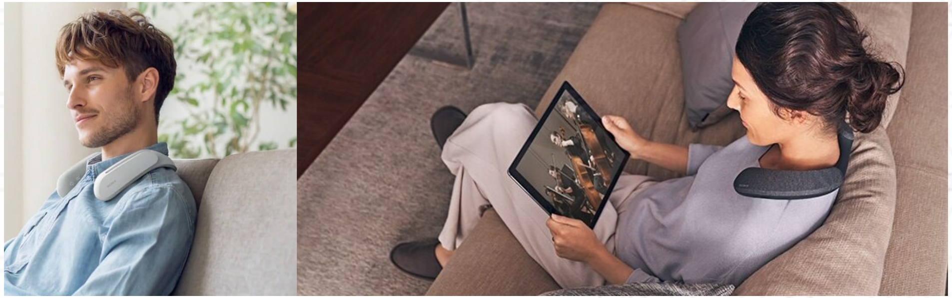
Personal active soundfield control (using wearable audio devices)

University of Maryland, College Park - Department of Computer Science
Shoken Kaneko, Irtaza Shahid, Nirupam Roy, Nail A. Gumerov, Ramani Duraiswami

UNIVERSITY OF MARYLAND

Emerging new mode of audio devices: “wearable loudspeakers”



* From SONY web site

* From BOSE web site

- Unique mode of audio presentation which is open and personal at the same time (which can be a dilemma)
- What can we do with this new unique device?

Sound control in open space

- Active sound control is expanding from closed spaces (ANC headphones / earphones) to loudspeaker arrays in open space

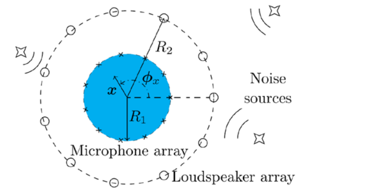
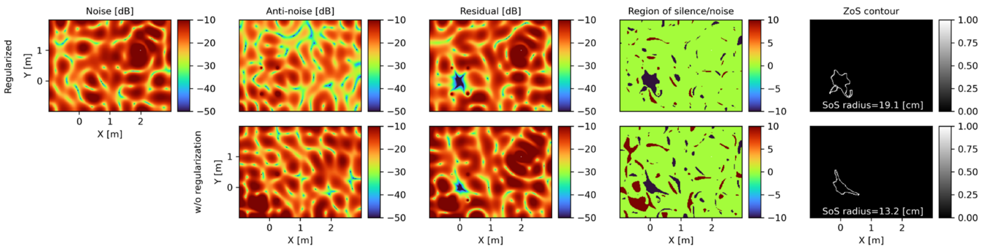


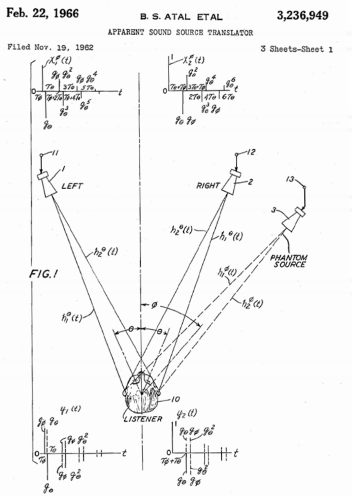
Fig. 1. A spatial ANC region (black) consists of a circular microphone array of radius R_1 and a circular loudspeaker array of radius R_2.



Zhang, J., Abhayapala, T. D., Zhang, W., Samarasinghe, P. N., & Jiang, S. (2018). "Active noise control over space: A wave domain approach". *IEEE/ACM Transactions on audio, speech, and language processing*, *26*(4), 774-786.

Kaneko, S., Roy, N., Gumerov, N., & Duraiswami, R. (2021, August). "Regularized spherical harmonics-domain spatial active noise cancellation in a reverberant room". In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 263, No. 2, pp. 4733-4742). Institute of Noise Control Engineering.

- Crosstalk cancellation (XTC; transaural audio playback)



Atal, Bishnu S., and Manfred R. Schroeder. "Apparent sound source translator." U.S. Patent No. 3,236,949. 22 Feb. 1966.

