UNIVERSITY OF
Southampton

Optimal Source Distribution for Multiple Listener Virtual Sound Imaging

P.A. Nelson, P. Couturier and X. Zhou
ISVR University of Southampton

T. Takeuchi
Opsodis Ltd.

3rd March 2022

3D sound with loudspeakers

B.B. Bauer, Stereophonic earphones and binaural loudspeakers, J. Audio Eng. Soc. 9 (2) (1961) 148–151.

B.S. Atal, M.R. Schroeder, United States Patent 3,236,949A, Apparent sound source translator, 1966.

P. Damaske, Head-related two-channel stereophony with loudspeaker reproduction, J. Acoust. Soc. Am. 50 (4B) (1971) 1109–1115.

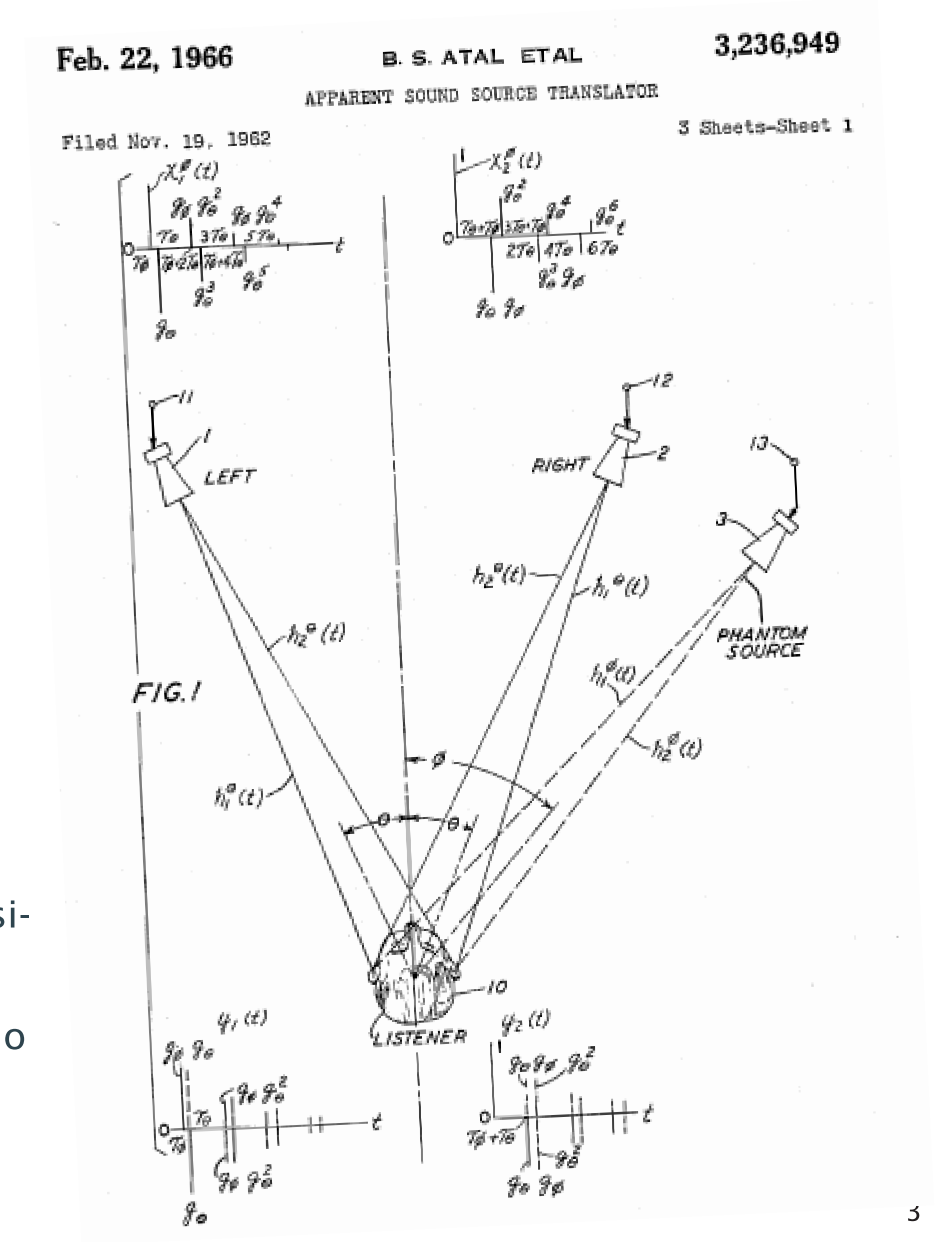
Y. Ando, S. Shidara, Z. Maekawa, K. Kido, Some basic studies on the acoustic design of room by computer, J. Acoust. Soc. Japan 29 (1973).

S. Sakamoto, T. Gotoh, T. Kogure, M. Shimbo, A. Clegg, Controlling sound-image localisation in stereophonic reproduction, J. Audio Eng. Soc. 29 (11) (1981) 794–799.

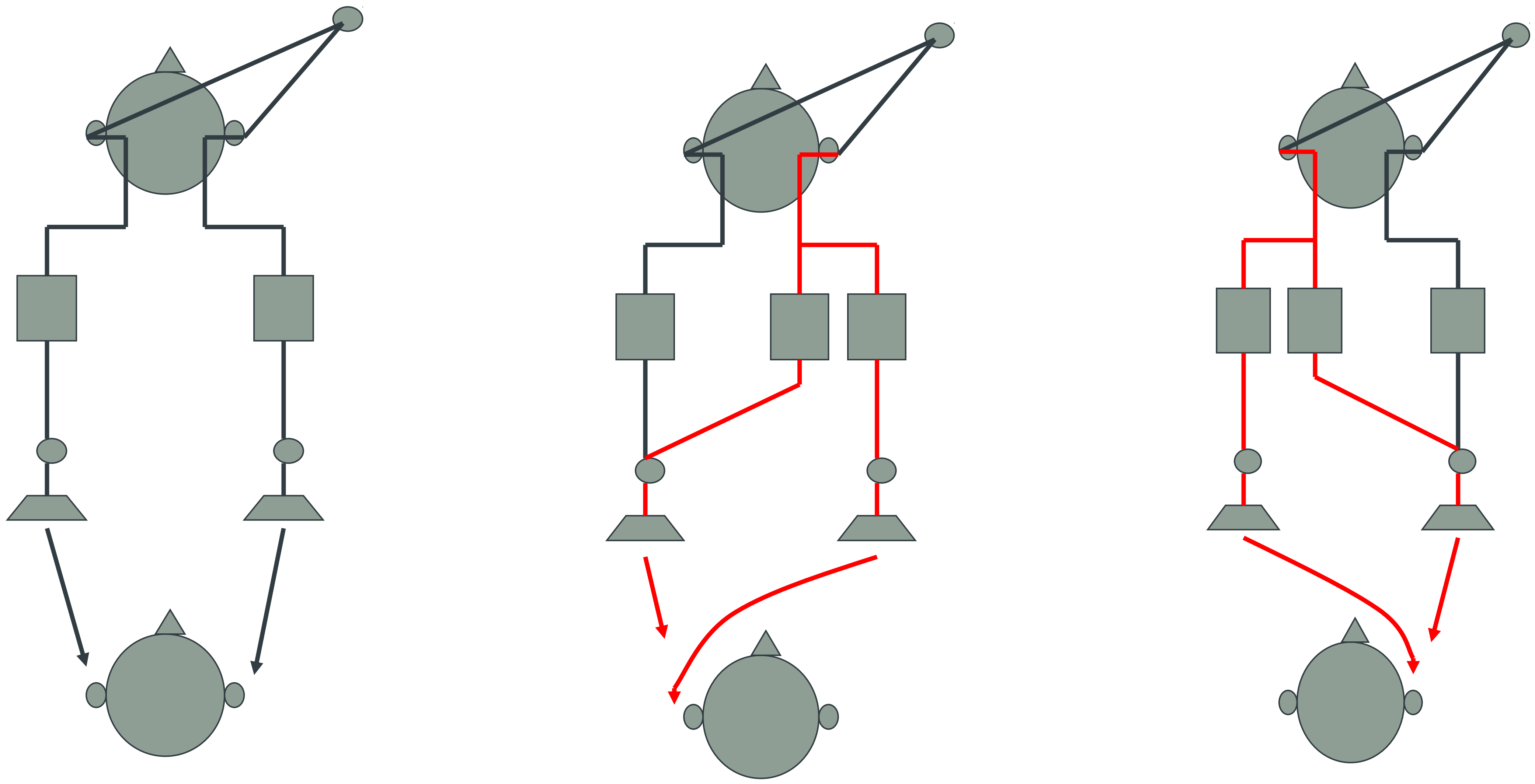
H. Hamada, Construction of orthosterophonic system for the purposes of quasi-insitu recording and reproduction, J. Acoust. Soc. Am. 39 (5) (1983) 337–348.

J.L. Bauck, D.H. Cooper, Generalised transaural stereo and applications, J. Audio Eng. Soc. 44 (9) (1996) 683–705.

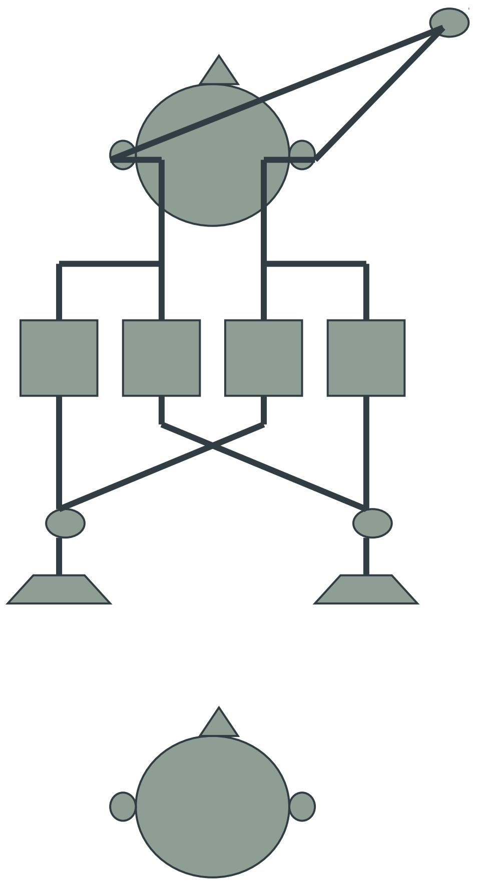
See also E.C. Hamdan. Theoretical Advances in Multichannel Crosstalk Cancellation Systems . PhD Thesis University of Southampton (2020)

