

A DIRECTIONALY CONSTRAINED DISTORTIONLESS MULTISTAGE LCMV BEAMFORMER

Daniel Wolff, Yaakov Buchris and Israel Cohen

Technion, Israel Institute of Technology

Technion City, Haifa 32000, Israel

{dwolff@tx, bucristx, icohen@ee}.technion.ac.il

ABSTRACT

In this paper, we generalize the recently proposed multi-stage minimum variance distortionless response beamformer to all distortionless linearly constrained minimum variance (LCMV) beamformers with directional constraints. Given N_c constraints and an M -element microphone array, we propose to divide this array into K -element microphone sub-arrays ($N_c \leq K \leq M$) on which LCMV beamforming is performed. The K -element outputs are then used as new sensor inputs, and the operation is performed recursively until there is only one output at the last stage. The multistage LCMV beamformer satisfies all the imposed directional constraints but with reduced complexity compared to the conventional one. Simulation results show that in the presence of diffuse noise, the multistage LCMV beamformer achieves higher white noise gain than that of the classic LCMV beamformer, although the directivity factor is slightly decreased.

Index Terms— Linearly constrained minimum variance, microphone arrays, multistage, beamforming.

1. INTRODUCTION

The linearly constrained minimum variance (LCMV) beamformer has been developed in order to minimize the power of the output signals received by a microphone array given some linear directional constraints. Although the optimum solution is well known, it involves the inversion of the correlation matrices of the noise and of the noisy signal that are often ill-conditioned [1], leading to performance degradation of the white noise gain (WNG).

There are two main traditional approaches to improve the WNG of the LCMV beamformer. The first approach is diagonal loading (DL) [2, 3] and the second approach is generalized sidelobe canceller (GSC) [4]. The main idea in DL is to add a small positive constant to the diagonal elements of the correlation matrix before inversion, which is equivalent to adding white noise in the observation. The drawback of this

This Research was supported by Qualcomm Research Fund and MAFAAT-Israel Ministry of Defense.

solution is sensitivity to the choice of the loading parameter. Too small values keep the matrix ill-conditioned, while much large values will impair the superdirectivity of the LCMV beamformer. Moreover, that solution has the same computational complexity as the conventional LCMV. The second method consists of formulating the linearly constrained problem of the LCMV in an equivalent unconstrained problem. However, this method suffers from high sensitivity to estimation error of the correlation matrix and high computational complexity.

A third approach to solve the WNG problem has been recently proposed for the minimum variance distortionless response (MVDR) beamformer [1], which is a particular case of the LCMV beamformer [5]. It recursively divides the array into smaller sub-arrays and implements a reduced size MVDR on each. The output signals from all the sub-arrays are then fed as input signals to another stage of sub-arrays MVDR beamformers. This process is repeated until only a single sub-array is left. Simulation results show that the proposed method improves the WNG and also reduces the computational complexity.

In this paper, we generalize the solution of the multistage MVDR beamformer and derive the multistage LCMV beamformer. We first present the signal model and review the original LCMV beamforming formulation. Then, we develop the multistage LCMV beamformer and propose an implementation. Finally, we compare its performance with that of the conventional LCMV beamformer in a typical example.

2. SIGNAL MODEL

We consider a classic room acoustic signal model in which a uniform linear array (ULA) of M sensors captures a desired farfield source signal. Using the formulation in [6], the received signals are expressed in the frequency domain as

$$Y_m(\omega) = X_m(\omega) + V_m(\omega), \quad m = 1, 2, \dots, M, \quad (1)$$

where $X_m(\omega)$, $V_m(\omega)$ and $Y_m(\omega)$ are the source signal, the additive noise and the sensor signal at the m th sensor, respectively. All these signals are assumed to be real valued and zero-mean.