Personal active soundfield control

- Using wearable <u>open</u> audio devices (loudspeaker / microphone arrays) to actively control <u>personal</u> sound fields.
- Control targets:
  - Outgoing fields:
    - Speech privacy
  - Incoming fields:
    - Personal active noise control
  - Local fields:
    - Immersive AR audio presentation

$$p^{(\mathrm{tot})} = p^{(\mathrm{rad})} + p^{(\mathrm{ext})} + p^{(\mathrm{can})} + p^{(\mathrm{tra})}$$

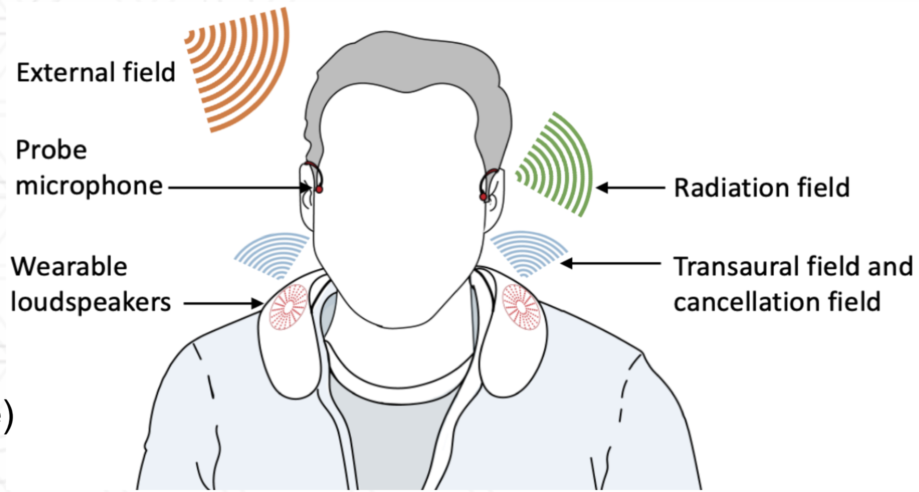
Total field

Radiation field

External field

Cancellation field (anti-noise)

Transaural field



Formulation: costs and solutions

- Task 1: Active radiation cancellation (ARC)

$$L_\lambda^{(\mathrm{ARC})} = ||\mathbf{p}_{\mathrm{test}}^{(\mathrm{rad})} + H_{\mathrm{ts}}\mathbf{c}^{(\mathrm{can})}||_2^2 + \lambda ||\mathbf{c}^{(\mathrm{can})}||_2^2 \Rightarrow \mathbf{c}_{\mathrm{opt}}^{(\mathrm{can})} = -(H_{\mathrm{ts}}^H H_{\mathrm{ts}} + \lambda I)^{-1} H_{\mathrm{ts}}^H \mathbf{p}_{\mathrm{test}}^{(\mathrm{rad})}$$

- Task 2: Personal open ANC

$$L_\lambda^{(\mathrm{ANC})} = ||\mathbf{p}_{\mathrm{probe}}^{(\mathrm{res})}||_2^2 + \lambda ||\mathbf{c}^{(\mathrm{can})}||_2^2. \Rightarrow \mathbf{c}_{\mathrm{opt}}^{(\mathrm{can})} = -(H_{\mathrm{ms}}^H H_{\mathrm{ms}} + \lambda I)^{-1} H_{\mathrm{ms}}^H \mathbf{p}_{\mathrm{probe}}^{(\mathrm{ext})}$$

- Task 3: XTC

$$L_\lambda^{(\mathrm{XTC})} = ||\mathbf{s}^{(\mathrm{bin})} - H_{\mathrm{rtf}} H_{\mathrm{xtc}} \mathbf{s}^{(\mathrm{bin})}||_2^2 + \lambda ||H_{\mathrm{xtc}} \mathbf{s}^{(\mathrm{bin})}||_2^2 \Rightarrow H_{\mathrm{xtc}} = (H_{\mathrm{rtf}}^H H_{\mathrm{rtf}} + \lambda I)^{-1} H_{\mathrm{rtf}}^H$$