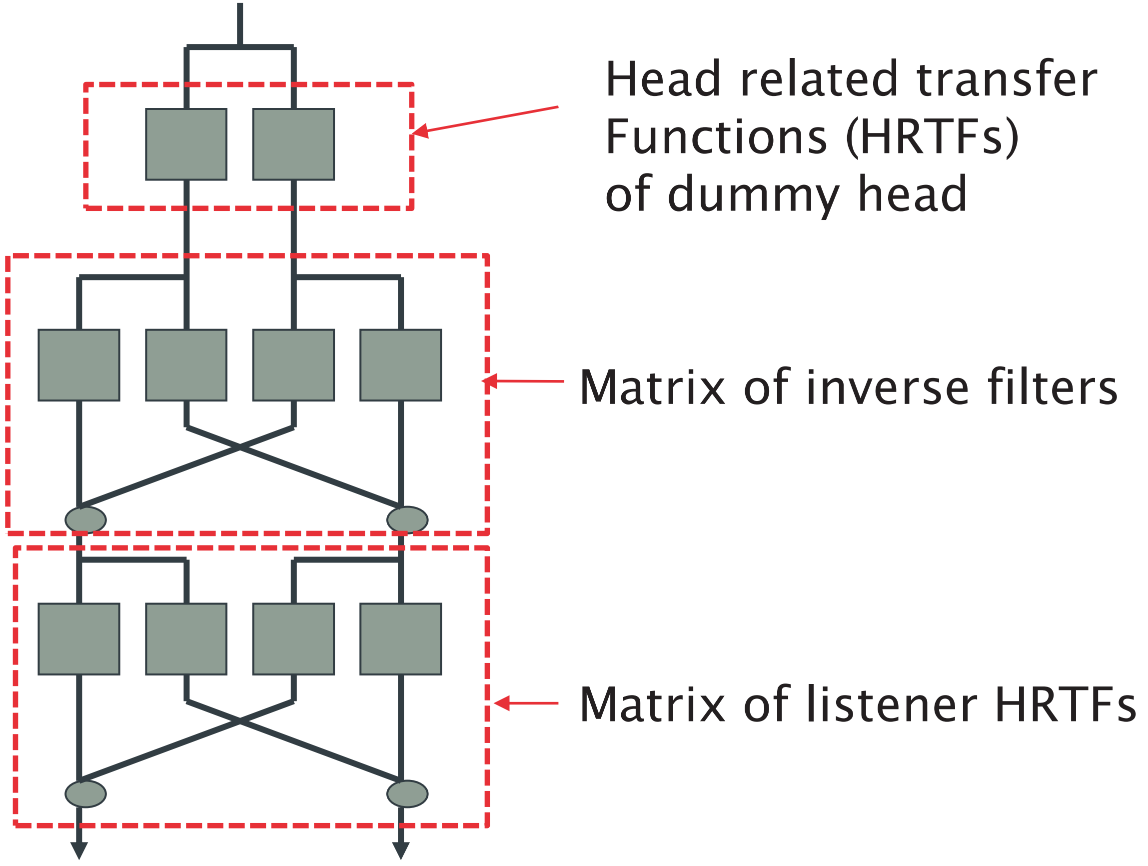


Crosstalk cancellation for 3D sound

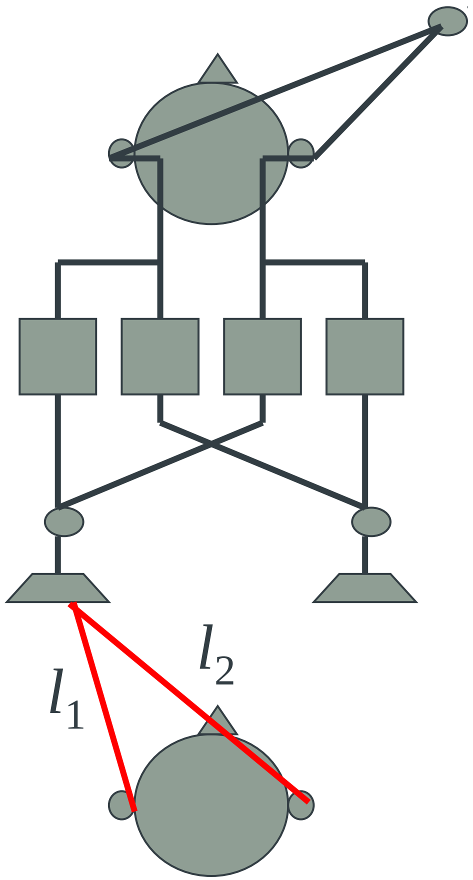


Block diagram

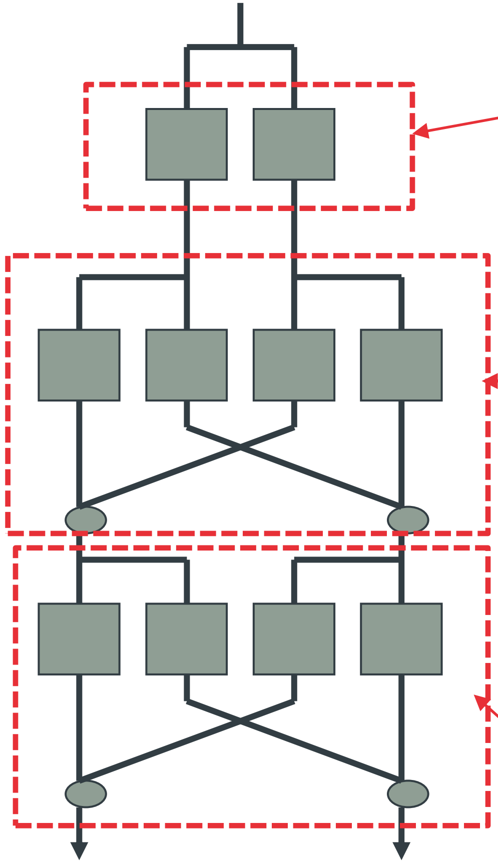




The half wavelength problem



Source signal



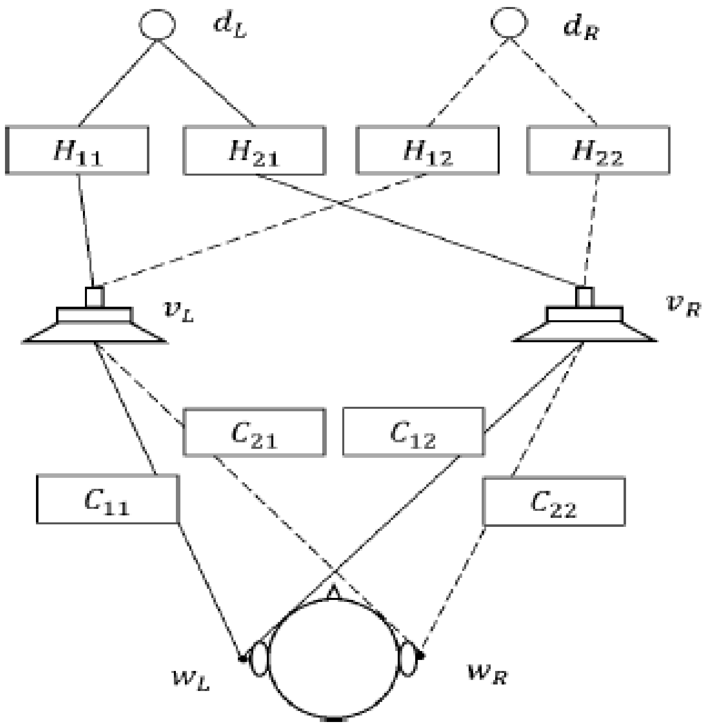
Head related transfer Functions (HRTFs) of dummy head

Matrix of inverse filters

Matrix of listener HRTFs becomes ill-conditioned when

$$l_2 - l_1 = \lambda / 2$$

Two source single listener geometry

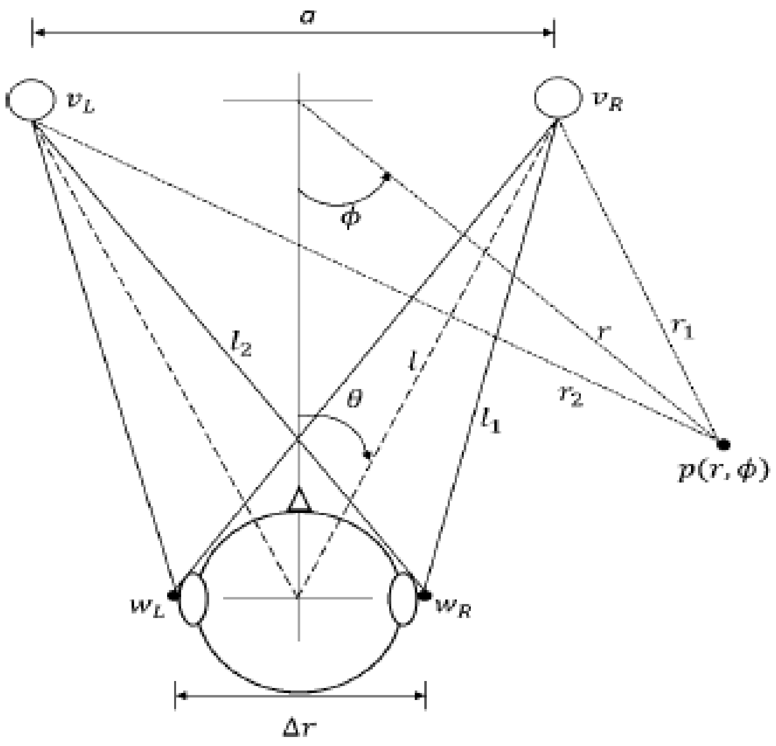


$$\mathbf{d}^{\mathrm{T}} = \begin{bmatrix} d_R & d_L \end{bmatrix}, \quad \mathbf{v}^{\mathrm{T}} = \begin{bmatrix} v_R & v_L \end{bmatrix}, \quad \mathbf{w}^{\mathrm{T}} = \begin{bmatrix} w_R & w_L \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

$$\mathbf{v} = \mathbf{Hd} \qquad \mathbf{w} = \mathbf{CHd}$$

Solution for the inverse filter matrix



$$\mathbf{C} = \frac{\rho_0 e^{-jkl_1}}{4\pi l_1} \begin{bmatrix} 1 & ge^{-jk\Delta l} \\ ge^{-jk\Delta l} & 1 \end{bmatrix}$$

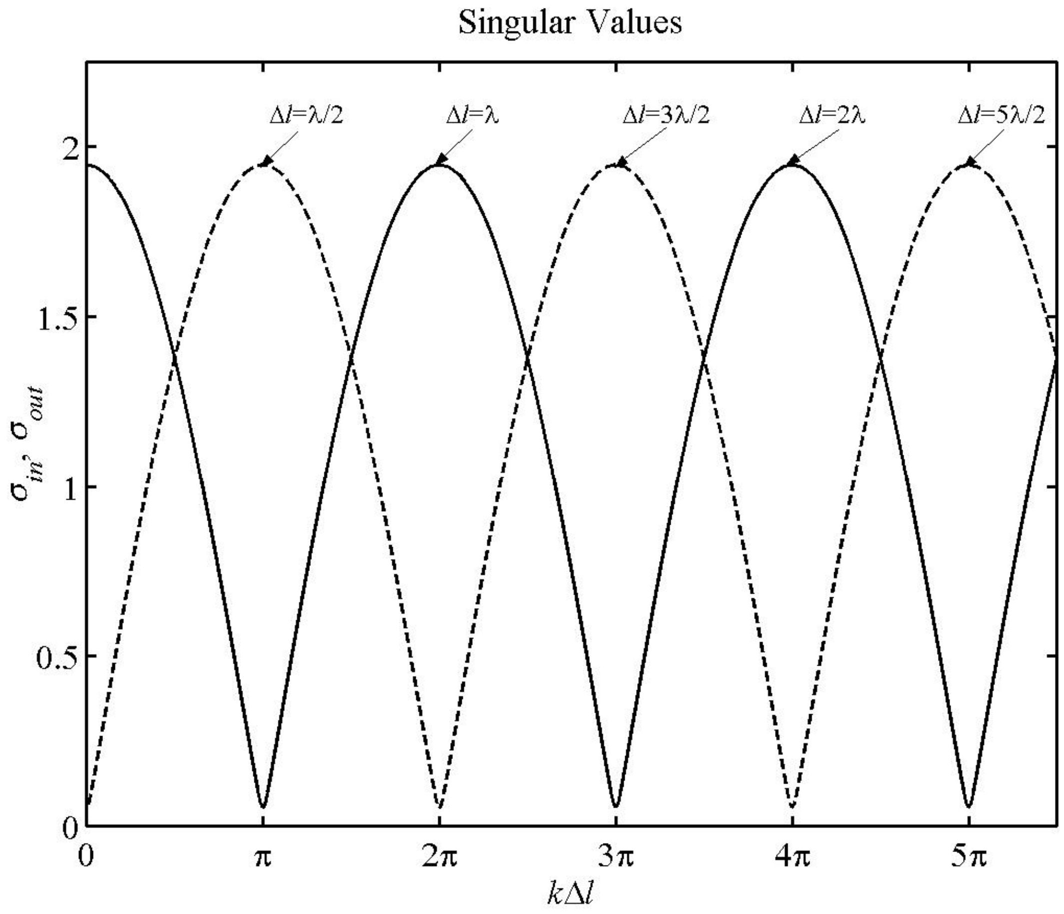
$$g = l_1/l_2 \qquad \Delta l = l_2 - l_1$$

$$\mathbf{d}^{\mathrm{T}} = \frac{\rho_0 e^{-jkl_1}}{4\pi l_1} \begin{bmatrix} d_R & d_L \end{bmatrix}$$

$\Delta l = \Delta r sin\theta$ when $l \gg \Delta r$

$$\mathbf{H} = \frac{1}{1 - g^2 e^{-2jkrsin\theta}} \begin{bmatrix} 1 & -ge^{-jk\Delta rsin\theta} \\ -ge^{-jk\Delta rsin\theta} & 1 \end{bmatrix}$$

Singular values



Singular values of transmission path matrix become equal when

$$k\Delta r sin\theta = \frac{n\pi}{2}$$

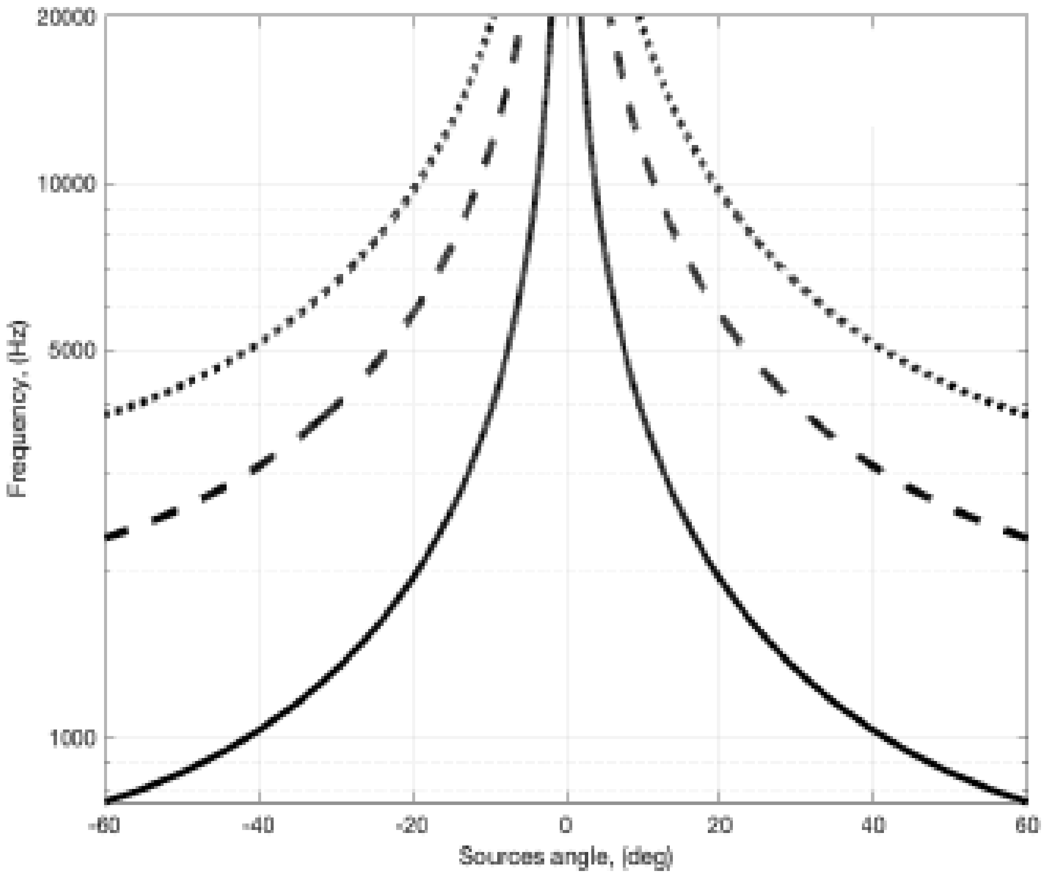
$$\Delta l = n\lambda/4$$

$n = \pm 1, \pm 3, \pm 5, \ldots\ldots.$

$$\mathbf{H} = \frac{1}{1+g^2}\begin{bmatrix} 1 & -jg \\ -jg & 1 \end{bmatrix}$$

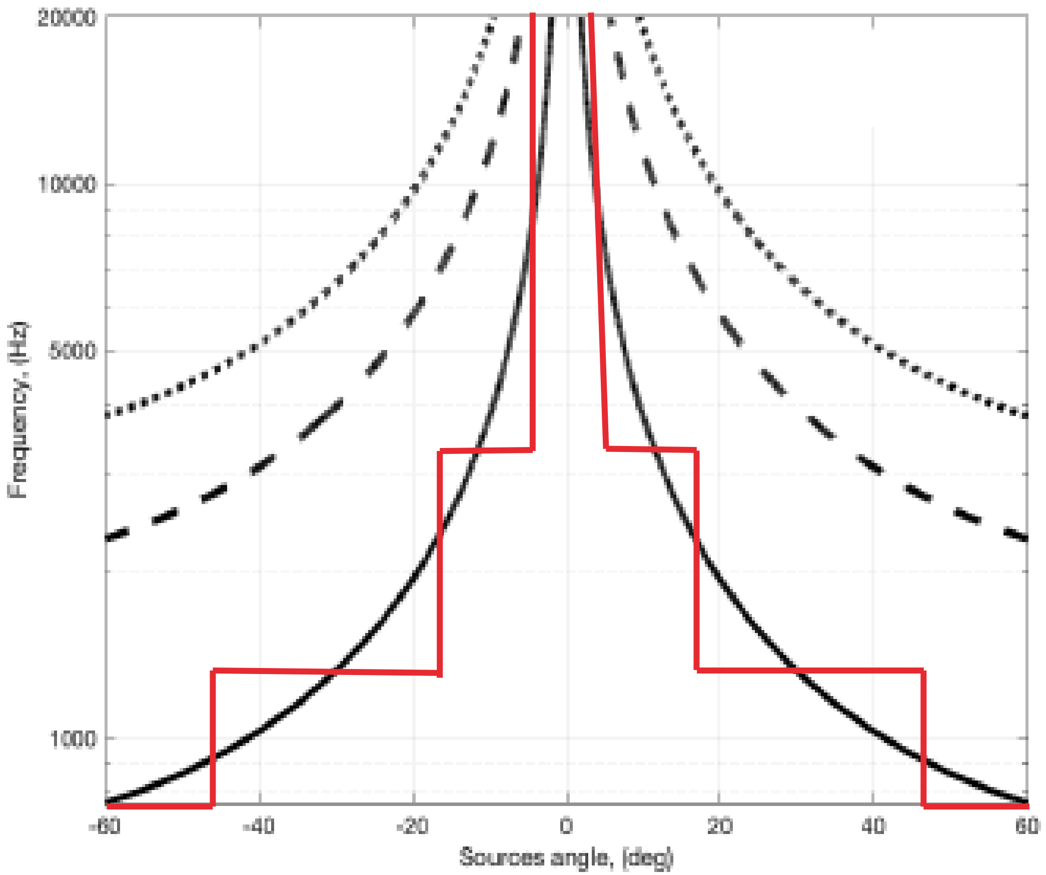
Takeuchi and Nelson JASA 112 (6) 2786-2797

