwage, S. Doclo, “Training Strategies for Own Voice Reconstruction in Hearing Protection Devices using an In-ear Microphone,” in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC*), Bamberg, Germany, Sep. 2022.
- M. Ohlenbusch, C. Rollwage, S. Doclo,“Speech-dependent Modeling of Own Voice Transfer Characteristics for In-ear Microphones in Hearables,” in *Proc. Forum Acusticum*, Torino, Italy, Sep. 2023.

DFG Deutsche Forschungsgemeinschaft

HAPPAA

Hearing 4all

DAAD Deutscher Akademischer Austausch Dienst German Academic Exchange Service

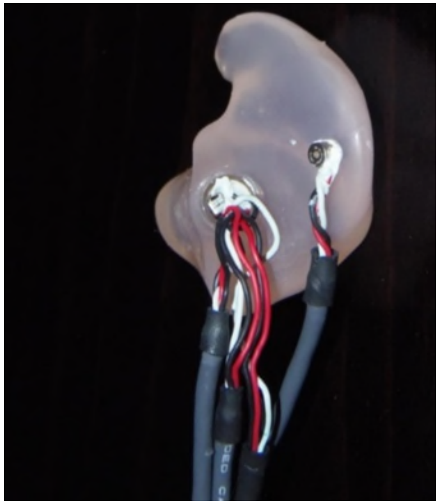
Questions ?

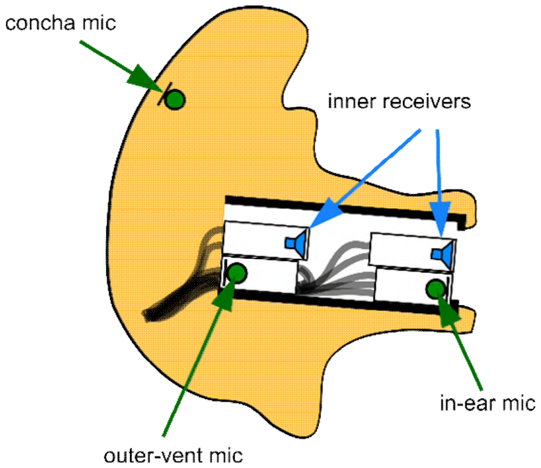
http://www.sigproc.uni-oldenburg.de

YouTube Signal Processing Uni Oldenburg

Acoustically Transparent Earpiece (v1)

- **Custom in-the-ear earpiece** with multiple integrated microphones and receivers and relatively open acoustics
  - Vent/core: **2 microphones and 2 receivers** (woofer/tweeter)
  - Concha: **1 microphone**
- Insertion into individual silicone ear mould or generic earplugs