The main objective is to extract the signal $X_m(\omega)$, received at any reference microphone. Without any loss of generality, we choose to recover the signal $X_1(\omega)$ of the first microphone. We can write (1) as

$$\mathbf{y}(\omega) = \mathbf{x}(\omega) + \mathbf{v}(\omega) = \mathbf{d}(\omega)X_1(\omega) + \mathbf{v}(\omega), \quad (2)$$

where

$$\begin{aligned}\mathbf{y}(\omega) &\triangleq [Y_1(\omega) \ Y_2(\omega) \ \dots \ Y_M(\omega)]^T, \\ \mathbf{x}(\omega) &\triangleq [X_1(\omega) \ X_2(\omega) \ \dots \ X_M(\omega)]^T, \\ \mathbf{v}(\omega) &\triangleq [V_1(\omega) \ V_2(\omega) \ \dots \ V_M(\omega)]^T,\end{aligned}$$

$\mathbf{d}(\omega)$ is the steering vector, and the superscript T denotes transpose of a vector or a matrix. We refer to the first microphone as the reference microphone, and write the steering vector $\mathbf{d}(\omega)$ for incidence angle θ and sensor spacing δ :

$$\mathbf{d}(\omega) = [1 \ e^{-j\omega\delta/c \cos\theta} \ \dots \ e^{-j(M-1)\omega\delta/c \cos\theta}]. \quad (3)$$

where c is the speed of sound.

The signal and the noise are assumed uncorrelated so we can write the power spectral density (PSD) matrix of $\mathbf{y}(\omega)$ as

$$\begin{aligned}\Phi_{\mathbf{y}}(\omega) &\triangleq E[\mathbf{y}(\omega)\mathbf{y}^H(\omega)] = \Phi_{\mathbf{x}}(\omega) + \Phi_{\mathbf{v}}(\omega) \\ &= \phi_{X_1}(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega) + \Phi_{\mathbf{v}}(\omega),\end{aligned} \quad (4)$$

where the superscript H denotes the complex conjugate transpose operator, $\Phi_{\mathbf{x}}(\omega) \triangleq E[\mathbf{x}(\omega)\mathbf{x}^H(\omega)]$ and $\Phi_{\mathbf{v}}(\omega) \triangleq E[\mathbf{v}(\omega)\mathbf{v}^H(\omega)]$ are the PSD matrices of the signal and the noise and $\phi_{X_m}(\omega)$ is the variance of $X_m(\omega)$, $m = 1, 2, \dots, M$. We then perform beamforming by applying a complex weight to the output of each sensor and summing them together, i.e., $Z(\omega) \triangleq \mathbf{h}^H(\omega)\mathbf{y}(\omega)$, where $\mathbf{h}(\omega) = [H_1(\omega) \ H_2(\omega) \ \dots \ H_M(\omega)]^T$ is the vector containing those weights. Using (2), we get: $Z(\omega) = X_o(\omega) + V_o(\omega)$, where $X_o(\omega) \triangleq X_1(\omega)\mathbf{h}^H(\omega)\mathbf{d}(\omega)$ and $V_o(\omega) \triangleq \mathbf{h}^H(\omega)\mathbf{v}(\omega)$ are the signal and noise at the output of the beamformer. We deduce that the PSD of the beamformer output is given by

$$\phi_Z(\omega) = \phi_{X_1}(\omega)|\mathbf{h}^H(\omega)\mathbf{d}(\omega)|^2 + \mathbf{h}^H(\omega)\Phi_{\mathbf{v}}(\omega)\mathbf{h}(\omega). \quad (5)$$

3. CONVENTIONAL LCMV BEAMFORMER WITH DIRECTIONAL CONSTRAINTS

The aim of the LCMV beamformer is to minimize the PSD of the output noise subject to several linear constraints. Let N_c be the number of constraints ($N_c \leq M$), we define an $M \times N_c$ matrix, $\mathbf{C}(\omega)$, whose columns are linearly independent. The problem can be written as:

$$\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega)\Phi_{\mathbf{v}}(\omega)\mathbf{h}(\omega) \quad s.t. \quad \mathbf{h}^H(\omega)\mathbf{C}(\omega) = \mathbf{q}^H(\omega), \quad (6)$$

where $\mathbf{q}(\omega)$ is an $N_c \times 1$ vector containing the constraints. The solution of this problem is known to be [6]

$$\mathbf{h}_L(\omega) = \Phi_{\mathbf{v}}^{-1}(\omega)\mathbf{C}(\omega)[\mathbf{C}^H(\omega)\Phi_{\mathbf{v}}^{-1}(\omega)\mathbf{C}(\omega)]^{-1}\mathbf{q}(\omega). \quad (7)$$