Optimal source angle



Optimal source angle giving perfect conditioning for
$n = \pm1, \pm3, \pm5$

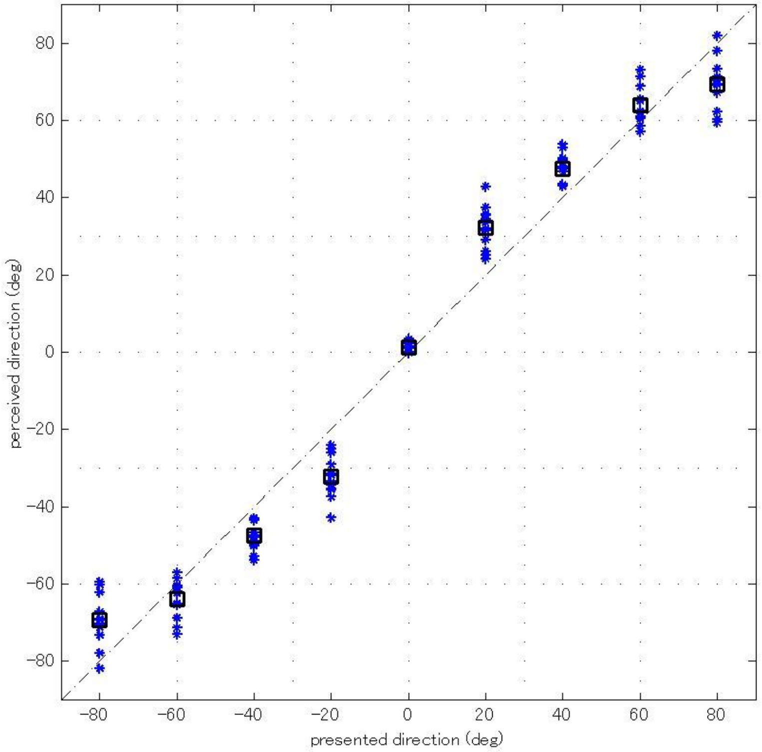
Discrete approximation to the optimal source angle



Optimal source angle giving perfect conditioning for
$n = \pm 1, \pm 3, \pm 5$

Subjective experiments

- Azimuth localisation from 3-way discrete optimal source distribution (OSD)



Some OSD adopters

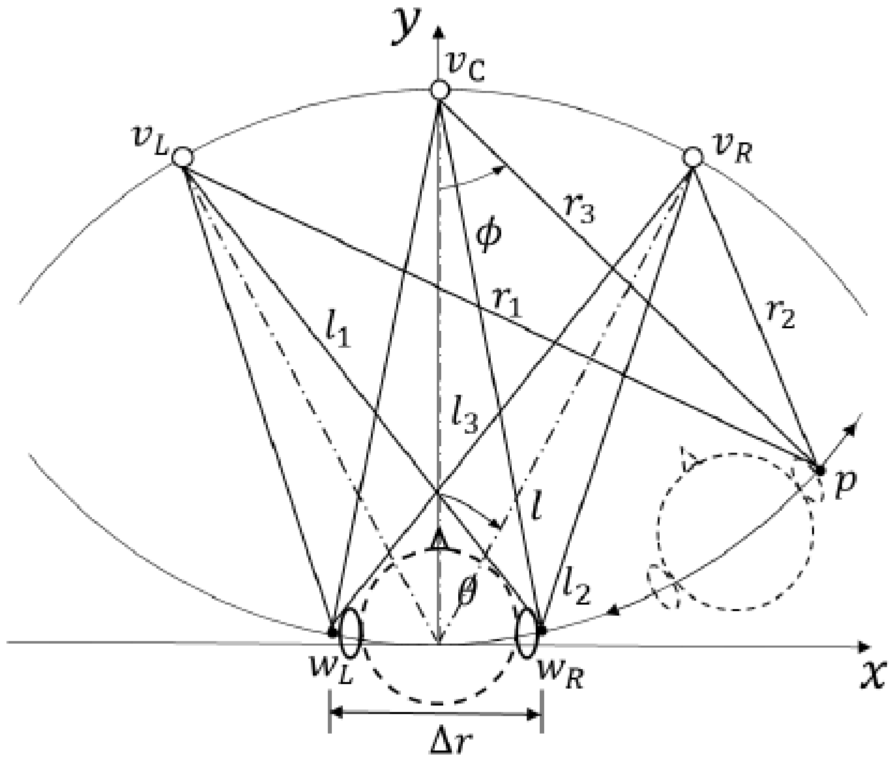
Marantz ES 7001

Sharp 8A-C22CX1

iF DESIGN AWARD 2021

Sherwood S7

Three Channel Optimal Source Distribution



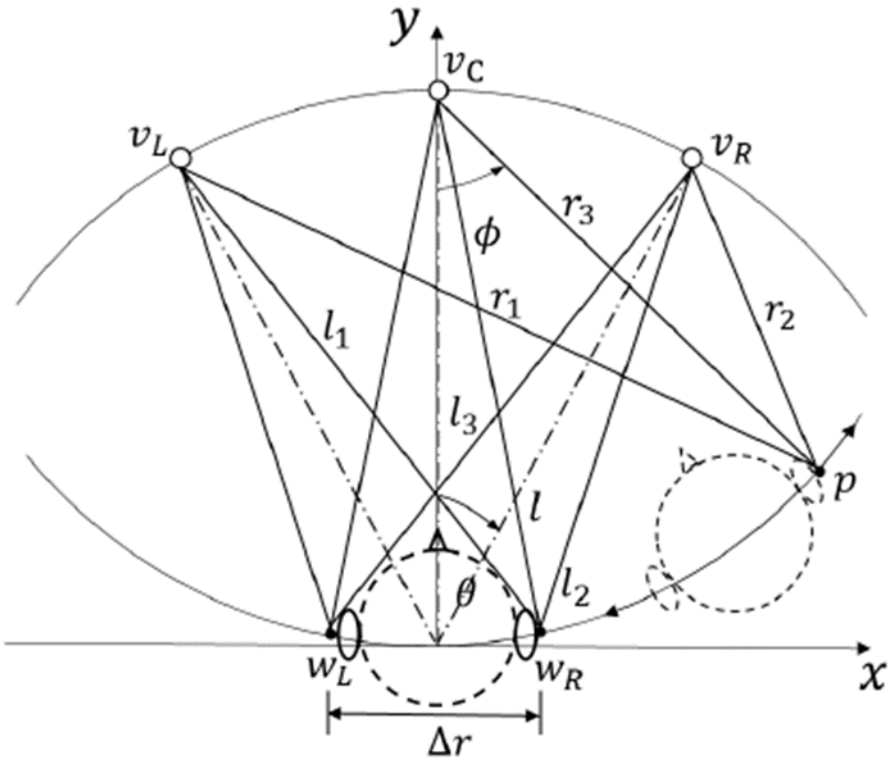
$$\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} v_R & v_C & v_L \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \\ H_{31} & H_{32} \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \end{bmatrix}$$

$$\mathbf{C} = \frac{\rho_0}{4\pi} \begin{bmatrix} \frac{e^{-jkl_1}}{l_1} & s\frac{e^{-jkl_3}}{l_3} & \frac{e^{-jkl_2}}{l_2} \\ \frac{e^{-jkl_2}}{l_2} & s\frac{e^{-jkl_3}}{l_3} & \frac{e^{-jkl_1}}{l_1} \end{bmatrix}$$

See Takeuchi and Nelson, Acta Acustica United with Acustica Vol. 94 (2008) 981 – 987

Singular values



$$\mathbf{C} = \frac{\rho_0}{4\pi}\frac{e^{-jkl_1}}{l_1}\begin{bmatrix} 1 & se^{-jk\Delta l/2} & e^{-jk\Delta l} \\ e^{-jk\Delta l} & se^{-jk\Delta l/2} & 1 \end{bmatrix}$$

Find eigenvalues from roots of determinant of

$$\lambda \mathbf{I} - \mathbf{C}\mathbf{C}^{\mathrm{H}}$$

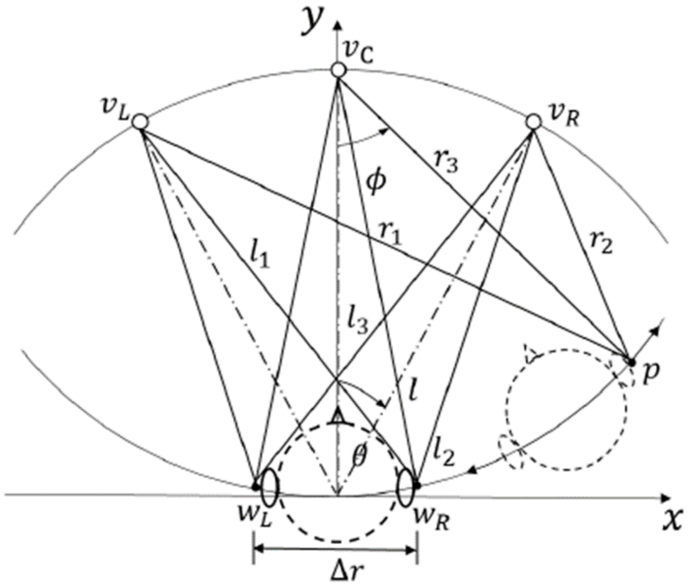
Singular values given by

$$\sigma_+ = \sqrt{[2 + 2s^2 + 2cosk\Delta l]}, \quad \sigma_- = \sqrt{[2 - 2cosk\Delta l]}$$

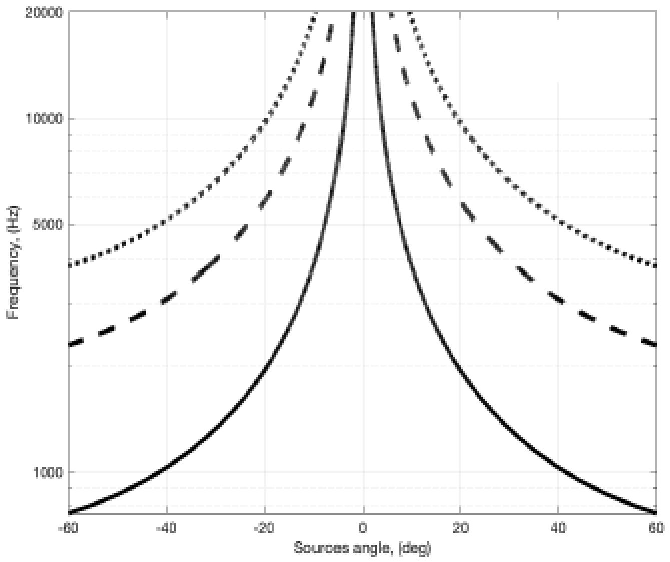
Optimal conditioning given by $s = \sqrt{2}$ which shows

$$\Delta l = n\lambda/2$$

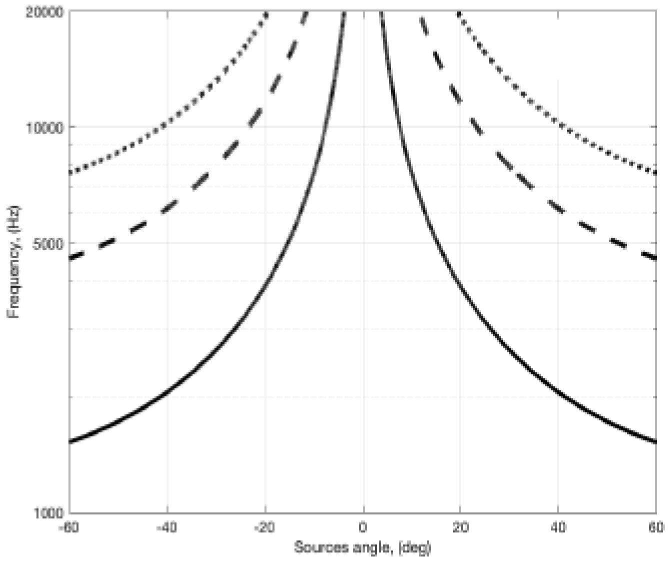
Optimal source angle as a function of frequency