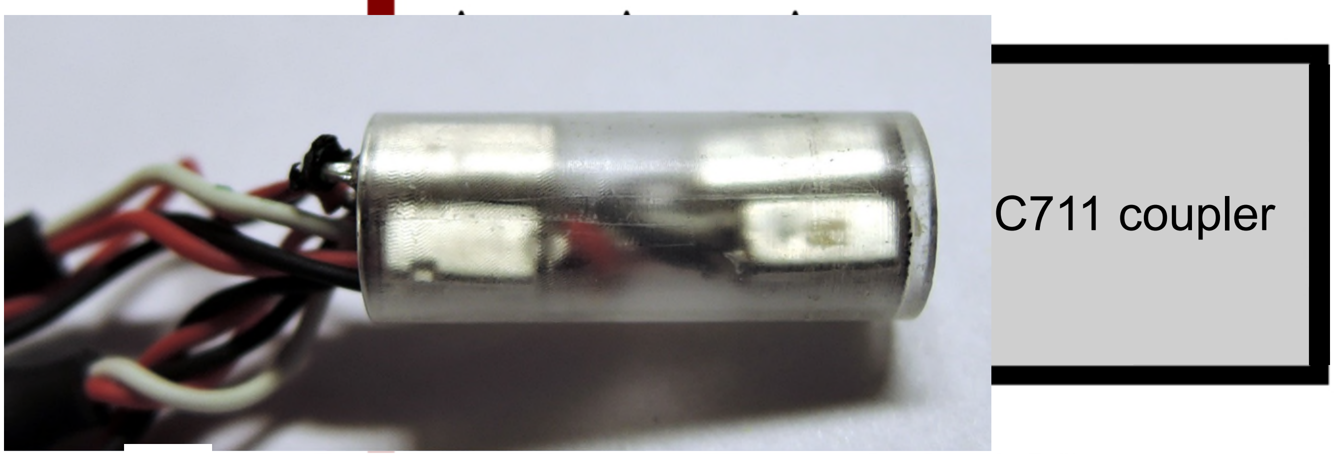


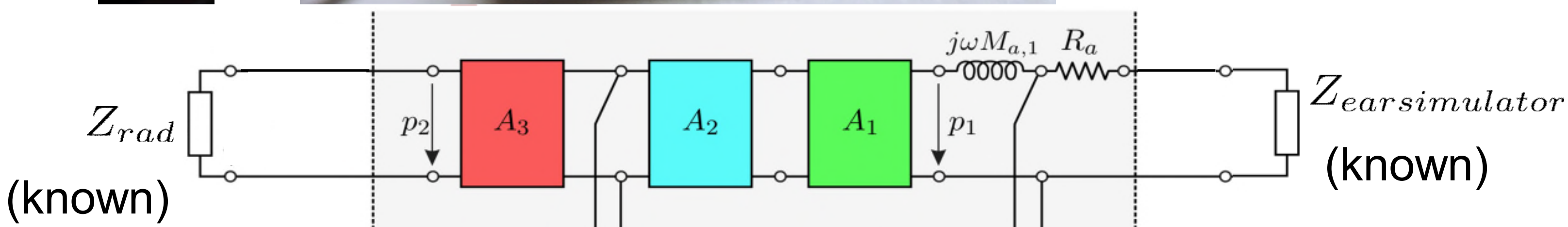


[Denk et al., *International Journal of Audiology*, 2018]

Electro-acoustic model

- Earpiece Model (Fixed)



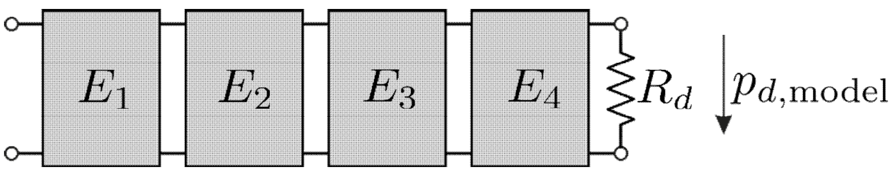


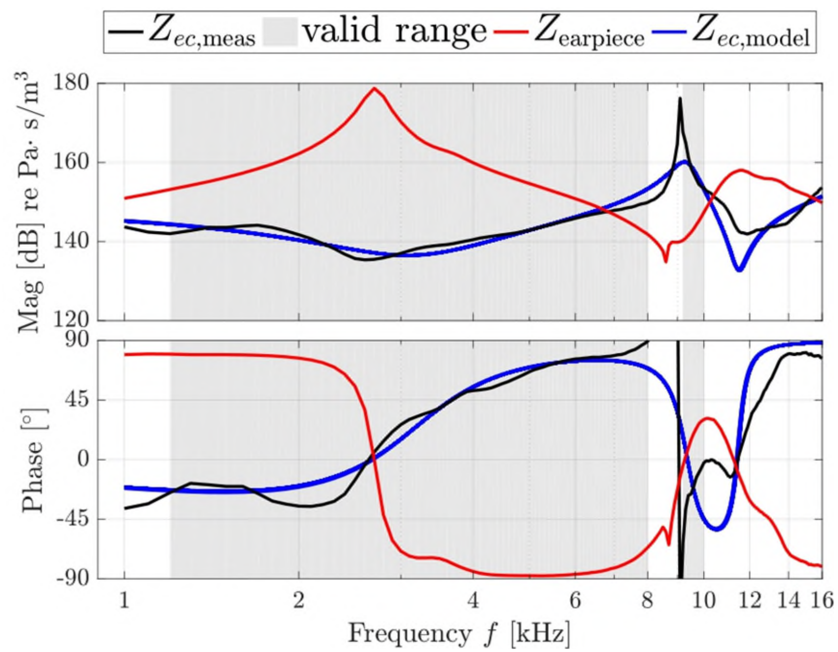
[Vogl and Blau 2019, JASA]

