A Broadband Beamformer Using Controllable Constraints and Minimum Variance

Karimian-Azari, Sam; Benesty, Jacob; Jensen, Jesper Rindom; Christensen, Mads Græsbøll

Published in:
2014 Proceedings of the 22nd European Signal Processing Conference (EUSIPCO 2014)

Publication date:
2014

Document Version
Accepted author manuscript, peer reviewed version

AUDIO ANALYSIS LAB

A Broadband Beamformer Using Controllable Constraints and Minimum Variance

Sam Karimian-Azari[†,1], Jacob Benesty[*], Jesper Rindom Jensen[†,2], and Mads Græsbøll Christensen[†]
email: [†]{ska, jrj, mgc}@create.aau.dk, and [*]benesty@emt.inrs.ca

AALBORG UNIVERSITY DENMARK

Introduction

The minimum variance distortionless response (MVDR) beamformer is an optimal approach to noise reduction:

- Achieves a high output SNR.
- It has degrees of freedom (DOF) corresponding to the number of microphones minus one.
- Its output is contaminated with both residual noise and interference.

The linearly constrained minimum variance (LCMV) beamformer reduces noise, and rejects interferers using linear constraints:

- Achieves a high output SIR.
- The number of constraints may degrade the DOF.
- It may amplify background noise which causes a lower output SNR.

To achieve a trade-off between attenuation of noise and interfering sources, we proposed the controllable LCMV (C-LCMV) beamformer in the frequency-domain.

Formulation

Multi-channel observed signals at the frequency f are observed, using an array of M microphones:

$$\mathbf{y}(f) = \mathbf{d}_1(f)X_1(f) + \sum_{n=2}^{N} \mathbf{d}_n(f)X_n(f) + \mathbf{v}(f) = \mathbf{D}_N(f)\mathbf{x}(f) + \mathbf{v}(f), \qquad (1)$$

where $\mathbf{d}_n(f) \in \mathbb{C}^M$ is the steering vector of signal source $X_n(f)$ (for $n = 1, \cdots, N$), $X_1(f)$ is the signal of interest (SOI), $\mathbf{x}(f) \in \mathbb{C}^N$ is the collected N signal sources, $\mathbf{v}(f) \in \mathbb{C}^M$ is noise, and

$$\mathbf{D}_N(f) = [\, \mathbf{d}_1(f)\ \mathbf{d}_2(f)\ \cdots\ \mathbf{d}_N(f)\,] \in \mathbb{C}^{M\times N}. \qquad (2)$$

The correlation matrix of $\mathbf{y}(f)$ (assuming uncorrelated signals) is

$$\mathbf{\Phi}_{\mathbf{y}}(f) = \mathbf{D}_N(f)\,\mathbf{\Phi}_{\mathbf{x}}(f)\,\mathbf{D}_N^H(f) + \mathbf{\Phi}_{\mathbf{v}}(f) = \mathbf{d}_1(f)\,\phi_{X_1}(f)\,\mathbf{d}_1^H(f) + \mathbf{\Phi}_{\mathrm{in}}(f), \qquad (3)$$

where $\mathbf{\Phi}_{\mathbf{x}}(f) = \mathrm{diag}[\,\phi_{X_1}(f)\ \phi_{X_2}(f)\ \ldots\ \phi_{X_N}(f)]$, $\mathbf{\Phi}_{\mathrm{in}}(f) = \mathbf{\Phi}_{\mathbf{i}}(f) + \mathbf{\Phi}_{\mathbf{v}}(f)$, and $\mathbf{\Phi}_{\mathbf{i}}(f) = \sum_{n=2}^{N} \mathbf{d}_n(f)\,\phi_{X_n}(f)\,\mathbf{d}_n^H(f)$ is the interference correlation matrix.

The output variance of the beamformer $\mathbf{h}(f)$ is

$$\phi_Z(f) = \mathbf{h}^H(f)\,\mathbf{d}_1(f)\phi_{X_1}(f)\,\mathbf{d}_1^H(f)\,\mathbf{h}(f) + \mathbf{h}^H(f)\,[\,\mathbf{\Phi}_{\mathbf{i}}(f) + \mathbf{\Phi}_{\mathbf{v}}(f)\,]\,\mathbf{h}(f). \qquad (4)$$