Two channel OSD

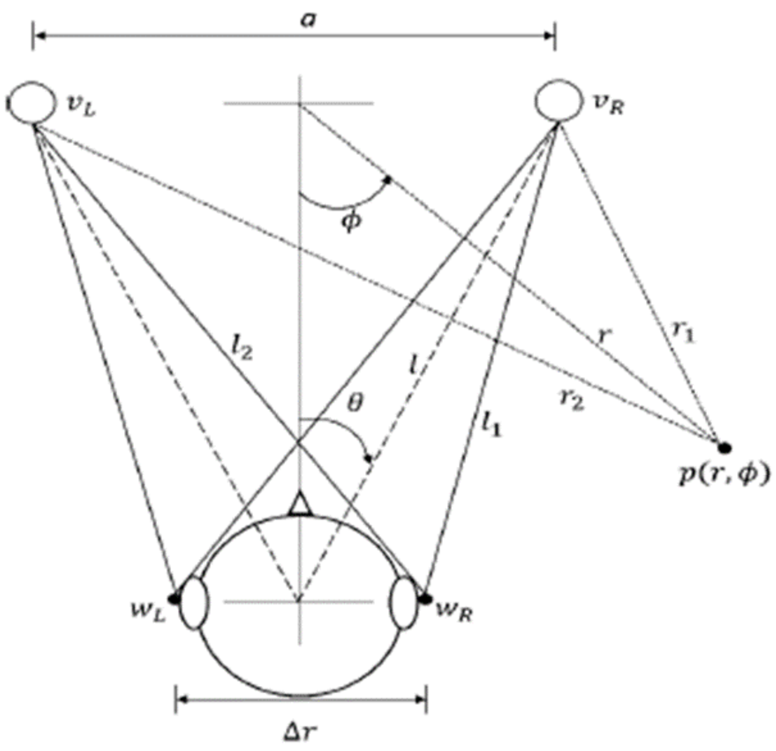


Three channel OSD



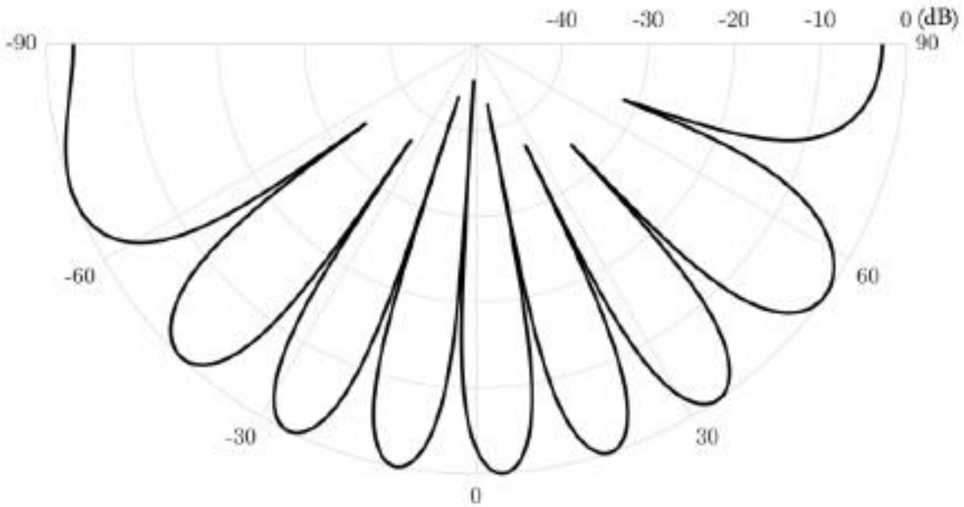
Optimal source angle giving perfect conditioning for $n = \pm 1, \pm 3, \pm 5$

Radiation properties of the two channel OSD



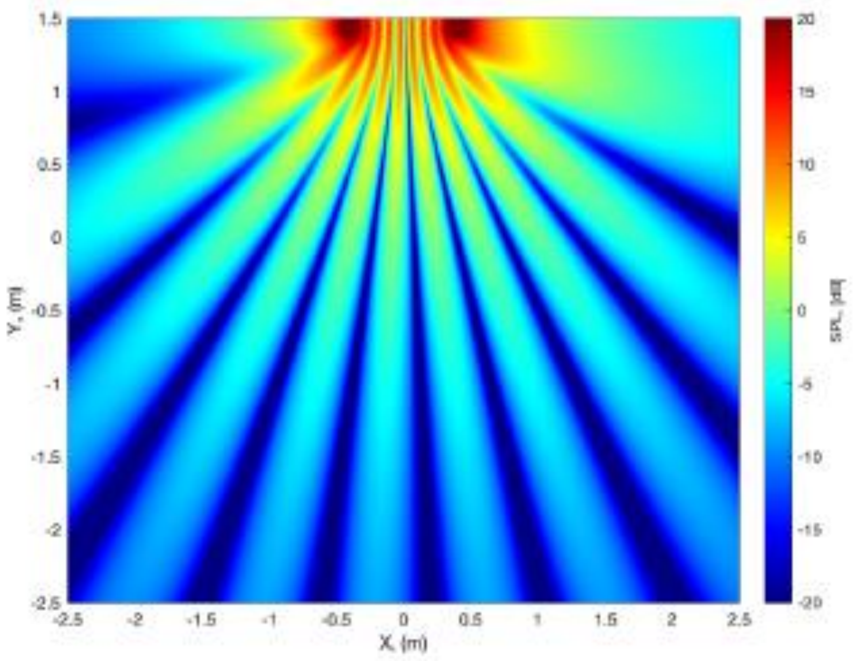
In the far field $r \approx r_1 \approx r_2$

$$|p(r,\phi)|^2 = \left(\frac{\rho_0}{4\pi r_1}\right)^2 (1 - sin(kasin\phi))$$

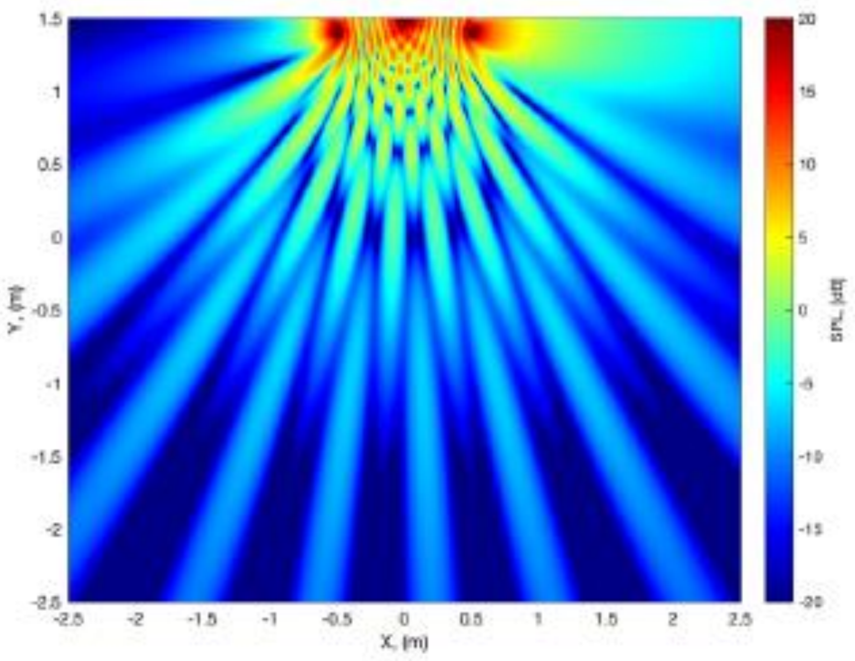


Radiation pattern examples

Two channel OSD at 2058 Hz

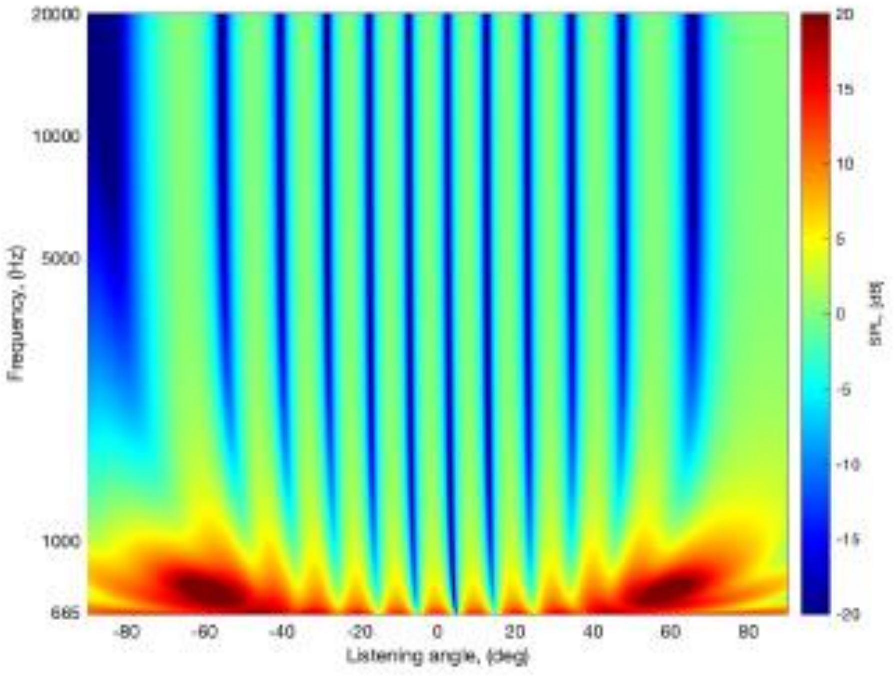


Three channel OSD at 3206 Hz

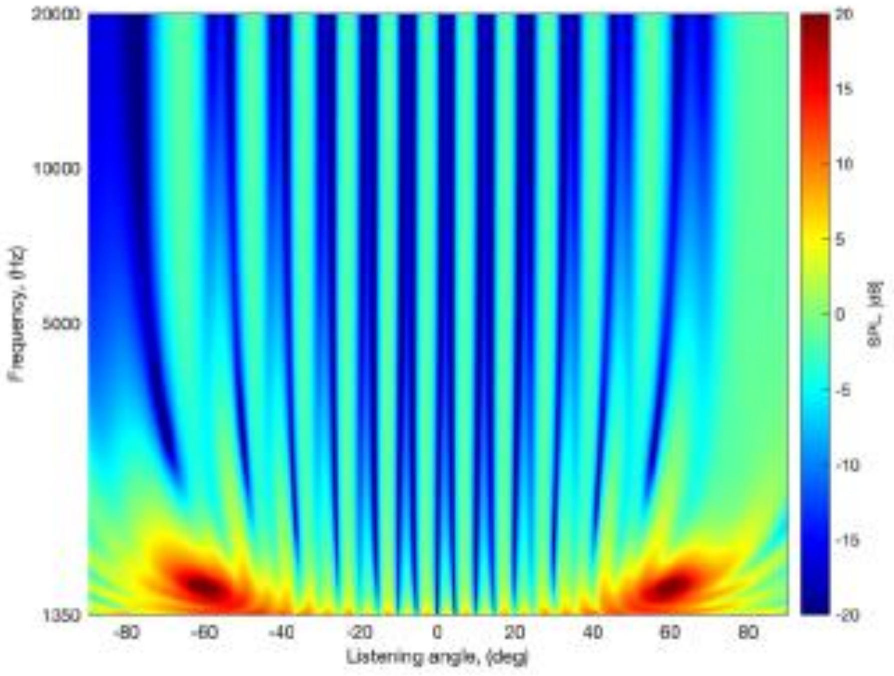


Radiation patterns as a function of source angle

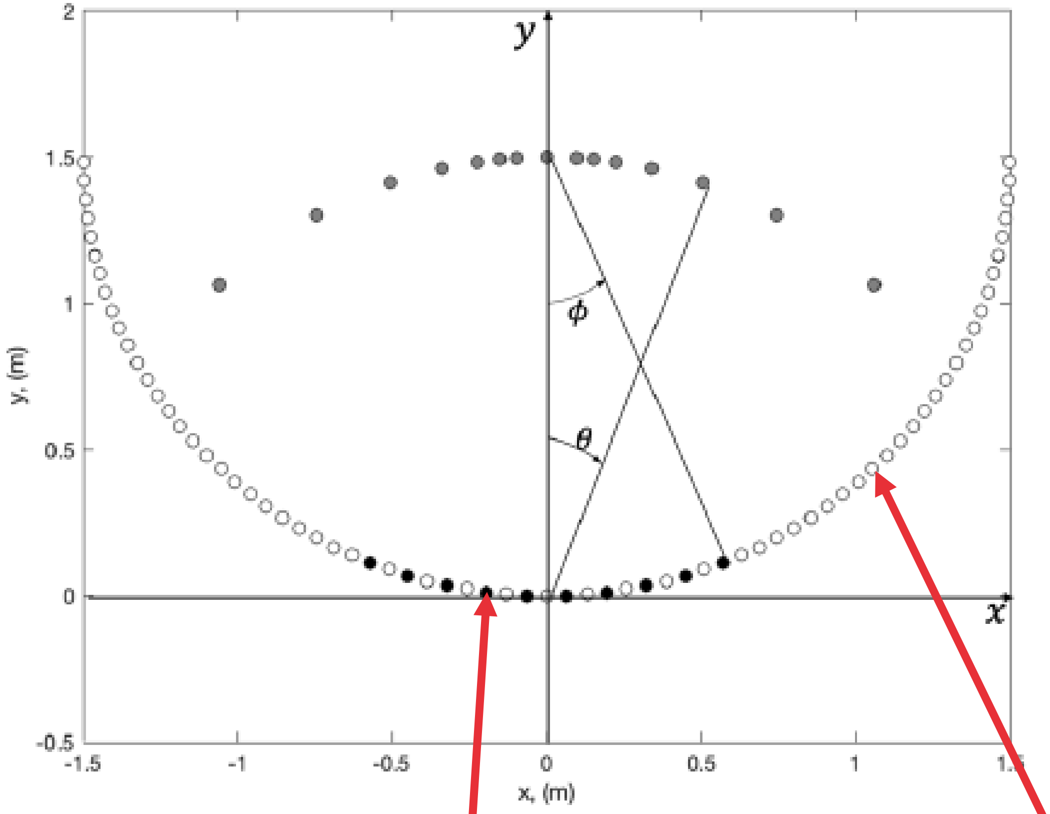
Two channel OSD



Three channel OSD



Discretisation of the Optimal Source Distribution



Desired signals $\hat{\mathbf{w}}_B$

Desired signals $\hat{\mathbf{w}}_A$

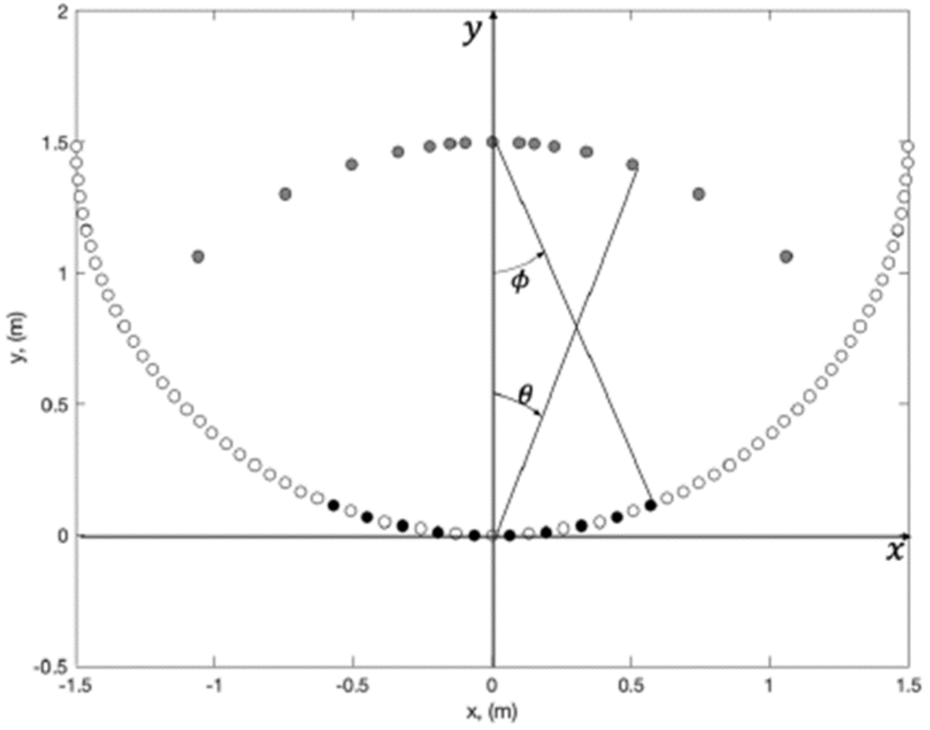
$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_A \\ \mathbf{w}_B \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathbf{v} = \mathbf{C}\mathbf{v}$$