Parameter optimization: minimize difference between all 4 measured and modeled transfer functions

Electro-acoustic model

- Ear Canal Model (Individualized)





Parameter optimization (4 radii, 1 length, 1 resistive load) by minimizing the difference between measured and modeled ear canal (Nelder-Mead simplex optimization procedure):

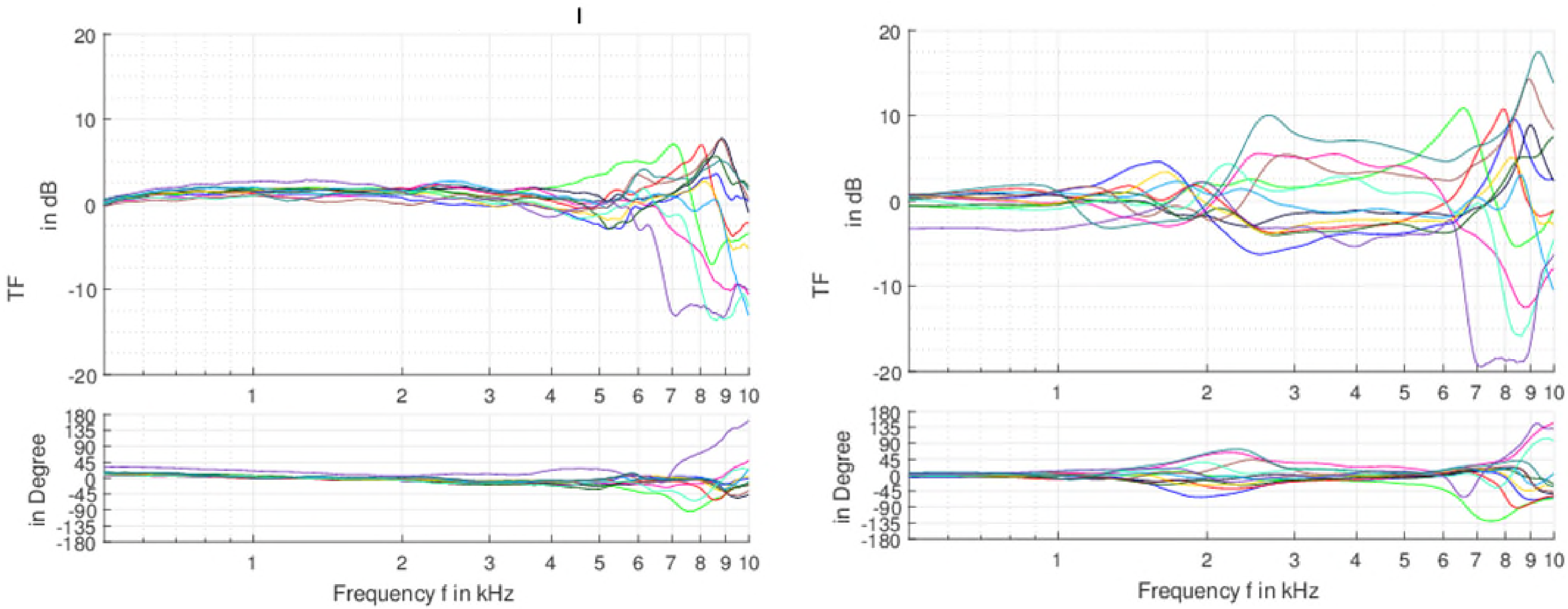
$$J(p) = \sum_{f_{\text{valid}}} \left(db(Z_{ec,\text{meas}}) - db(Z_{ec,\text{model}}(p))\right)^2 + 10 \cdot \left(arg(Z_{ec,\text{meas}}) - arg(Z_{ec,\text{model}}(p))\right)^2$$