Here, we only consider distortionless beamformers, meaning that when the sensors receive the desired signal without any noise, the signal must be perfectly recovered, i.e., $X_o(\omega) = X_1(\omega)$. This implies the condition $\mathbf{h}^H(\omega)\mathbf{d}(\omega) = 1$, that can be summarized as $\mathbf{C}_{(1)}(\omega) = \mathbf{d}(\omega)$ and $\mathbf{q}_1(\omega) = 1$ where $\mathbf{C}_{(l)}(\omega)$ is the l th column of $\mathbf{C}(\omega)$ and $\mathbf{q}_l(\omega)$ the l th element of $\mathbf{q}(\omega)$. The MVDR beamformer is the LCMV beamformer with the distortionless constraint only.

Among the typical constraints that can be added in an LCMV design, we choose to focus on directional constraints. Given a particular direction by the steering vector $\mathbf{d}_{(l)}(\omega)$, $l = 1, \dots, N_c$, we assign a response $\mathbf{q}_l(\omega)$, i.e., $\mathbf{h}^H(\omega)\mathbf{d}_{(l)}(\omega) = \mathbf{q}_l(\omega)$. Different values for $\mathbf{q}_l(\omega)$ can be chosen. For example, we can flatten the beampattern around the signal direction by adding constraints at two symmetric directions around the main direction [7]. We can suppress directional noise by setting a null ($\mathbf{q}_l(\omega) = 0$) in the unwanted direction [8]. Combining together all these constraints, we obtain:

$$\mathbf{C}(\omega) = [\mathbf{d}_{(1)}(\omega) \ \mathbf{d}_{(2)}(\omega) \ \dots \ \mathbf{d}_{(N_c)}(\omega)], \quad (8)$$

$$\mathbf{q}(\omega) = [\mathbf{q}_1(\omega) \ \mathbf{q}_2(\omega) \ \dots \ \mathbf{q}_{N_c}(\omega)]^T. \quad (9)$$

4. MULTISTAGE LCMV BEAMFORMER

4.1. Derivation

To compute the LCMV beamformer, we need to invert the $M \times M$ matrix $\Phi_{\mathbf{v}}(\omega)$ in each frequency. The first issue is that usually, this matrix is ill-conditioned and lead to very unstable results. Moreover, the computational complexity of the matrix inversion goes like $\mathcal{O}(M^3)$. That motivates to generalize the multistage MVDR beamformer, which is proposed to solve these issues, to all LCMV beamformers with directional constraints [1]. Indeed, we propose to apply the beamformer to smaller sets of K sensors in $N_s = \log M / \log K$ successive stages ($K \geq N_c$). For every stage n , we apply LCMV to M/K^n sets of K elements. This allows to compute inversions of $K \times K$ matrices instead of $M \times M$ matrices.

For every stage $n = 0, 1, 2, \dots, N_s - 1$, we perform the LCMV to sets of K sensors, such that the output $Z_i^{(n+1)}(\omega)$ is a new estimate of $X_i(\omega)$. For $i = 1, \dots, N/K^{n+1}$, we define sets of K sensors input:

$$\begin{aligned}\mathbf{z}_i^{(n)}(\omega) &= [Z_{K(i-1)+1}^{(n)}(\omega) \ \dots \ Z_{Ki}^{(n)}(\omega)]^T \\ &= [X_{K(i-1)+1}^{(n)}(\omega) \ \dots \ X_{Ki}^{(n)}(\omega)]^T + \mathbf{v}_i^{(n)}(\omega)\end{aligned} \quad (10)$$