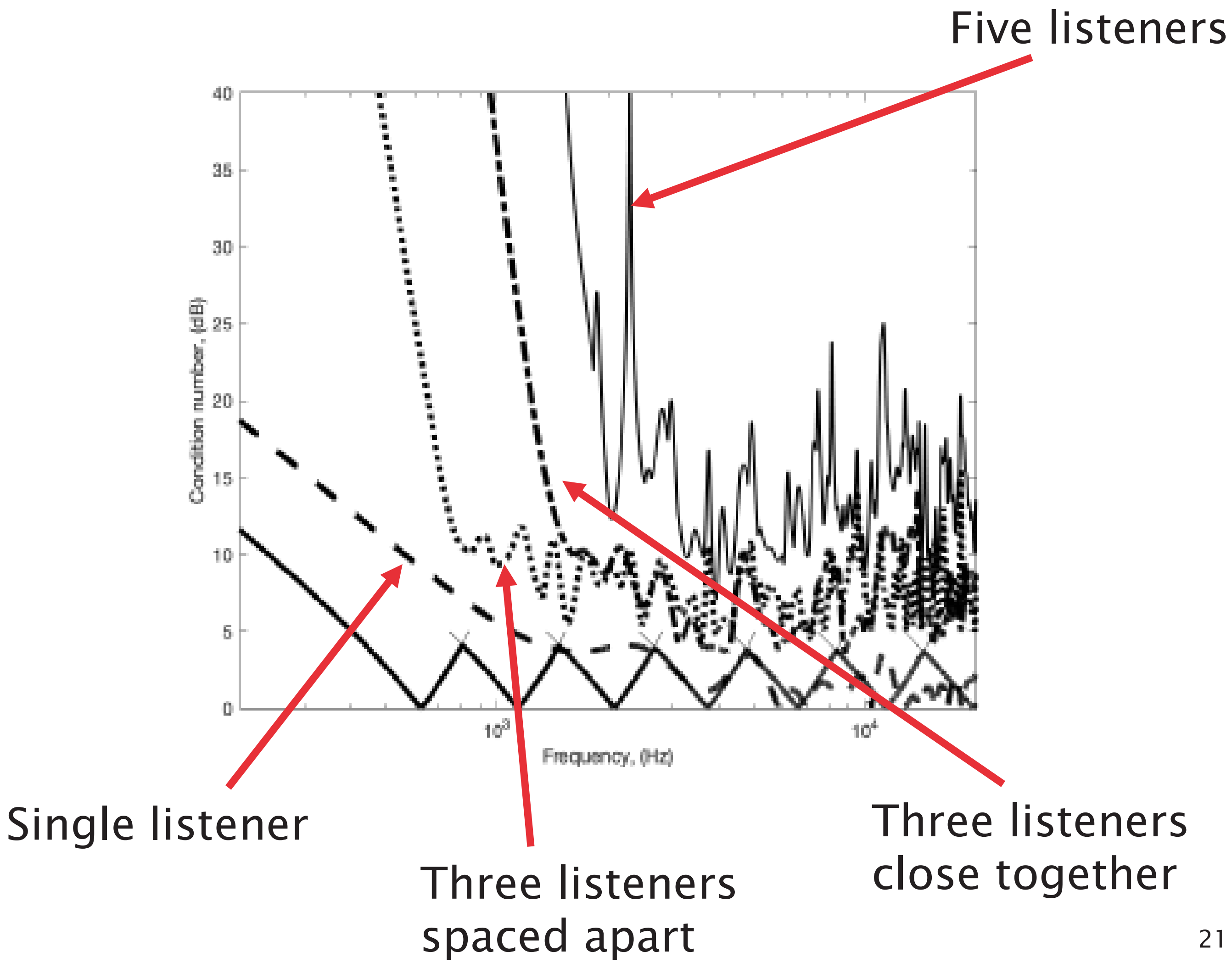
$$\hat{\mathbf{w}}_B = \mathbf{D}\mathbf{d}$$

$$\begin{bmatrix} \hat{w}_{B1} \\ \hat{w}_{B2} \\ \hat{w}_{B3} \\ \hat{w}_{B4} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_R \\ d_L \end{bmatrix}$$

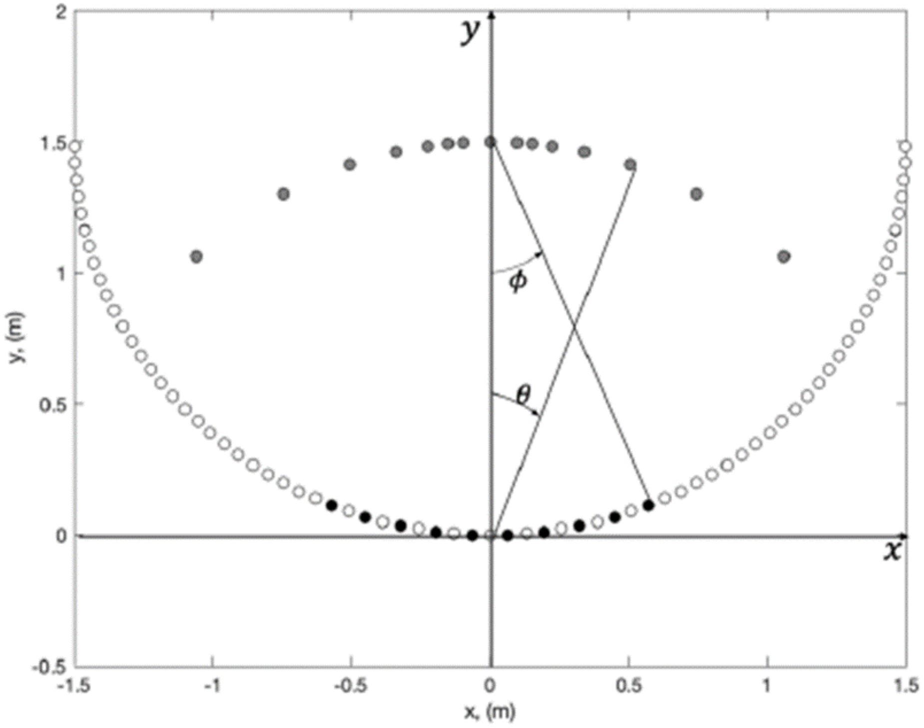
See P.A. Nelson, T. Takeuchi, P. Couturier, X. Zhou, J. Sound Vib.539 (2022) 117259

Condition number of matrices **B** and **C** for 2Ch OSD

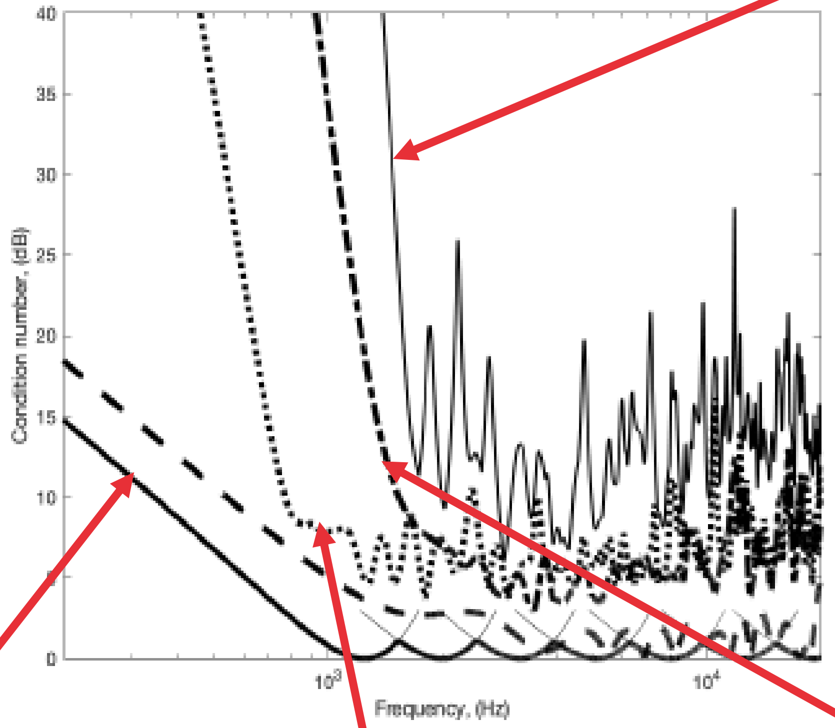




Condition number of matrices B and C for 3Ch OSD



Five listeners

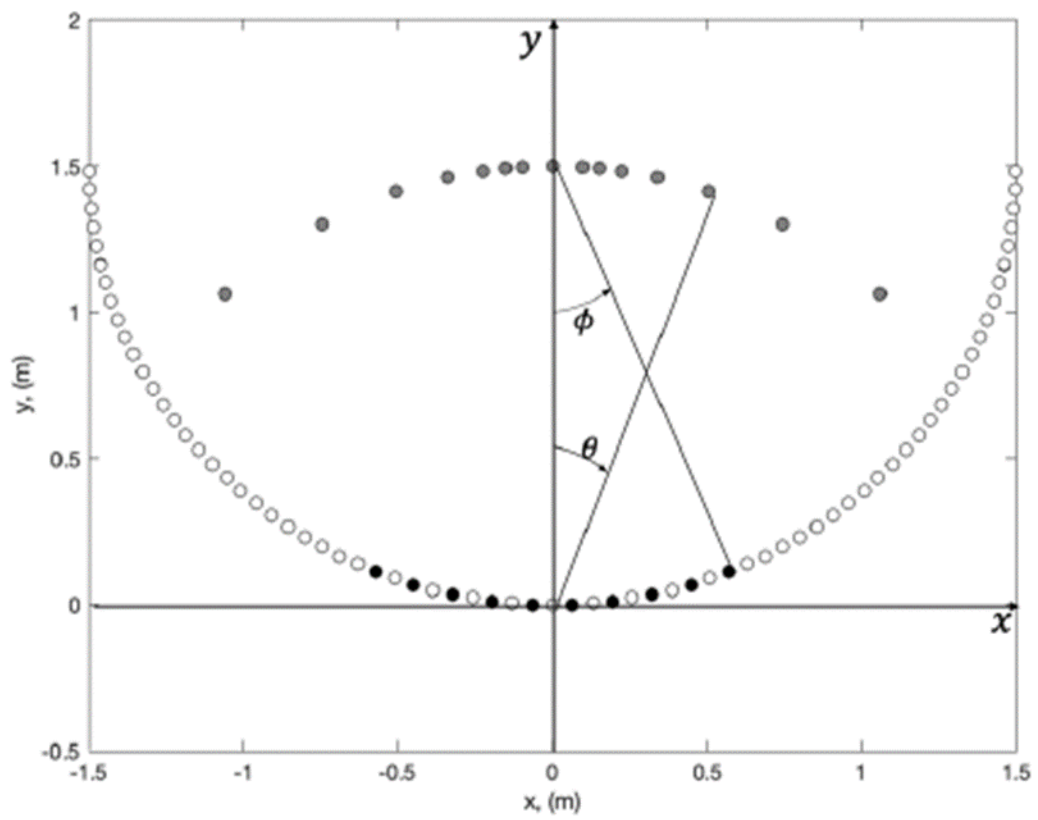


Single listener

Three listeners spaced apart

Three listeners close together

Method (1) Minimise deviation from OSD sound field



$$min\left[\|\mathbf{Av}-\hat{\mathbf{w}}_A\|_2^2+\beta\|\mathbf{v}\|_2^2\right]$$