Global cost function for hyperparameter optimization

- Global cost function (for task 1):

$$L^{(\mathrm{glob})}(\lambda) = -R(\mathbf{p}_{\mathrm{opt}})$$

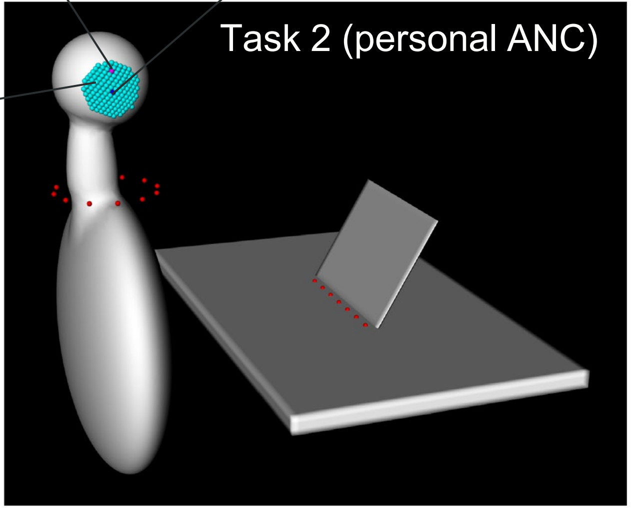
$$R(\mathbf{p}_Q) = 20(\log_{10}||\mathbf{p}_Q^{(\mathrm{rad})}||_2 - \log_{10}||\mathbf{p}_Q^{(\mathrm{res})}||_2)$$

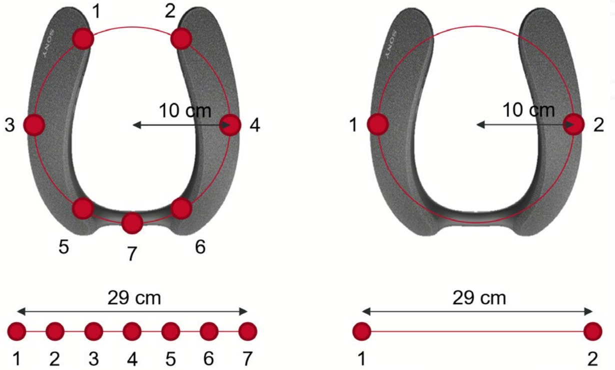
… reduction level on “optimization points” Q, a set of points which is distinct from the test (control) points

- Hyper parameter λ is optimized by grid search to minimize the global cost function

Numerical experiment

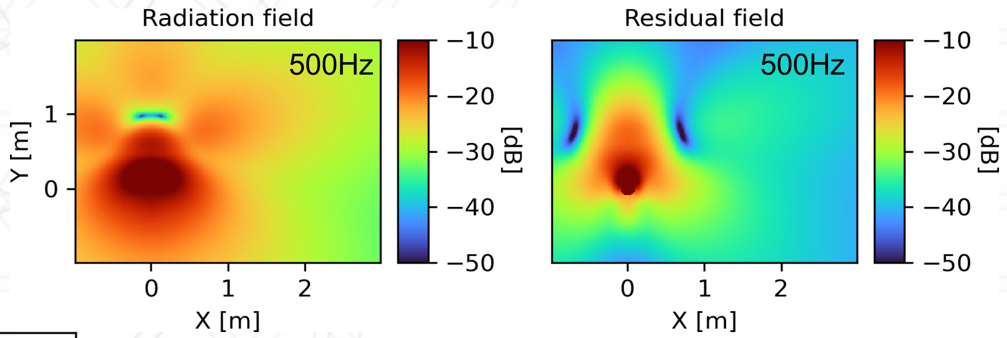
- Geometry:
  - A humanoid in front of a desk with a laptop screen on top
  - Loudspeakers:
    - neckband loudspeaker array or loudspeaker array installed in laptop (2ch or 7ch)
- Transfer function computation:
  - Fast-multipole BEM