Then, we perform the LCMV beamformer, choosing the i th microphone as reference. We can rewrite $\mathbf{x}(\omega)$ as

$$\mathbf{x}(\omega) = X_1(\omega)\mathbf{d}(\omega) = \frac{X_i(\omega)}{D_i(\omega)}\mathbf{d}(\omega), \quad (11)$$

where $D_i(\omega)$ is the i th element of $\mathbf{d}(\omega)$, for $i = 1, \dots, M/K^n$. We deduce that for all constraints $l = 1, \dots, N_c$, the steering vector of the K -sensor input $\mathbf{z}_i^{(n)}(\omega)$ is

$$\mathbf{d}_{(l),i}(\omega) = (D_{(l),i}(\omega))^{-1} \cdot [D_{(l),K(i-1)+1}(\omega) \ \dots \ D_{(l),Ki}(\omega)]^T, \quad (12)$$

We can apply to this set the LCMV beamformer for the new estimate $Z_i^{(n+1)}(\omega)$ of $X_i(\omega)$:

$$Z_i^{(n+1)}(\omega) = \mathbf{h}_i^{(n)H}(\omega) \mathbf{z}_i^{(n)}(\omega), \quad (13)$$

with

$$\mathbf{h}_i^{(n)}(\omega) = \Phi_{\mathbf{v}_i^{(n)}}^{-1}(\omega) \mathbf{C}_i(\omega) \left[\mathbf{C}_i^H(\omega) \Phi_{\mathbf{v}_i^{(n)}}^{-1}(\omega) \mathbf{C}_i(\omega) \right]^{-1} \tilde{\mathbf{q}}(\omega), \quad (14)$$

where

$$\mathbf{C}_i(\omega) = [\mathbf{d}_{(1),i}(\omega) \ \dots \ \mathbf{d}_{(N_c),i}(\omega)]^T, \quad (15)$$

$$\tilde{\mathbf{q}}(\omega) = [\mathbf{q}_1^{1/N_s}(\omega) \ \dots \ \mathbf{q}_{N_c}^{1/N_s}(\omega)]^T. \quad (16)$$

Using (5) we get the output noise PSD matrix:

$$\Phi_{\mathbf{v}_i^{(n+1)}}(\omega) = \mathbf{h}_i^{(n)H}(\omega) \Phi_{\mathbf{v}_i^{(n)}}(\omega) \mathbf{h}_i^{(n)}(\omega). \quad (17)$$

The replacement of $\mathbf{q}(\omega)$ with $\tilde{\mathbf{q}}(\omega)$ enables to keep the same directional constraints after N_s LCMV beamforming.

4.2. Implementation

In Algorithm 1, we present the multistage LCMV beamformer implementation. The noise signal at the n th stage is stacked into a vector and a global filter $\mathbf{H}_n(\omega)$ of size $K^{N_s-n} \times K^{N_s-n-1}$ is defined such that $N_s - n - 1$ beamformers are applied to groups of K sensors:

$$\mathbf{H}^{(n)}(\omega) = \begin{pmatrix} \mathbf{h}_1^{(n)}(\omega) & 0_{K \times 1} & \dots & 0_{K \times 1} \\ 0_{K \times 1} & \mathbf{h}_2^{(n)}(\omega) & \dots & 0_{K \times 1} \\ \dots & \dots & \ddots & \dots \\ 0_{K \times 1} & 0_{K \times 1} & \dots & \mathbf{h}_{K^{N_s-n-1}}^{(n)}(\omega) \end{pmatrix}. \quad (18)$$

The combined filter is given by

$$\mathbf{h}_{multi}(\omega) = \mathbf{H}^{(0)}(\omega) \mathbf{H}^{(1)}(\omega) \dots \mathbf{H}^{(N_s-1)}(\omega). \quad (19)$$

If we have null constraints, the signal from that direction will be nulled after the first stage. Therefore, we can remove the null constraints in $\mathbf{C}(\omega)$ in the next stages. That is why we remove the columns of $\mathbf{C}(\omega)$ corresponding to the null constraints in the second stage of the algorithm.