Solution given by

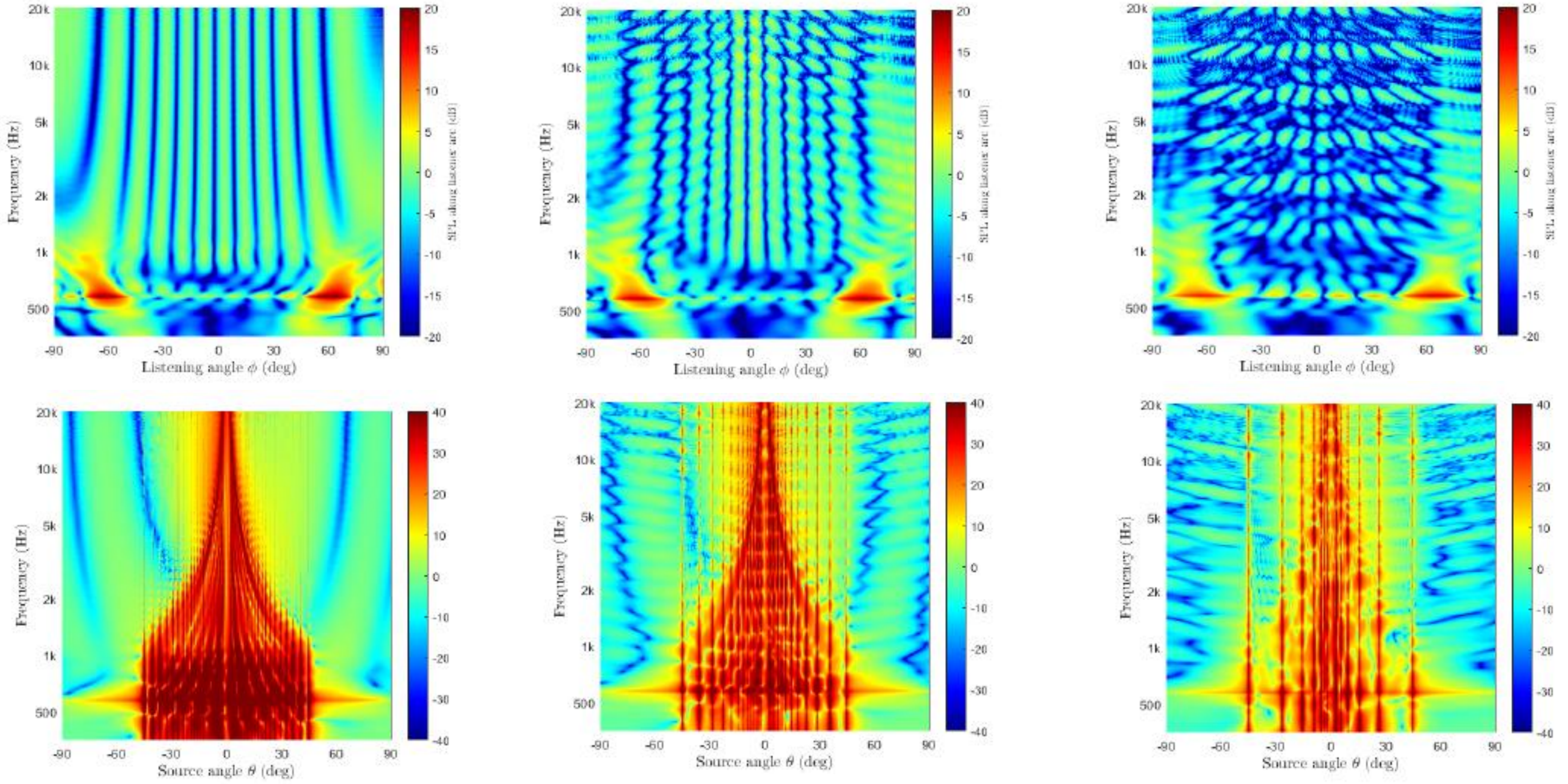
$$\mathbf{v}_{opt}=[\mathbf{A}^{\mathrm{H}}\mathbf{A}+\beta\mathbf{I}]^{-1}\mathbf{A}^{\mathrm{H}}\hat{\mathbf{w}}_A$$

Numerical experiments with 95 equispaced far field sensors

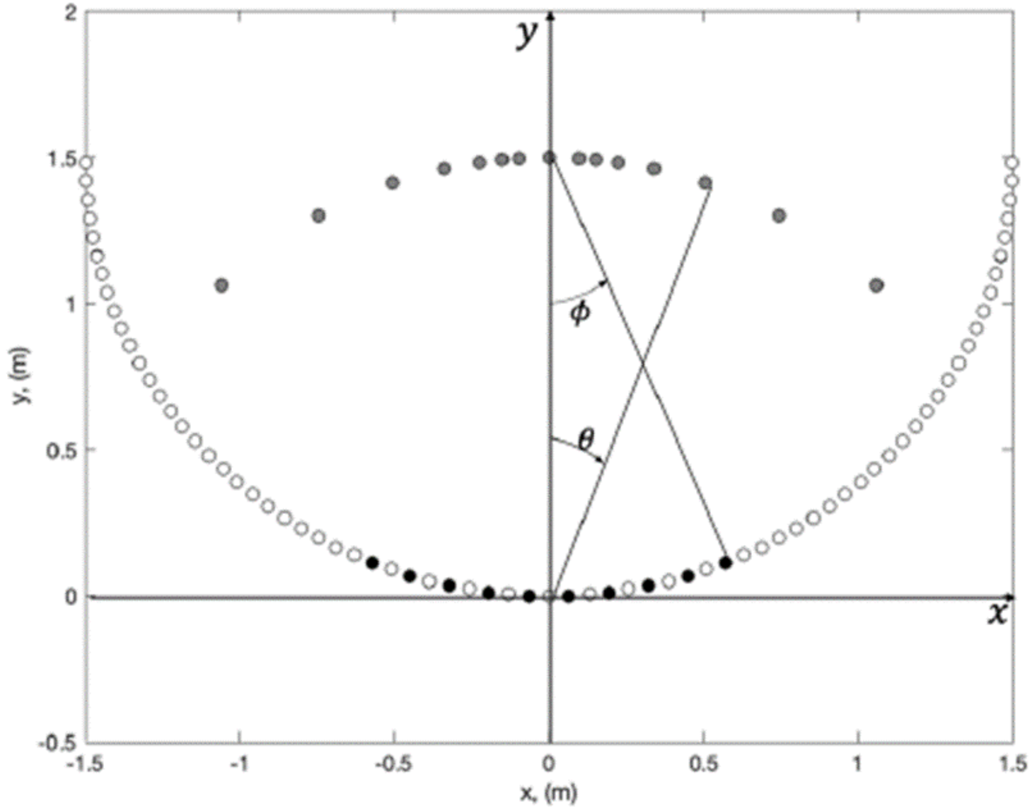
Variation in number of sources

All results with regularisation $\beta=0.01$

62 Sources | 30 Sources | 14 Sources



Method (2) Minimise deviation from OSD sound field with the constraint of crosstalk cancellation



$$min\left\|\mathbf{A}\mathbf{v} - \hat{\mathbf{w}}_A\right\|_2^2 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v}$$

Solution via QR decomposition given by

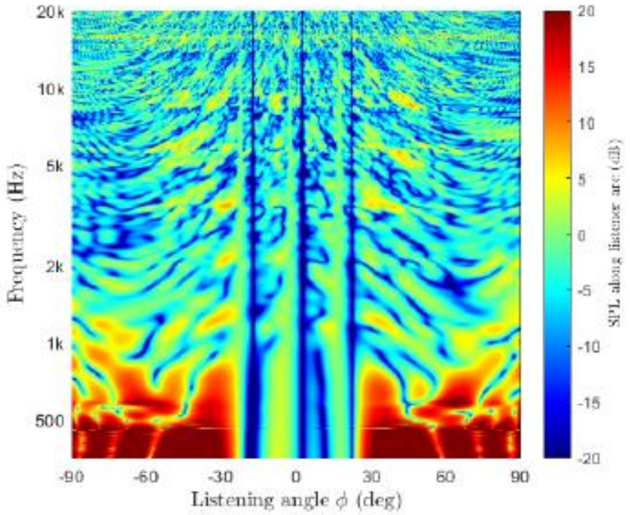
$$\mathbf{v}_{opt} = \mathbf{Q}_2\mathbf{A}_2^{\dagger}\hat{\mathbf{w}}_A + (\mathbf{Q}_1\mathbf{R}^{\mathrm{H}-1} - \mathbf{Q}_2\mathbf{A}_2^{\dagger}\mathbf{A}_1\mathbf{R}^{\mathrm{H}-1})\hat{\mathbf{w}}_B$$

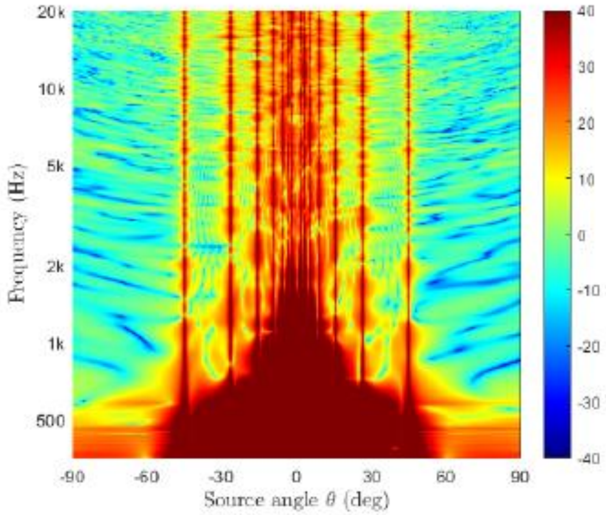
$$\mathbf{B}^{\mathrm{H}} = \mathbf{Q}\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix} \quad \mathbf{A}\mathbf{Q} = [\mathbf{A}_1 \quad \mathbf{A}_2] \quad \mathbf{Q} = [\mathbf{Q}_1 \quad \mathbf{Q}_2]$$

$$\mathbf{A}_2^{\dagger} = [\mathbf{A}_2^{\mathrm{H}}\mathbf{A}_2]^{-1}\mathbf{A}_2^{\mathrm{H}}$$

Results of imposing crosstalk cancellation constraint

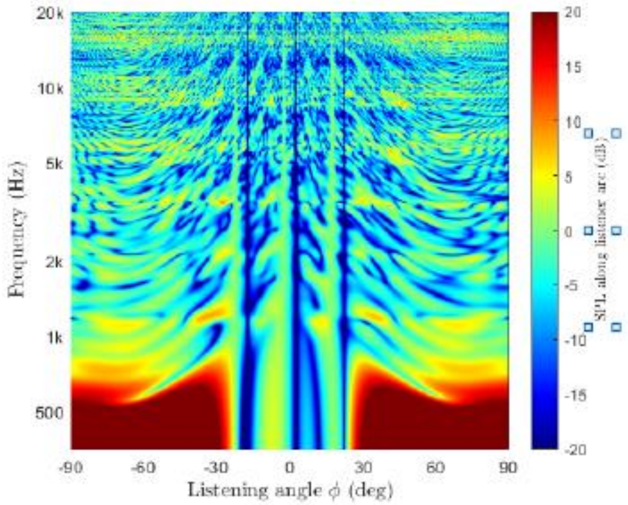
Without regularisation

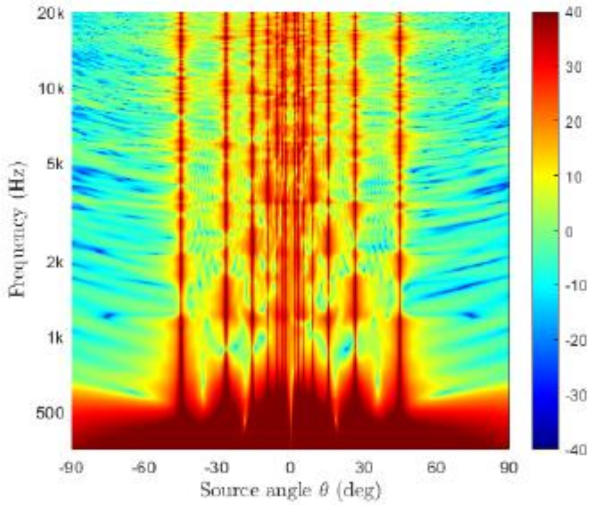




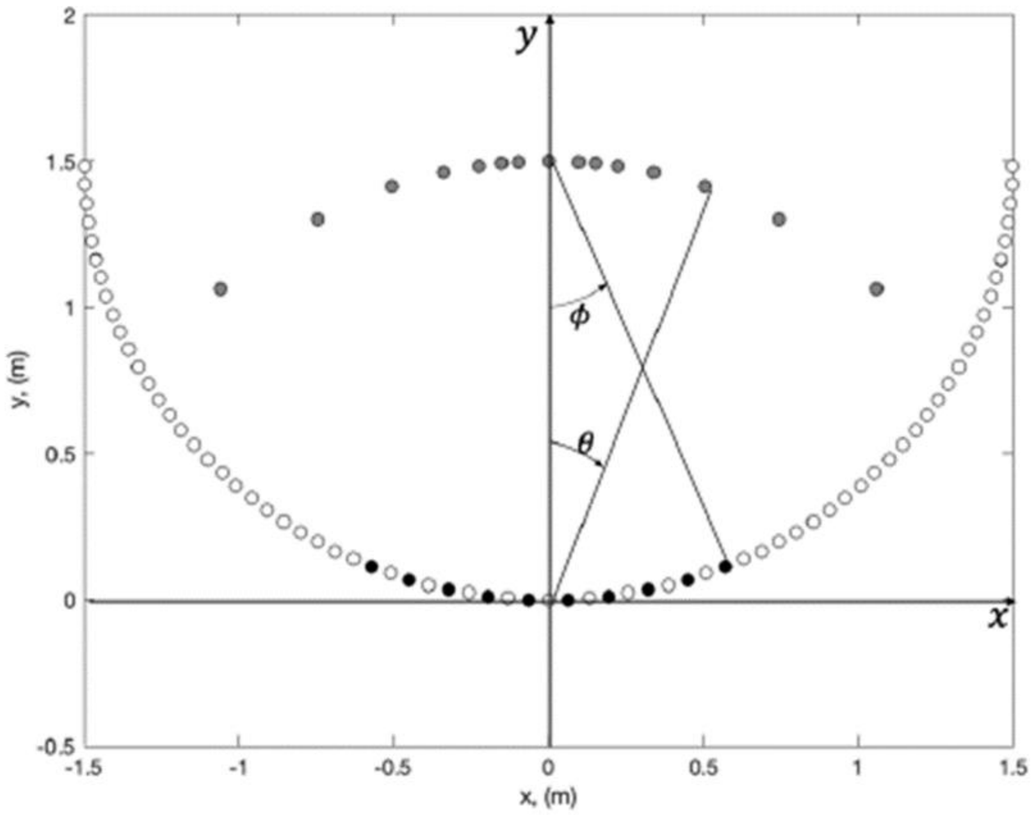
With regularisation of

$$\mathbf{A}_2^{\dagger} = [\mathbf{A}_2^{\mathrm{H}}\mathbf{A}_2]^{-1}\mathbf{A}_2^{\mathrm{H}}$$





Method (3) Minimise the L2 norm of the source strength vector



$$min\|\mathbf{v}\|_2^2 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v}$$

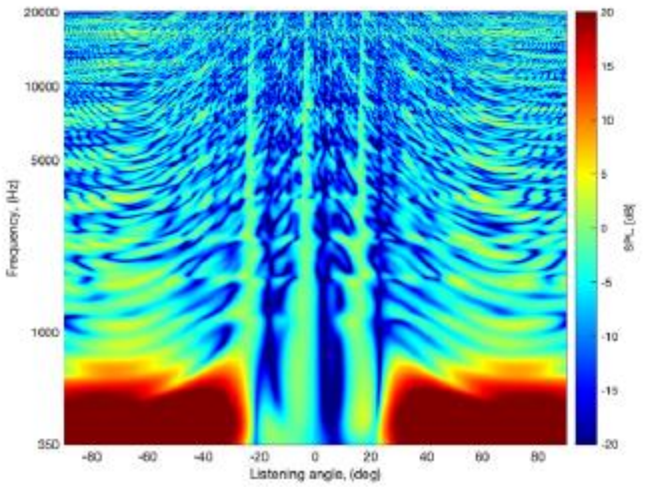
Solution given by

$$\mathbf{v}_{opt} = \mathbf{B}^{\mathrm{H}}[\mathbf{B}\mathbf{B}^{\mathrm{H}} + \beta\mathbf{I}]^{-1}\mathbf{D}\mathbf{d}$$

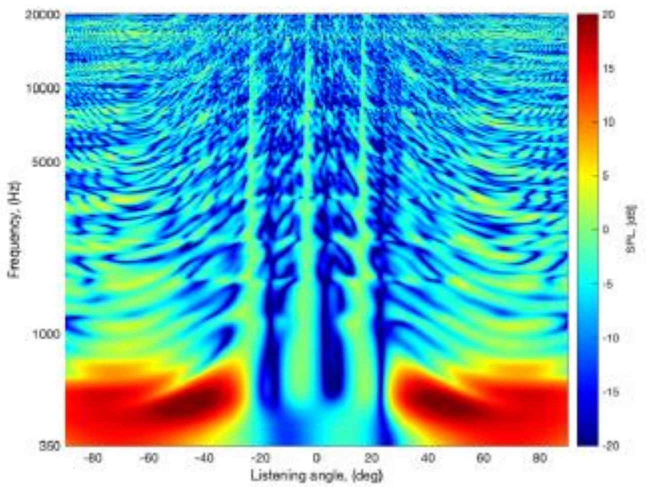
See also Holleborn et al, JAES 69 (3) (2021) 191-203.