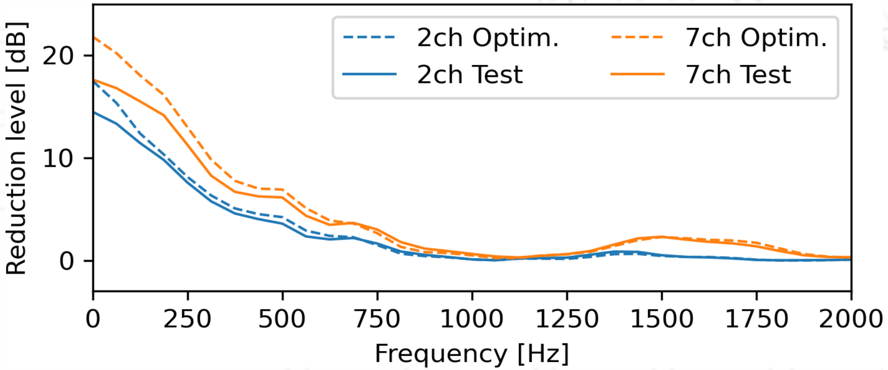


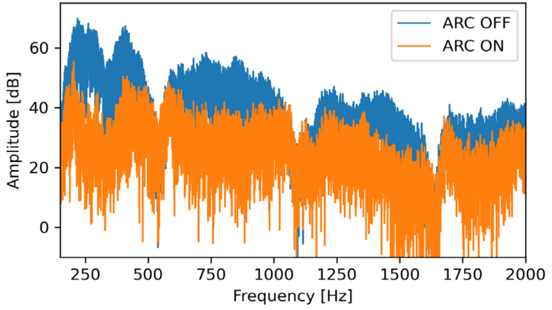


Results: task 1 - speech privacy

- Point source in the vicinity of the mouth
- Test & optimization points: 2048 points (spherical Fibonacci grid) at distance 0.5m and 0.7m, respectively.
- Reduction observed at low frequencies (< 1 kHz)



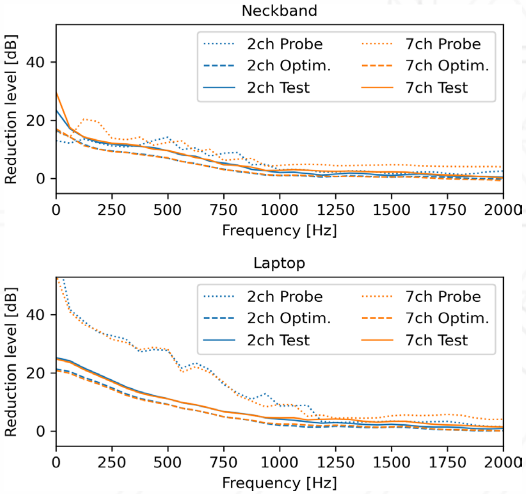


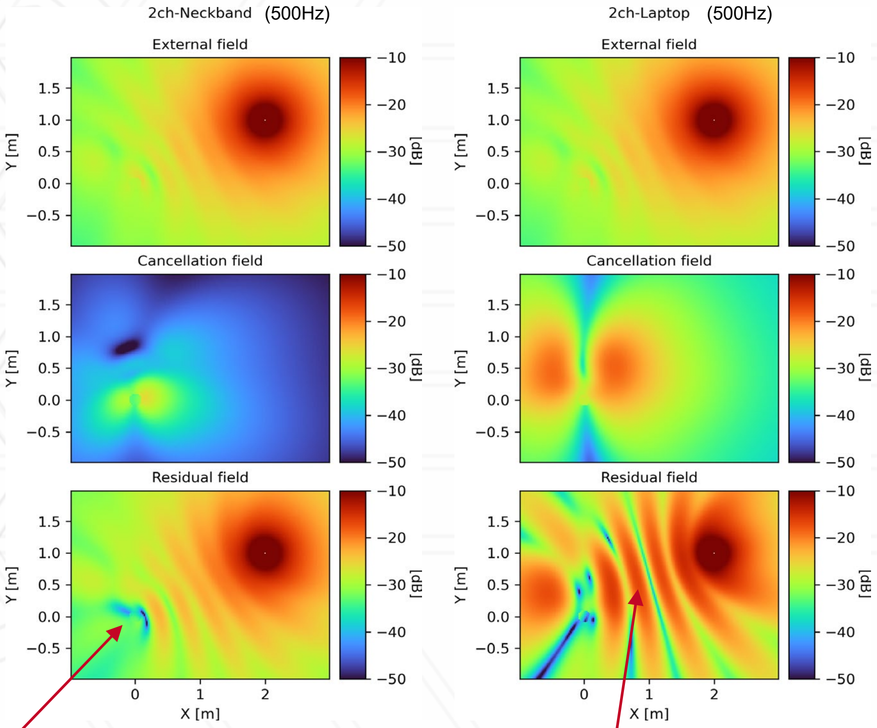


Simulation using measured IRs and human speech as input

Results: task 2 - personal open ANC

- Similar reduction observed at low frequencies (< 1 kHz) for both worn and frontal loudspeakers
- Spatial characteristics are very different