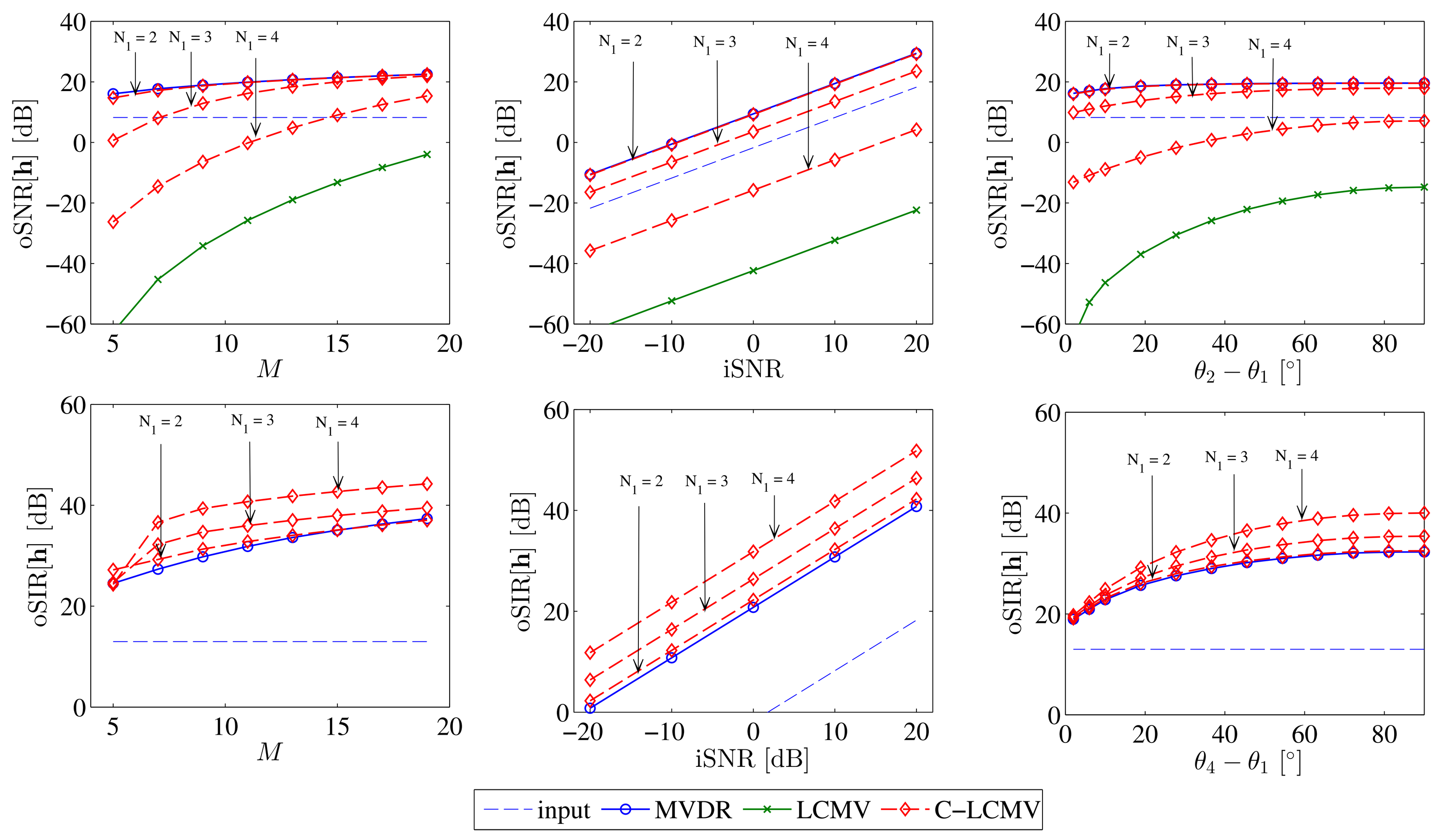
With the distortionless constraint that $\mathbf{h}^H(f)\mathbf{d}_1(f) = 1$, we can write

$$\mathrm{oSINR}[\mathbf{h}(f)] = \frac{\phi_{X_1}(f)}{\mathbf{h}^H(f)\,[\,\mathbf{\Phi}_{\mathrm{in}}(f) + \mathbf{\Phi}_{\mathbf{v}}(f)\,]\,\mathbf{h}(f)}, \qquad (5)$$

$$\mathrm{oSIR}[\mathbf{h}(f)] = \frac{\phi_{X_1}(f)}{\mathbf{h}^H(f)\,\mathbf{\Phi}_{\mathrm{in}}(f)\,\mathbf{h}(f)}. \qquad (6)$$

Experiment: Synthetic Signal

Using a uniform linear array (ULA) with $M = 10$ microphones and Gaussian noise signal sources in $\theta_1 = 0$, $\theta_2 = \pi$, $\theta_3 = 5\pi/6$, $\theta_4 = 4\pi/6$, and $\theta_5 = \pi/2$, where iSINR = 8 dB and iSIR = 13 dB:



C-LCMV Beamformer

MDVR:
$$\min_{\mathbf{h}(f)}\ \mathbf{h}^H(f)\,[\,\mathbf{\Phi}_{\mathbf{i}}(f) + \mathbf{\Phi}_{\mathbf{v}}(f)\,]\,\mathbf{h}(f) \qquad (7)$$
$$\text{subject to}\quad \mathbf{h}^H(f)\,\mathbf{d}_1(f) = 1,$$

LCMV:
$$\min_{\mathbf{h}(f)}\ \mathbf{h}^H(f)\,\mathbf{\Phi}_{\mathbf{v}}(f)\,\mathbf{h}(f) \qquad (8)$$
$$\text{subject to}\quad \mathbf{h}^H(f)\,\mathbf{D}_N(f) = \mathbf{i}_N^T,$$

where $\mathbf{i}_N$ is the first column of a $N \times N$ identity matrix.

We divide N signal sources into two sets of N_1 sources, containing SOI, and $N_2 = N - N_1$ remaining signal sources;

$$\mathbf{x}(f) = [\,\mathbf{x}_{N_1}^T(f)\ \mathbf{x}_{N_2}^T(f)\,]^T, \text{ and} \qquad (9)$$
$$\mathbf{D}_N(f) = [\,\mathbf{D}_{N_1}(f)\ \mathbf{D}_{N_2}(f)\,]. \qquad (10)$$

Therefore,

$$\mathbf{y}(f) = \mathbf{D}_{N1}(f)\,\mathbf{x}_{N_1}(f) + [\,\mathbf{D}_{N_2}(f)\,\mathbf{x}_{N_2}(f) + \mathbf{v}(f)\,], \qquad (11)$$

and

$$\mathbf{\Phi}_{\mathbf{y}}(f) = \mathbf{D}_{N_1}(f)\,\mathbf{\Phi}_{\mathbf{x}_{N_1}}(f)\,\mathbf{D}_{N_1}^H(f) + \mathbf{\Phi}_{\mathrm{in},N_2}(f), \qquad (12)$$

where $\mathbf{\Phi}_{\mathrm{in},N_2}(f) = \mathbf{D}_{N_2}(f)\,\mathbf{\Phi}_{\mathbf{x}_{N_2}}(f)\,\mathbf{D}_{N_2}^H(f) + \mathbf{\Phi}_{\mathbf{v}}(f)$ is the correlation matrix of N_2 signal sources plus background noise.

The C-LCMV beamformer is designed as

$$\min_{\mathbf{h}(f)}\ \mathbf{h}^H(f)\,\mathbf{\Phi}_{\mathrm{in},N_2}(f)\,\mathbf{h}(f) \qquad (13)$$
$$\text{subject to}\quad \mathbf{h}^H(f)\,\mathbf{D}_{N_1}(f) = \mathbf{i}_{N_1}^T.$$

Then the solution is given like

$$\mathbf{h}_{\mathrm{C}}(f) = \mathbf{\Phi}_{\mathrm{in},N_2}^{-1}(f)\,\mathbf{D}_{N_1}(f)\,[\,\mathbf{D}_{N_1}^H(f)\,\mathbf{\Phi}_{\mathrm{in},N_2}^{-1}(f)\mathbf{D}_{N_1}(f)\,]^{-1}\mathbf{i}_{N_1}. \qquad (14)$$

- Properties:

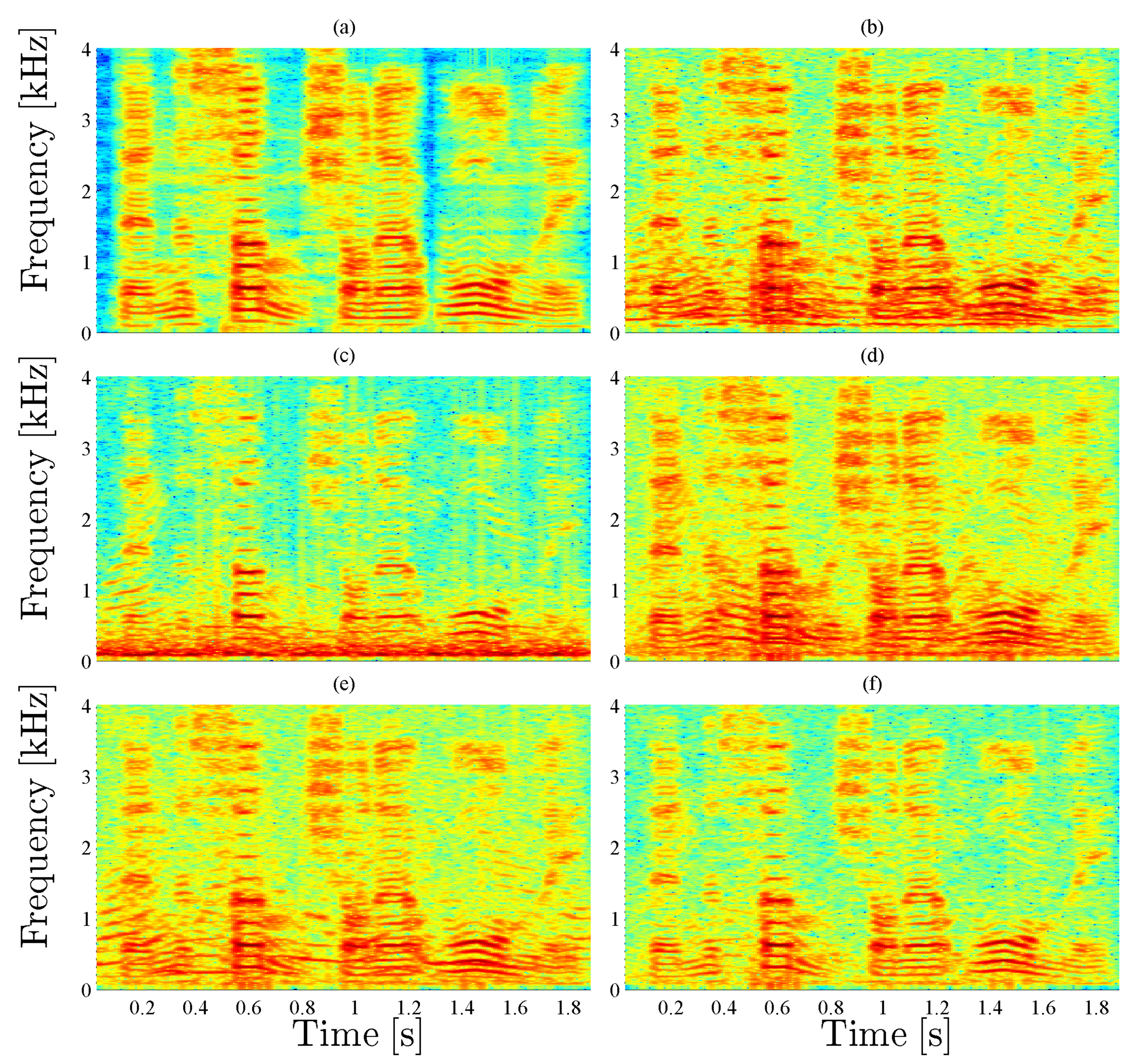
$$\mathrm{oSINR}\,[\mathbf{h}_{\mathrm{L}}(f)] \le \mathrm{oSINR}\,[\mathbf{h}_{\mathrm{C}}(f)] \le \mathrm{oSINR}\,[\mathbf{h}_{\mathrm{M}}(f)], \qquad (15)$$
$$\mathrm{oSIR}\,[\mathbf{h}_{\mathrm{M}}(f)] \le \mathrm{oSIR}\,[\mathbf{h}_{\mathrm{C}}(f)] \le \mathrm{oSIR}\,[\mathbf{h}_{\mathrm{L}}(f)]. \qquad (16)$$

- Optimal steering matrix (in practice):

$$\mathbf{D}_{N_1}^{\mathrm{opt}}(f) = \arg\max_{\mathbf{D}_{N_1}(f)} \mathrm{oSINR}\,[\mathbf{h}_{\mathrm{C}}(f)]. \qquad (17)$$

Experiment: Real Scenario

We simulated a room ($6 \times 7 \times 3$ m) with reverberation time $T_{60} = 0.25$ s, and $N = 3$ speech signals, and used a ULA with $M = 5$ hypercardioid microphones, SNR= 20 dB.



(a) SOI, (b-c) output of MVDR and LCMV beamformers, and (d-f) output of C-LCMV beamformers, where $\mathbf{D}_{N_1}(f) = [\mathbf{d}_1(f)\ \mathbf{d}_2(f)]$, $\mathbf{D}_{N_1}(f) = [\mathbf{d}_1(f)\ \mathbf{d}_3(f)]$, and $\mathbf{D}_{N_1}(f) = \mathbf{D}_{N_1}^{\mathrm{opt}}(f)$.

Conclusion

- The C-LCMV beamformer generalizes the MVDR and LCMV beamformers with the ability to control the quality of the output signal.
- Experiment results indicate that the C-LCMV beamformer using optimal steering matrix gets better results than the other MV beamformers.

Thanks to the Villum Foundation[1], and the Danish Council for Independent Research grant DFF-1337-00084[2]