Effect of regularisation

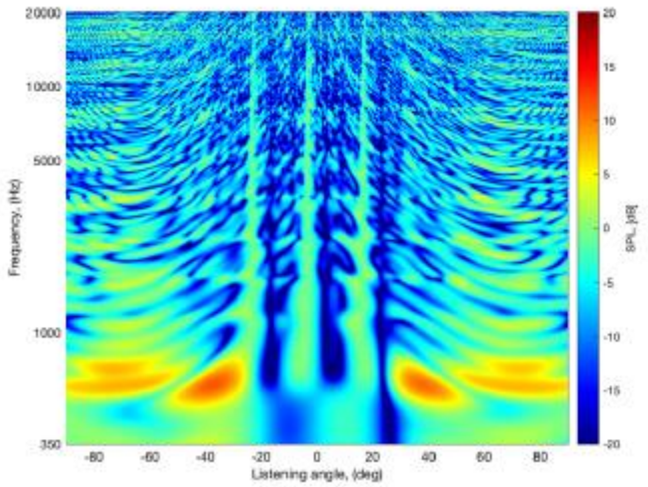
$\beta=0$



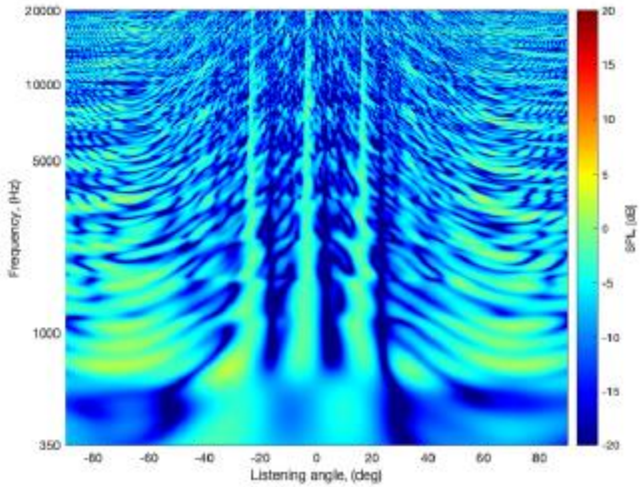
$\beta=0.0001$



$\beta=0.001$

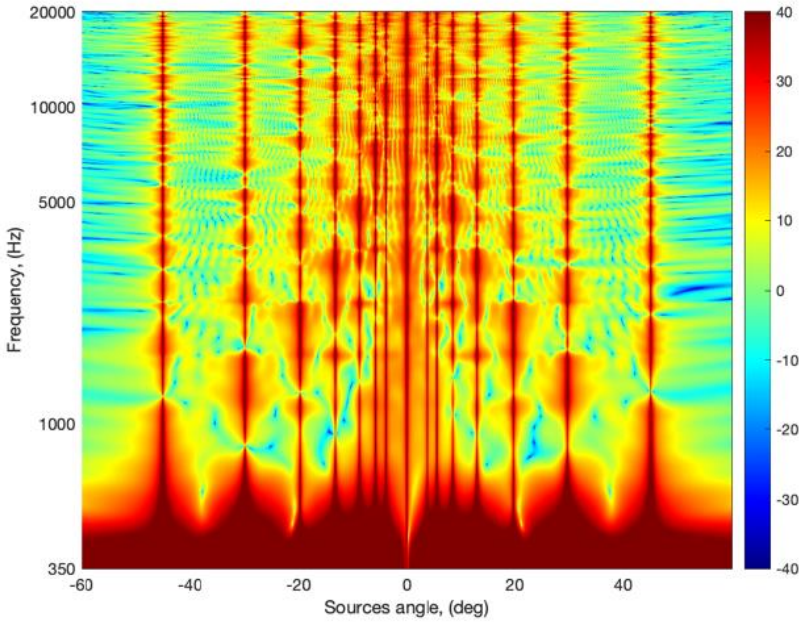


$\beta=0.01$

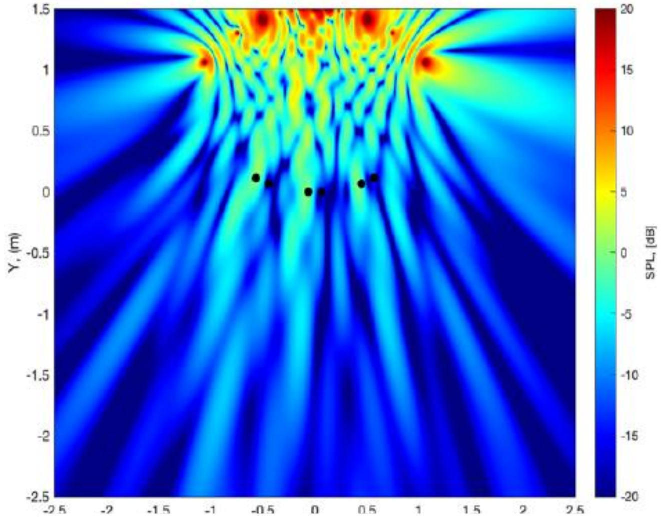


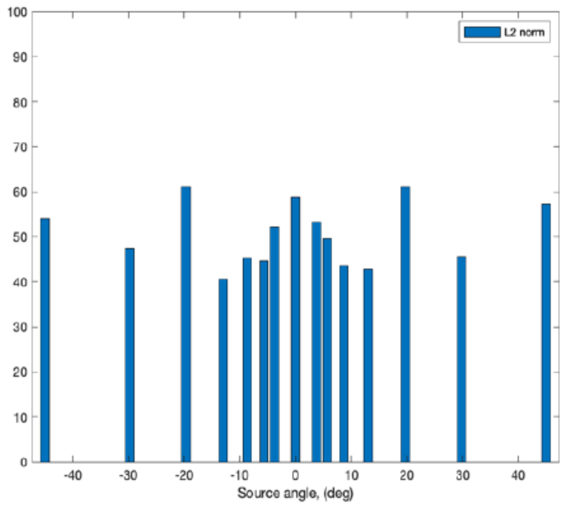
Results of L2 norm minimisation (3Ch OSD)

Using 15 sources and 3 spaced apart listeners

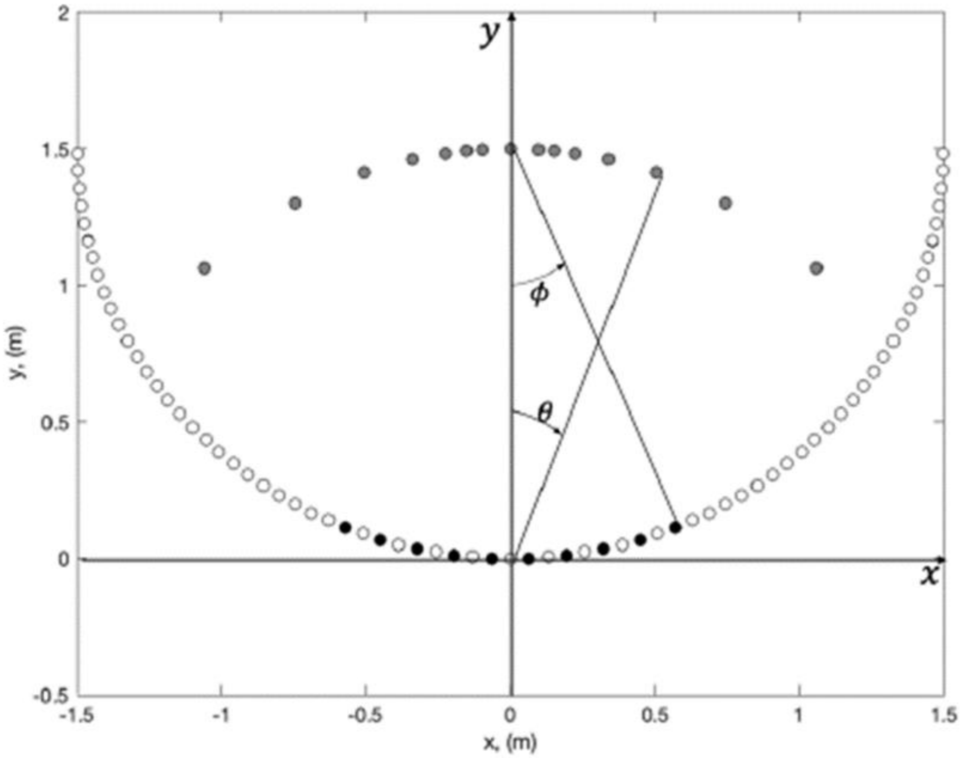


Results at 1995 Hz





Method (4) Minimise the L1 norm of the source strength vector

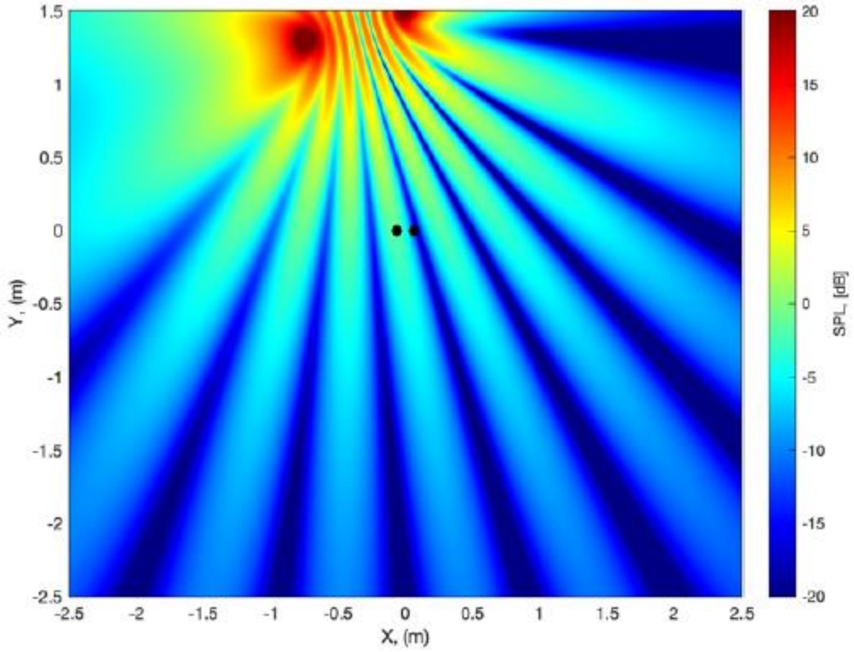


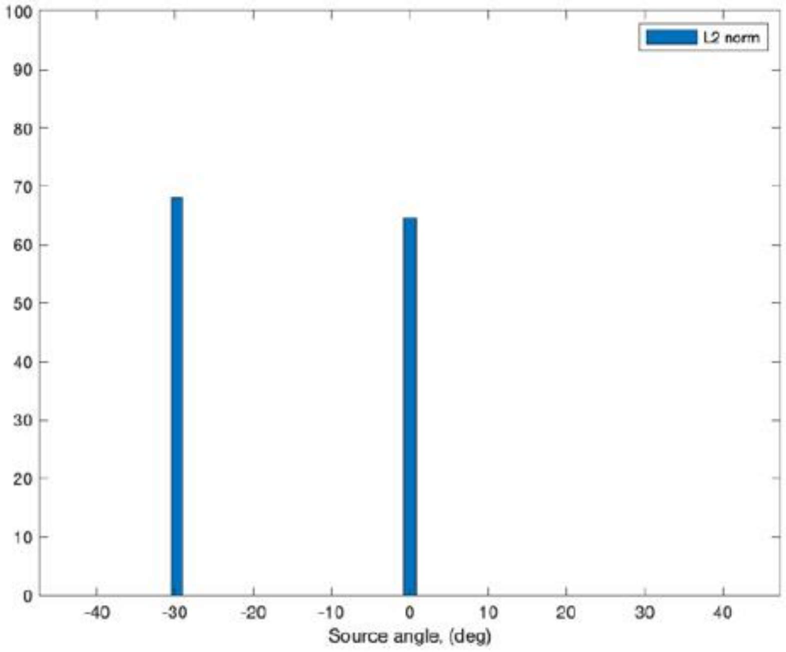
$$min\left\|\mathbf{v}\right\|_1 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v}$$

Solutions computed numerically using CVX package

Single listener crosstalk cancellation (3Ch OSD)