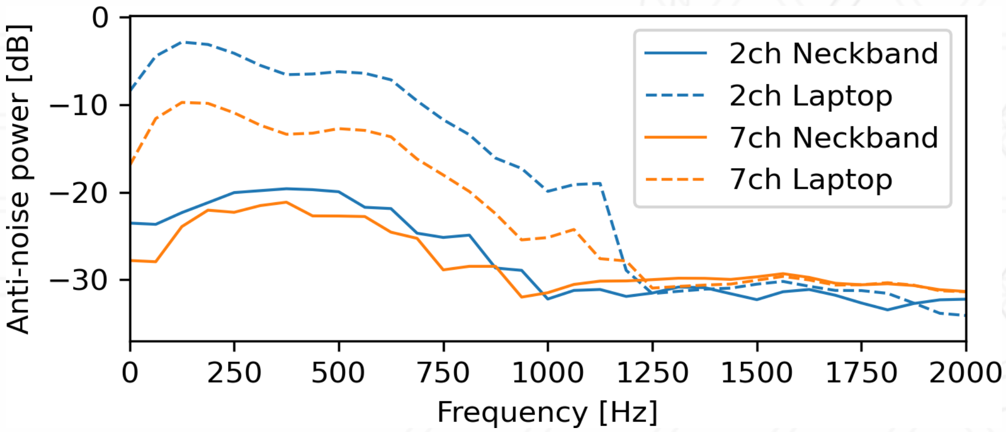


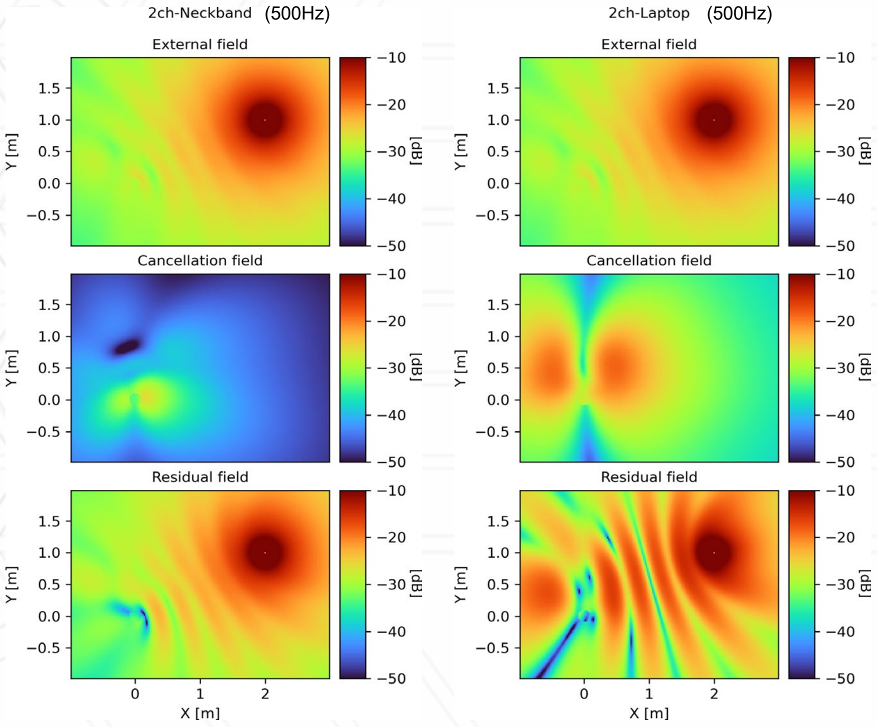
Reduction in the vicinity of the ears

Significant noise injection into the space

Results: task 2 - personal open ANC

- While the reduction level was similar, worn loudspeakers required significantly less energy





Conclusion

- Two personal active sound control tasks (personal ANC & ARC) using a wearable loudspeaker array have been studied by means of numerical simulation.
- Successful control was observed at low frequencies (< 1kHz)
- Wearable loudspeakers were found to achieve personal ANC with less energy compared to frontal loudspeaker arrays
  - This can be understood as a result of the proximity of the loudspeakers and the head/neck being natural acoustic obstacles reducing crosstalk
    → good indication for reducing the noise injection side-effect and for the XTC task
  - Wearable loudspeakers seem to solve the “personal-while-open” objective better
- Future research:
  - Developing signal processing which considers nonlinearities of the small loudspeakers in wearable devices
  - Content-dependent smart control of radiated/incident/local sounds using statistical learning from data

End

Thank you very much for your attention!

UNIVERSITY OF MARYLAND

Shoken Kaneko

8125 Paint Branch Dr., College Park, MD 20742

kaneko60@umd.edu