Electro-acoustic model

- Evaluation (sound pressure at ear drum) for 12 subjects

$$\frac{p_{d,\text{Model (i.e. individualized)}}}{p_{d,\text{Probe Tube Meas}}}$$

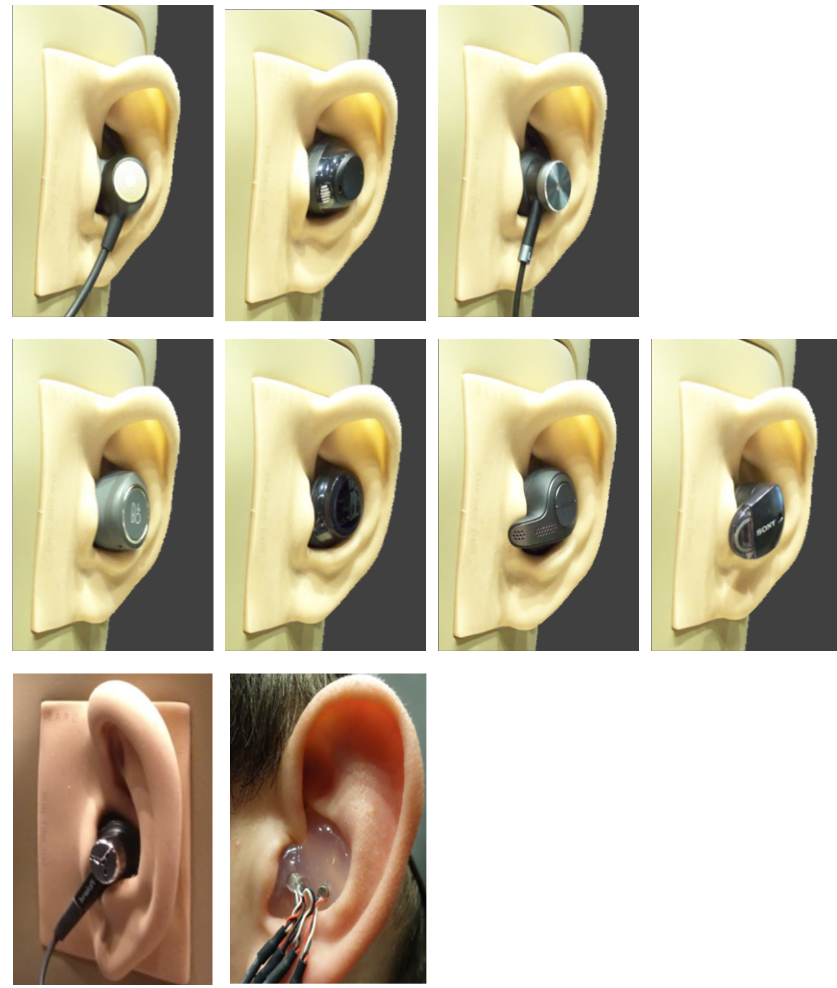
$$\frac{p_{d,\text{from average correction of p1 (i.e. non-individualized)}}}{p_{d,\text{Probe Tube Meas}}}$$



accurate prediction of sound pressure at ear drum possible using individualized electro-acoustic model up to about 6 kHz