This beamformer follows the condition $\mathbf{h}^H(\omega) \mathbf{C}(\omega) = \mathbf{q}(\omega)$ although it does not minimize the variance of the residual noise as the conventional LCMV beamformer, but the simulations show that the loss of performances is limited. Moreover, the complexity of the multistage beamformer is decreased by the fact we invert $K \times K$ instead of $M \times M$ matrices. In the case that the number of constraints N_c is small compared to M , the computational complexity of the beamformer at every frequency falls from $\mathcal{O}(M^3)$ to $\mathcal{O}(K^2 M)$.

Algorithm 1 Multistage Directional LCMV Beamformer

Input: input noise PSD $\Phi_{\mathbf{v}^{(0)}}(\omega) = \Phi_{\mathbf{v}}(\omega)$, input signals $\mathbf{z}^{(0)}(\omega) = \mathbf{y}(\omega)$, LCMV constraints $\mathbf{C}(\omega)$ and $\mathbf{q}(\omega)$

Filtering:

for $n = 0, 1, \dots, N_s - 1$ **do**

if $\mathbf{n}=1$ **then**

for all l where $\mathbf{q}_l = 0$ **do** remove column \mathbf{C}_l from \mathbf{C} and \mathbf{q}_l from \mathbf{q}

end for

end if

for $i = 0, 1, \dots, K^{N_s-n-1}$, **do**

$$\Phi_{\mathbf{v}_i^{(n)}}(\omega) = [\Phi_{\mathbf{v}_i^{(n)}}(\omega)]_{K(i-1)+1:Ki, K(i-1)+1:Ki}$$

$$\mathbf{h}_i^{(n)}(\omega) = \Phi_{\mathbf{v}_i^{(n)}}^{-1}(\omega) \mathbf{C}_i(\omega) \left[\mathbf{C}_i^H(\omega) \Phi_{\mathbf{v}_i^{(n)}}^{-1}(\omega) \mathbf{C}_i(\omega) \right]^{-1} \tilde{\mathbf{q}}(\omega)$$

end for

$$\mathbf{H}^{(n)}(\omega) = \begin{pmatrix} \mathbf{h}_1^{(n)}(\omega) & 0_{K \times 1} & \dots & 0_{K \times 1} \\ 0_{K \times 1} & \mathbf{h}_2^{(n)}(\omega) & \dots & 0_{K \times 1} \\ \dots & \dots & \ddots & \dots \\ 0_{K \times 1} & 0_{K \times 1} & \dots & \mathbf{h}_{K^{N_s-n-1}}^{(n)}(\omega) \end{pmatrix},$$

$$\mathbf{z}^{(n+1)}(\omega) = \mathbf{H}^{(n)H}(\omega) \mathbf{z}^{(n)}(\omega)$$

$$\Phi_{\mathbf{v}^{(n+1)}}(\omega) = \mathbf{H}^{(n)H}(\omega) \Phi_{\mathbf{v}^{(n)}}(\omega) \mathbf{H}^{(n)}(\omega)$$

end for

Output: $Z_1^{N_s}(\omega) = \mathbf{z}^{N_s}(\omega)$ {Signal Estimation}

5. PERFORMANCE MEASURES

The array gain is defined as the ratio between the output and the input signal-to-noise ratios (SNR). Using (5), and the distortionless condition, we obtain the array gain:

$$\begin{aligned} \mathcal{A}(\omega) &\triangleq oSNR(\omega)/iSNR(\omega) \\ &= \phi_{V_o}(\omega) [\mathbf{h}^H(\omega) \Phi_{\mathbf{v}}(\omega) \mathbf{h}(\omega)]^{-1}. \end{aligned} \quad (20)$$

In order to see the sensor noise amplification, we also compute the WNG defined as the gain in presence of white noise, $\mathcal{A}_{wn}(\omega) = [\mathbf{h}^H(\omega) \mathbf{h}(\omega)]^{-1}$.

