Binaural Beamforming Taking into Account Spatial Release from Masking

de Vries, Johannes W.; van de Par, Steven; Leus, Geert; Heusdens, Richard; Hendriks, Richard C.

DOI
10.1109/TASLP.2024.3451988

Publication date
2024

Document Version
Final published version

Published in
IEEE/ACM Transactions on Audio Speech and Language Processing

Citation (APA)
de Vries, J. W., van de Par, S., Leus, G., Heusdens, R., & Hendriks, R. C. (2024). Binaural Beamforming Taking into Account Spatial Release from Masking. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *32*, 4002-4012. https://doi.org/10.1109/TASLP.2024.3451988

Binaural Beamforming Taking into Account Spatial Release from Masking

J.W. de Vries,[@],[*] S. van de Par,[†] G.J.T. Leus,[*] R. Heusdens,[‡] R.C. Hendriks[*]

[@]j.w.devries-1@tudelft.nl, [*]Signal Processing Systems, TU Delft, [†]Acoustics Group, Carl von Ossietzky University, [‡]Netherlands Defence Academy

Problem Statement

Spatial filters (beamformers) in hearing aids typically focus on minimising noise and interference power in order to improve intelligibility. The binaural MVDR beamformer can do so optimally, but with a drawback: all the remaining interference is virtually moved to the target location. This eliminates any spatial release from masking (SRM), which also has a positive effect on intelligibility and is usually not taken into account in beamformer design. Additional beamformer constraints on the interferer binaural cues (leading to JBLCMV beamforming) can restore SRM, but at the cost of less effective noise and interference reduction.

A trade-off for intelligibility seems to exist between optimal beamforming and spatial release from masking!

Signal Model Extension

To take SRM into account, the beamformer signal model of (1) and (2) is extended with the binaural intelligibility prediction model of Figure 1, leading to (3).

HA recorded signals: $$\boldsymbol{x} = \boldsymbol{a}s + \boldsymbol{u}_0 + \sum_{i=1}^{I} \boldsymbol{b}_i u_i, \quad (1)$$

presented signals: $$\boldsymbol{y} = \begin{pmatrix} y_{\mathrm{L}} \\ y_{\mathrm{R}} \end{pmatrix} = \boldsymbol{W}^{\mathrm{H}}\boldsymbol{x}, \quad (2)$$

'perceived' signal: $$z = \boldsymbol{v}^{\mathrm{H}}(\boldsymbol{y} + \boldsymbol{n}). \quad (3)$$

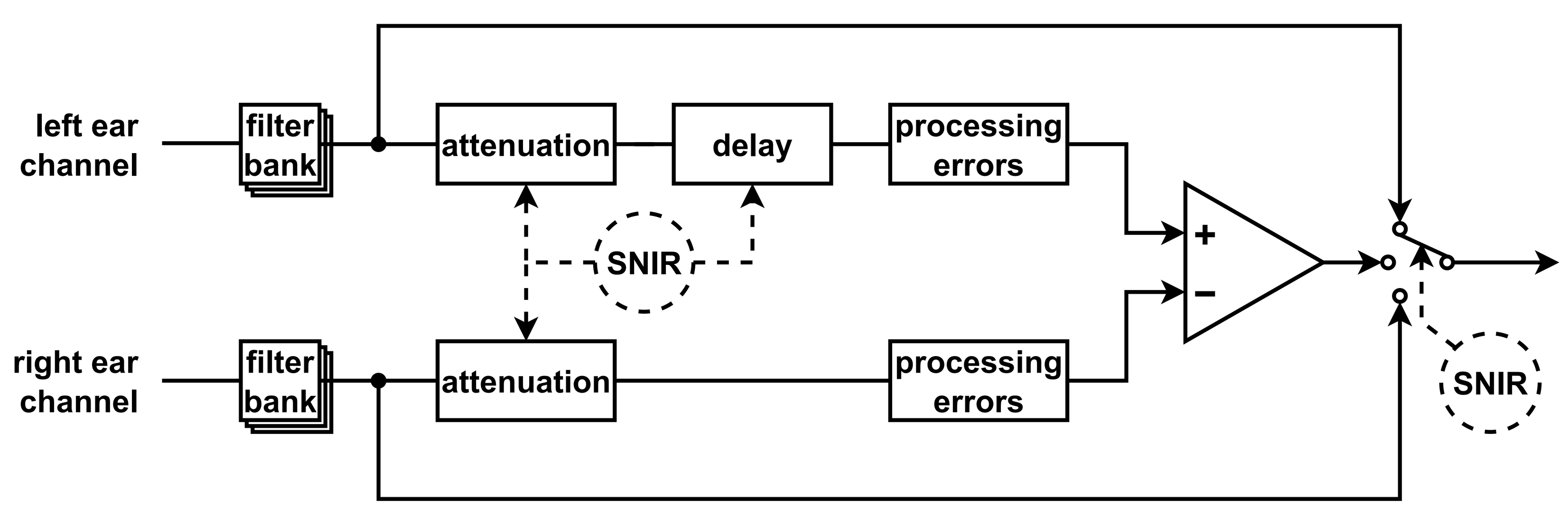

Figure 1: Schematic overview of the binaural intelligibility prediction model of Beutelmann and Brand (2006, JASA 120:331).

To investigate the trade-off, the BMVDR (optimal noise reduction) and JBLCMV (with room for SRM) beamformers are simulated in various acoustic scenes, both in presented and 'perceived' domain.