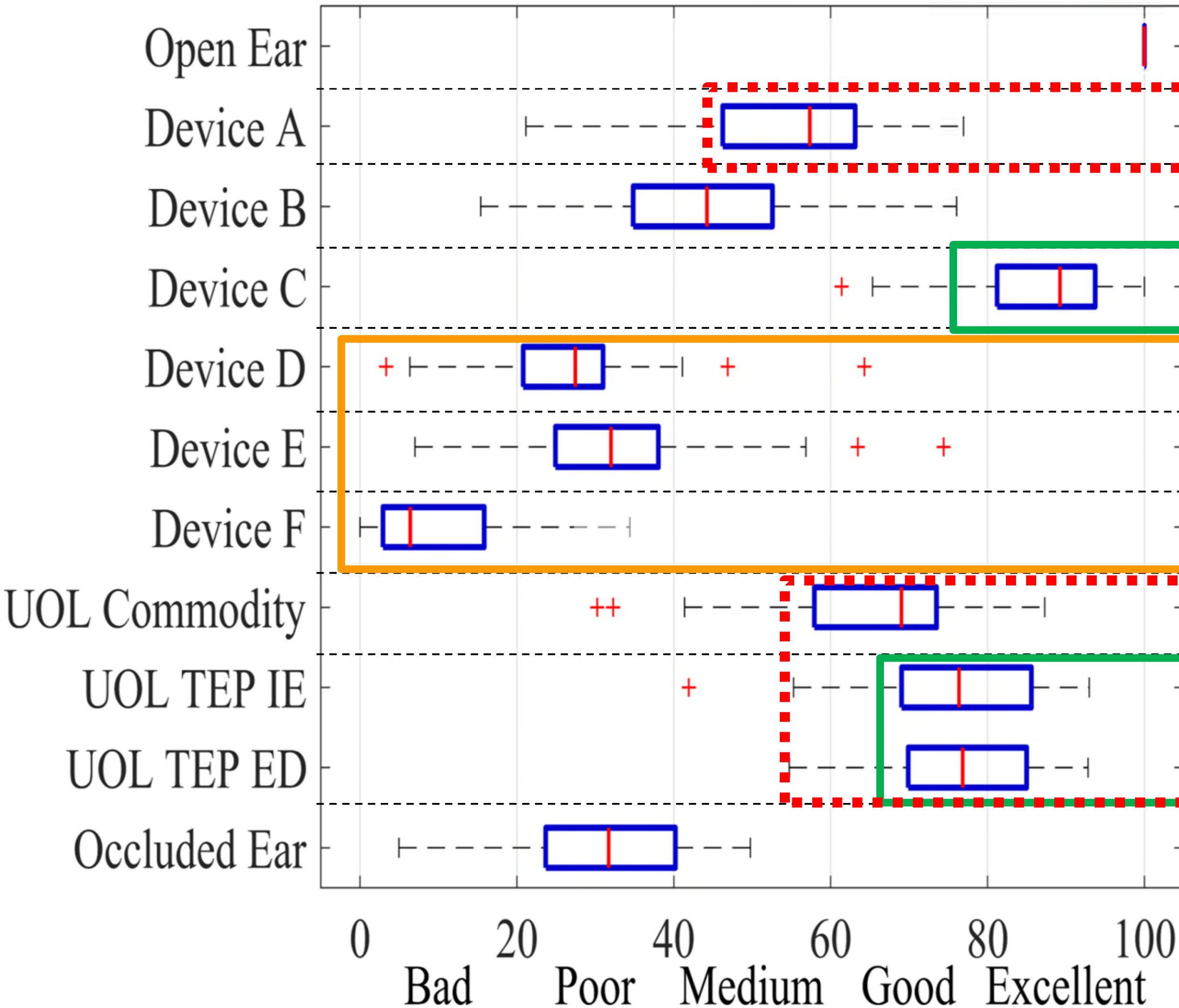
Evaluation of hear-through feature

- **7 commercial hearables**
  - 3 hearing support: **Devices A-C**
  - 4 wireless earbuds: **Devices D-G**
- **2 research prototypes**
  - UOL Commodity: consumer hardware based hearing aid prototype [Schädler 2017, Buhl, Denk et al. 2019]
  - UOL Acoustically Transparent Earpiece: Adaptation to individual ear acoustics [Denk et al. 2018, Schepker, Denk et al. 2019]