Frequency of 1995 Hz

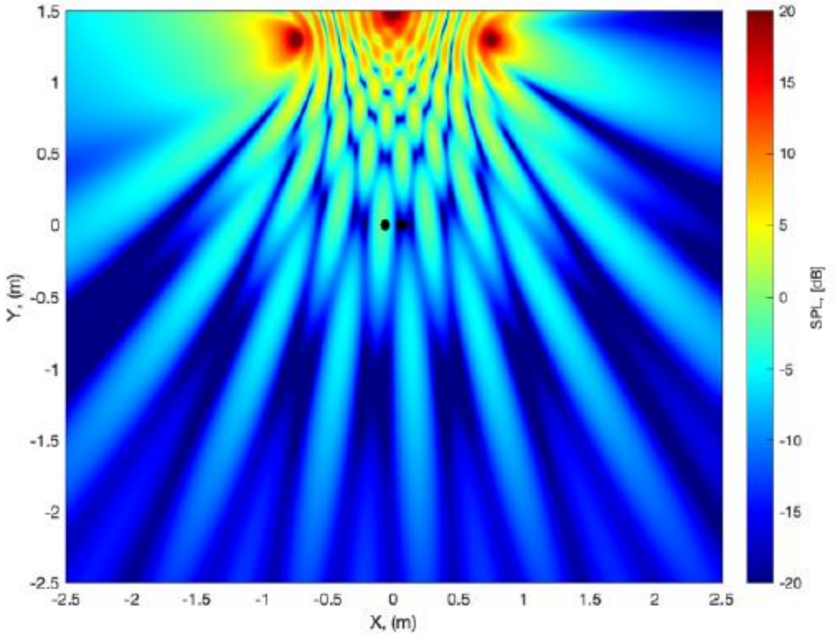


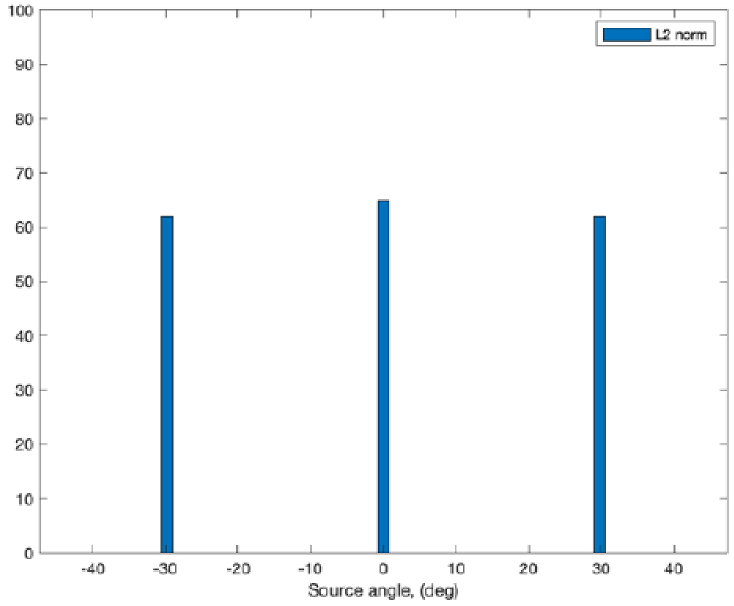


Single listener crosstalk cancellation (3Ch OSD)

Results with additional symmetry constraint imposed

$$min\left\|\mathbf{v}\right\|_1 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v} \quad \text{and} \quad \upsilon_m = -\upsilon_{-m}$$

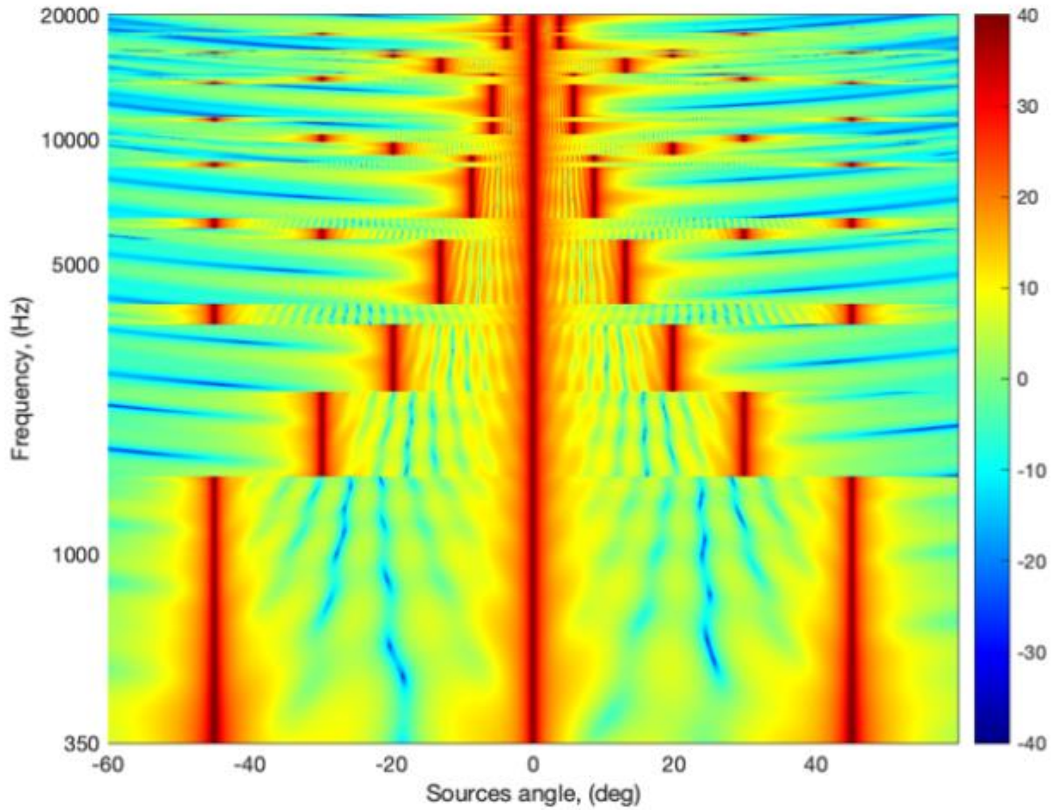




Single listener crosstalk cancellation (3Ch OSD)

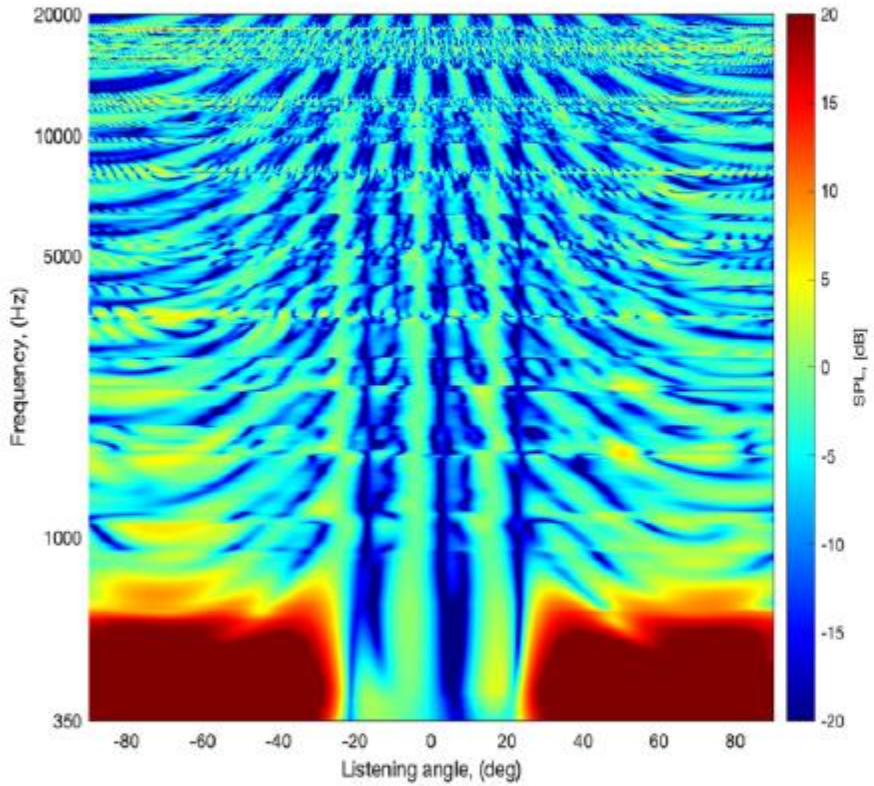
Results with additional symmetry constraint imposed

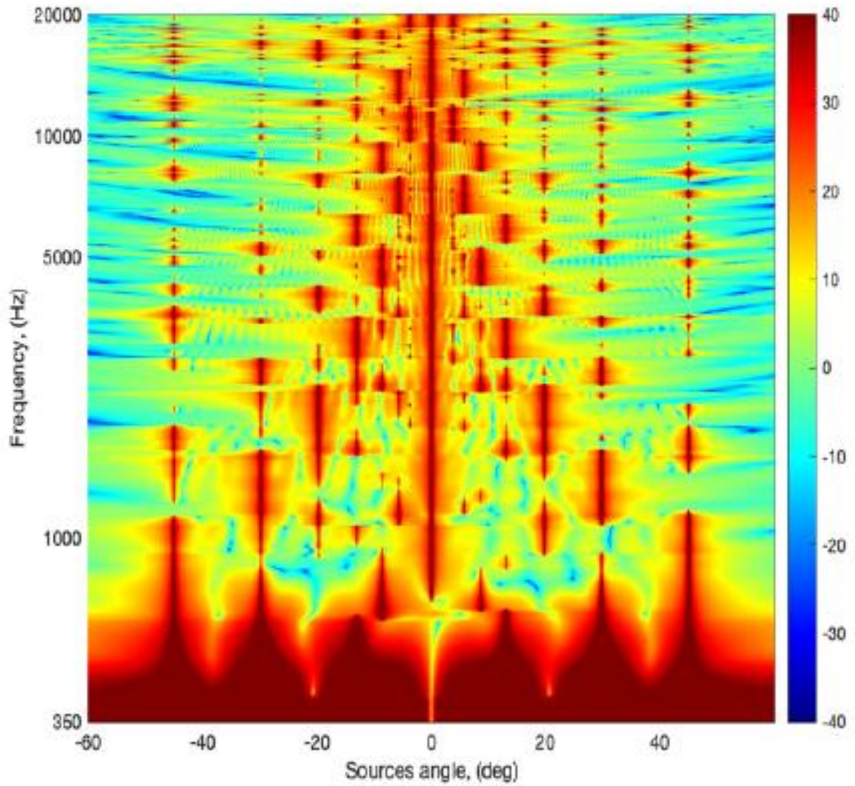
$$min\left\|\mathbf{v}\right\|_1 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v} \quad \text{and} \quad v_m = -v_{-m}$$



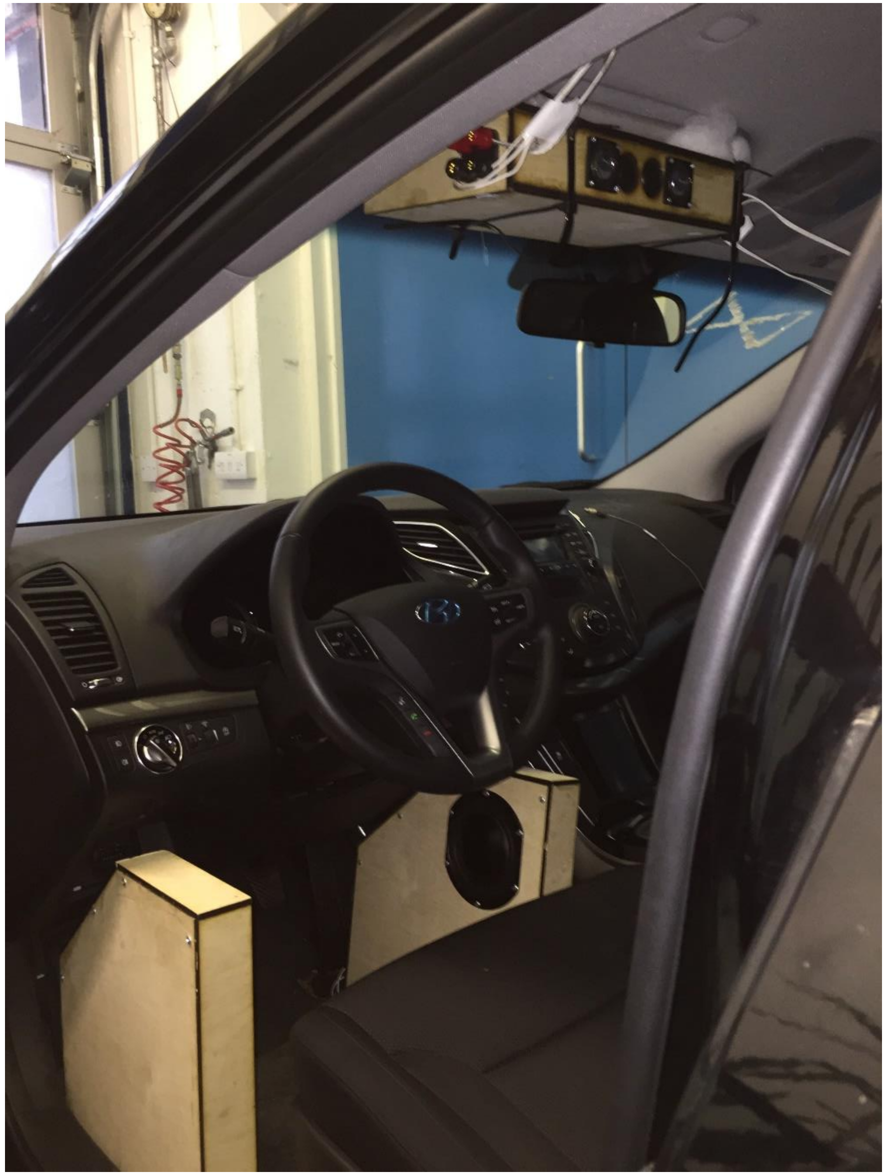
Results for crosstalk cancellation at three spaced apart listeners

$$min\|\mathbf{v}\|_1 \quad \text{subject to} \quad \hat{\mathbf{w}}_B = \mathbf{B}\mathbf{v} \quad \text{and} \quad v_m = -v_{-m}$$





The OSD in automotive applications



Conclusions

- A number of strategies for making use of the remarkable properties of the Optimal Source Distribution (OSD) have been analysed by using numerical simulations
- Replicating the OSD sound field can be achieved in principle, but requires a large number of sources
- Minimising the deviation from the OSD sound field with the constraint of crosstalk cancellation at a number of listeners seems feasible
- Minimising the L2 norm of the source strength vector appears also to give good results
- Minimising the L1 norm of the source strength vector with a symmetry constraint appears to yield the same solution as the OSD for a single listener
- Minimising the L1 norm of the source strength vector seems also to provide a sparse solution for multiple listeners

Further work

- Evaluate the subjective performance of such systems, paying particular attention to the time domain response associated with the different strategies available
- Understand better the process of L1 norm minimisation (a variety of algorithms are available – some apparently simple and efficient)
- Investigate further the optimal geometrical disposition of sources and listeners and the influence of such factors on subjective performance

YOUR QUESTIONS