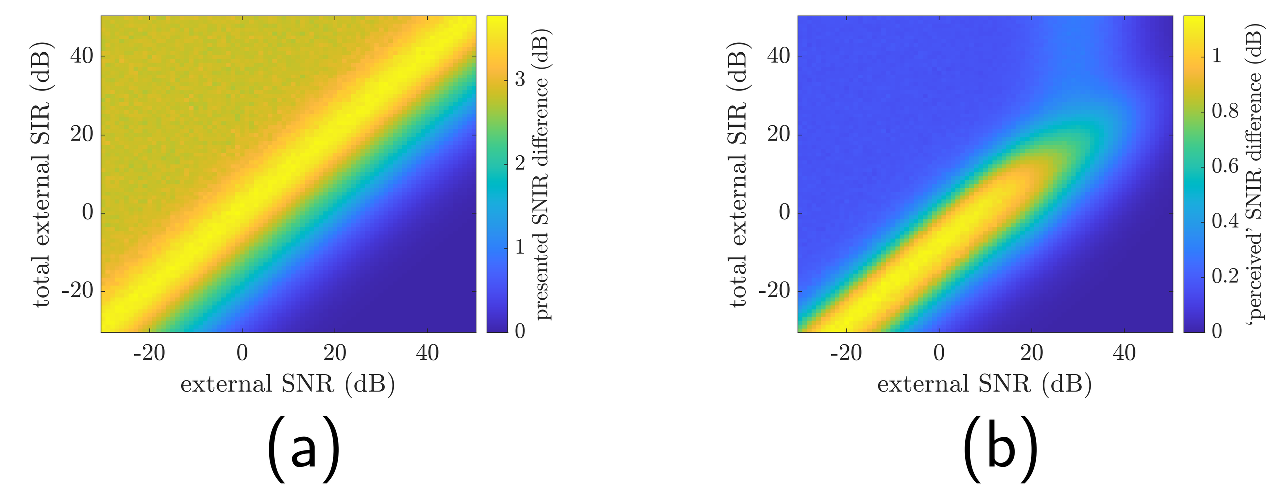

Figure 2: Simulated BMVDR–JBLCMV beamforming difference in SNIR of (a) the presented signals $\boldsymbol{y}$ and (b) the 'perceived' signal z.

SRM seems to have no significant effect on the intelligibility after beamforming is performed.

Optimal Beamforming

A novel beamformer is designed by optimising the 'perceived' SNIR while preserving the target,

$$\underset{\boldsymbol{v},\,\boldsymbol{W}}{\text{maximise}} \quad \frac{p_s \boldsymbol{v}^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{v}}{\boldsymbol{v}^{\mathrm{H}}(\boldsymbol{W}^{\mathrm{H}}\boldsymbol{P}_u\boldsymbol{W} + \boldsymbol{P}_n)\boldsymbol{v}} \qquad \text{subject to} \quad \boldsymbol{W}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{a}_{\mathrm{ref}}. \quad (4)$$

It can be shown both mathematically and through simulations that the resulting beamformer attains an identical 'perceived' SNIR to the BMVDR beamformer. There are, however, $M-1$ degrees of freedom in the solution that can be used to constrain interferer binaural cues.

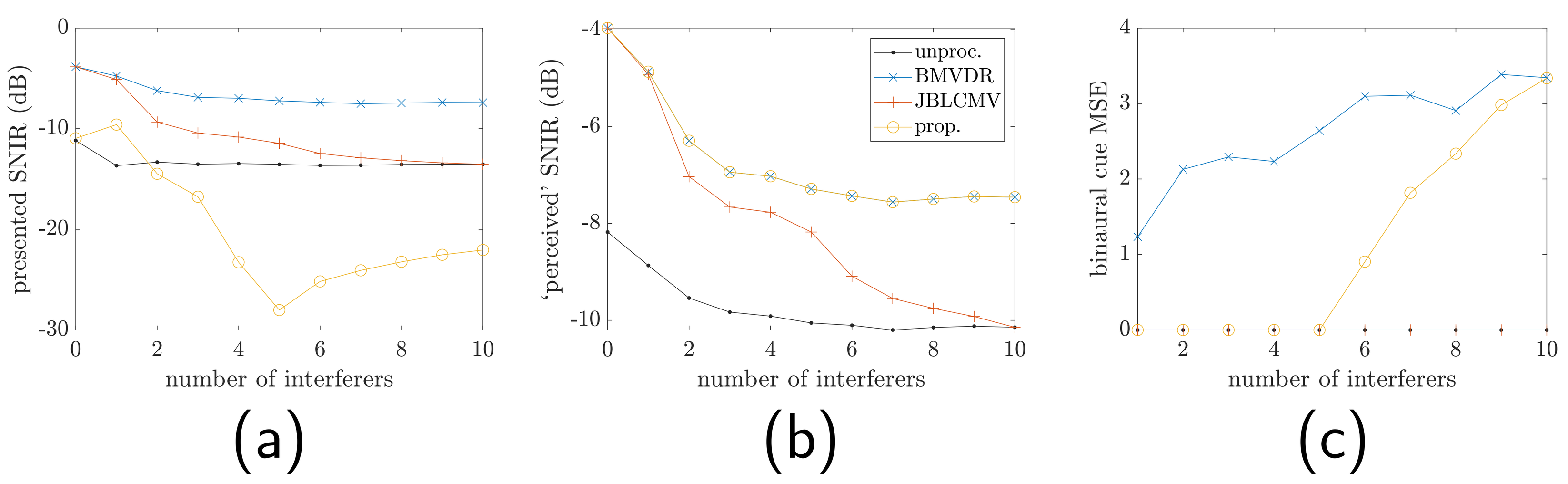

Figure 3: Simulated SNIR of (a) the presented signals $\boldsymbol{y}$ and (b) the 'perceived' signal z, and (c) simulated MSE on interferer binaural cues.

The proposed beamformer has an optimal performance in the 'perceived' domain, while also preserving the binaural cues of up to $M-1$ interferers.