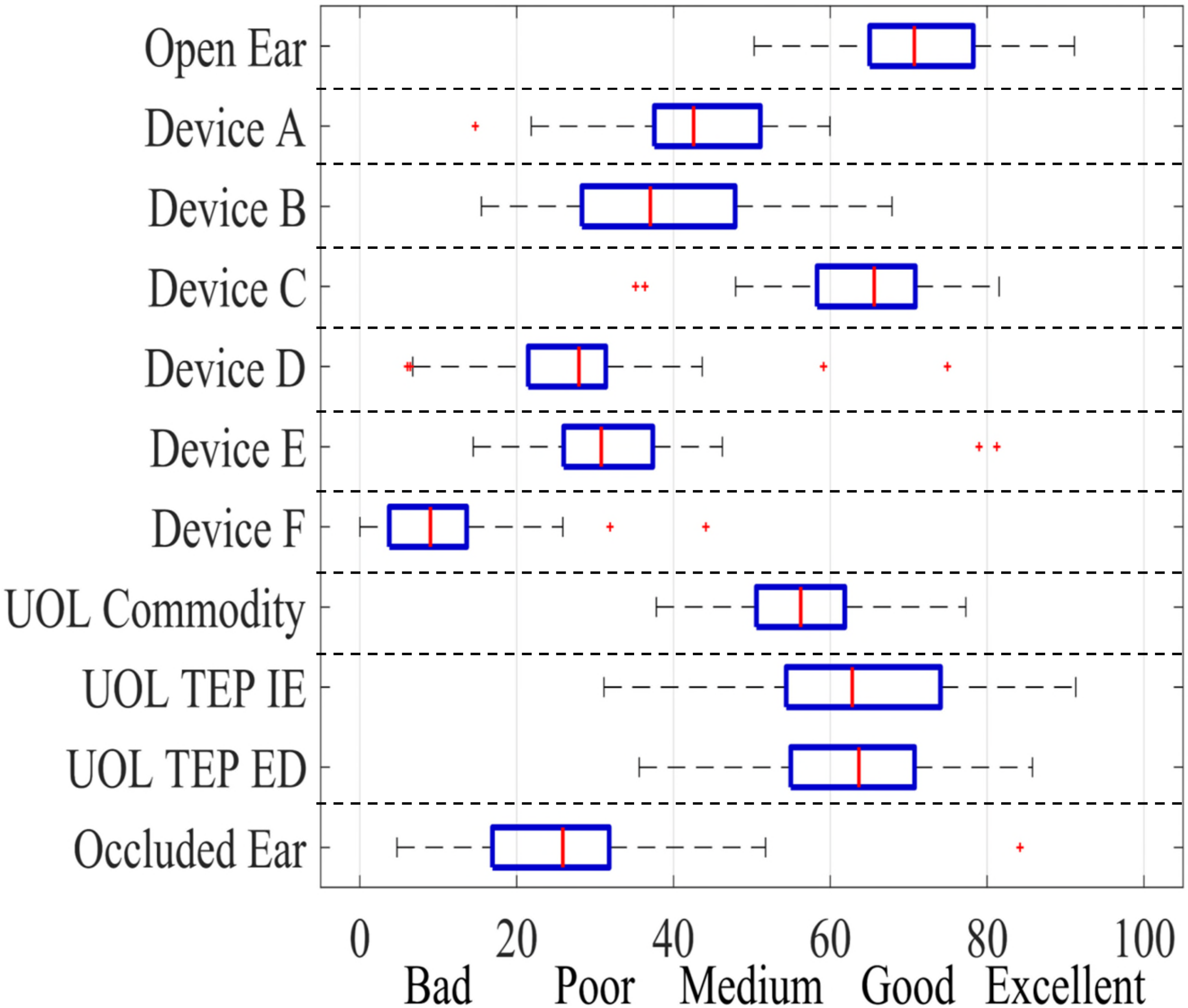
[Denk et al., J. Audio Eng. Soc., 2020] [Schepker et al., J. Audio Eng. Soc., 2020]

Results: Open-ear reference given



	Response Match	Comb Filters
Open Ear		
Device A	+	-
Device B	-	+
Device C	++	++
Device D	--	++
Device E	-	++
Device F	--	-
UOL Commodity	+	-
UOL TEP IE / UOL TEP ED	++	-
Occluded Ear	--	++

Results: *No* reference given



	Response Match	Comb Filters
Device A	+	-
Device B	-	+
Device C	++	++
Device D	--	++
Device E	-	++
Device F	--	-
UOL Commodity	+	-
UOL TEP IE / UOL TEP ED	++	-
Occluded Ear	--	++

[Schepker et al., J. Audio Eng. Soc., 2020]