We analyze the performance in presence of diffuse noise, because it involves the inversion of very ill-conditioned matrices as shown in [1]. In that case we take advantage of the dimensional reduction of the problem. The PSD has the form $\Phi_{\mathbf{v}}(\omega) = \phi_{V_1}(\omega) \Gamma(\omega)$, with the pseudo-coherence matrix $[\Gamma(\omega)]_{i,j} = \text{sinc}(2\omega\delta_{i,j}/c)$, where $\delta_{i,j}$ is the distance between the i th and the j th sensors. The array gain for this choice of $\Phi_{\mathbf{v}}(\omega)$ is known as the directivity index (DI).

6. EXPERIMENTAL RESULTS

To evaluate the performance, we choose a ULA with a sensor spacing of 2 cm. We apply a superdirective LCMV beamformer, meaning that we set the input noise in the filter as diffuse noise. We assume that the desired signal is in the end-fire direction, where the results are optimal. Besides the dis-

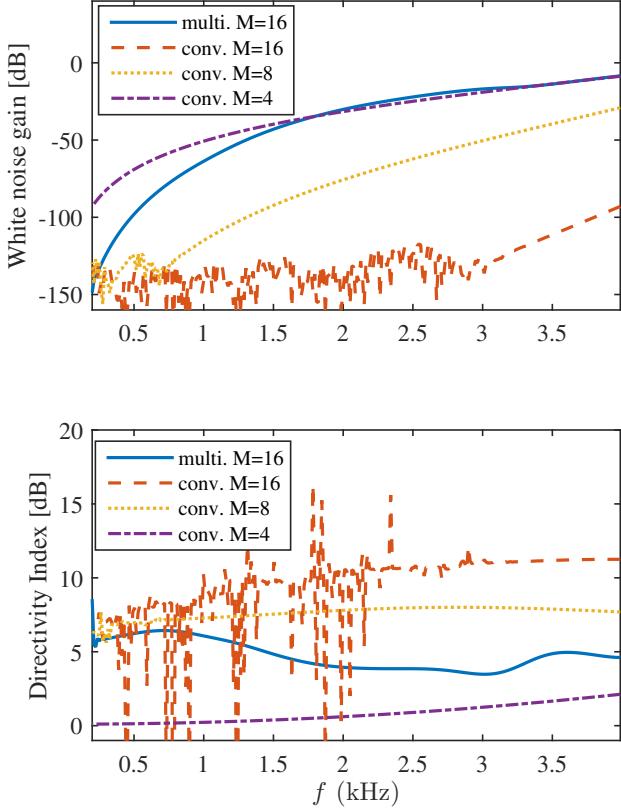


Fig. 1: Performance of the multistage LCMV beamformer compared to that of the conventional implementation.

tortionless constraint, we set a null constraint at angle θ_1 to deal with one farfield directional interference source. This solution is very practical when the noise signals are highly non-stationary because we cannot estimate their PSD [8]. We also want to force the beamformer with a gain of 0.9 at angle θ_2 . As we have three constraints, we choose to perform this test with sub-arrays with respect to one degree of freedom, i.e., with $K = 4$ microphones. The set of parameters is

$$\begin{cases} M = 16 \quad K = 4 \quad N_s = 2, \quad \theta_1 = \pi/3 \quad \theta_2 = \pi/2 - 0.1 \\ C(\omega) = \begin{bmatrix} d_{\theta=0}(\omega) & d_{\theta_1}(\omega) & d_{\theta_2}(\omega) \end{bmatrix}^T, q = \begin{bmatrix} 1 & 0 & 0.9 \end{bmatrix}^T. \end{cases} \quad (21)$$

Figure 1 compares the performances of the multistage beamformer and the conventional LCMV beamformer for the cases of 4, 8 and 16 sensors. Firstly, we see that the conventional 8 and 16 microphone beamformer is stable only above 0.8 and 3 kHz, respectively. Indeed, Figure 2 shows the eigenvalues of $\Phi_v(\omega)$, and as M increases and f decreases, the eigenvalues come closer to zero. With $K = 4$, the multistage beamformer remains stable for the whole frequency range.

The performance of the multistage beamformer with 16 sensors in white noise gain is very close to that of the conventional beamformer with 4 sensors, with a gain of more than 20

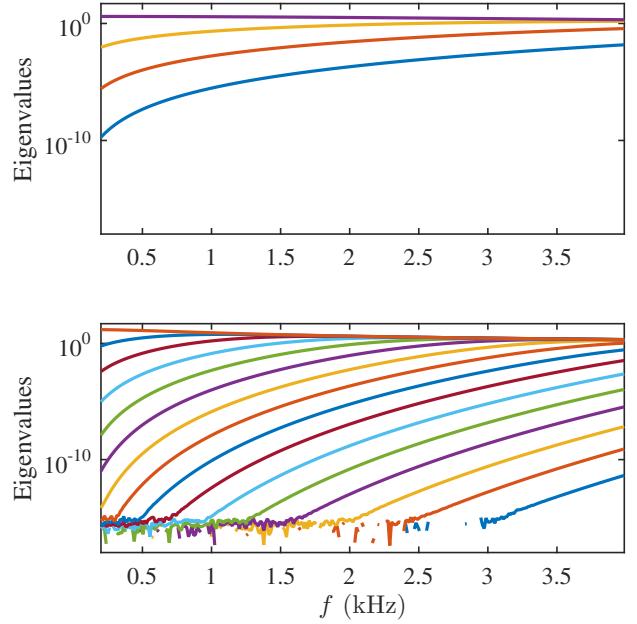


Fig. 2: Eigenvalues of the pseudocoherence matrix of the diffuse noise with 4 (top) and 16 (bottom) microphones.

dB and 80 dB compared to the conventional beamformer with 8 and 16 microphones, respectively. In contrast, the DI of the multistage LCMV beamformer is slightly decreased with respect to the conventional 16-sensor LCMV. This result is expected since in the multistage, we implement several beamformers, but with reduced DI with respect to the conventional one.

7. CONCLUSION

We have developed a multistage LCMV beamformer with directional constraints as a generalization of the multistage MVDR beamformer. We have divided the sensor array into smaller sub-arrays, that produced a new small sized microphone array. This recursive action leads to a beamformer that preserves the conditions of the conventional LCMV beamformer. This method improves the robustness and computational cost of the LCMV beamformer for a high number of microphones. Despite a slight decrease of the DI, this method considerably restricts the white noise amplification.

8. REFERENCES

- [1] C. Pan, J. Chen and J. Benesty, "A multistage minimum variance distortionless response beamformer for noise reduction," *The Journal of the Acoustical Society of America*, vol. 137, no. 3, pp. 1377-1388, 2015.
- [2] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," in *IEEE*

Transactions on Aerospace and Electronic Systems,
vol. 24, no. 4, pp. 397-401, Jul 1988.

- [3] J. Li, P. Stoica and Z. Wang, "On robust Capon beam-forming and diagonal loading," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2003. , pp. V-337-340.
- [4] B. R. Breed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Lett.*, vol. 9, pp. 168-169, 2002
- [5] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," in *Proc. of the IEEE*, vol. 60, no. 8, pp. 926-935, Aug. 1972.
- [6] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.
- [7] H. L. Van-trees, *Detection, Estimation and Modulation Theory, Part IV - Optimum Array Processing*. New York: Wiley Interscience, 2002.
- [8] E. A. P. Habets, J. Benesty, S. Gannot, P. A. Naylor, and I. Cohen, "On the application of the LCMV beamformer to speech enhancement," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA)*, New York, Oct. 2009, pp. 141-144.