ON THE APPLICATION OF THE LCMV BEAMFORMER TO SPEECH ENHANCEMENT

Emanuël A. P. Habets[1], *Jacob Benesty*[2], *Sharon Gannot*[3], *Patrick A. Naylor*[1], *and Israel Cohen*[4]

[1] Imperial College
London, SW7 2AZ, United Kingdom
{e.habets,p.naylor}@ic.ac.uk

[2] INRS-EMT, University of Quebec
Montreal, H5A 1K6, Canada
benesty@emt.inrs.ca

[3] Bar-Ilan University
Ramat-Gan, 52900, Israel
gannot@eng.biu.ac.il

[4] Technion - IIT
Haifa, 32000, Israel
i.cohen@technion.ac.il

ABSTRACT

In theory the linearly constrained minimum variance (LCMV) beamformer can achieve perfect dereverberation and noise cancellation when the acoustic transfer functions (ATFs) between all sources (including interferences) and the microphones are known. However, blind estimation of the ATFs remains a difficult task. In this paper the noise reduction of the LCMV beamformer is analyzed and compared with the noise reduction of the minimum variance distortionless response (MVDR) beamformer. In addition, it is shown that the constraint of the LCMV can be modified such that we only require relative transfer functions rather than ATFs to achieve perfect coherent noise reduction. Finally, we evaluate the noise reduction performance achieved by the LCMV and MVDR beamformers for two coherent sources: one desired and one undesired.

Index Terms— Linearly constrained minimum variance (LCMV) filter, noise reduction, speech enhancement, microphone arrays, beamforming

1. INTRODUCTION

Distant or hands-free audio acquisition is required in many applications such as teleconferencing. Microphone arrays are often used for the acquisition and consist of sets of microphone sensors that are arranged in specific geometries. The received sensor signals usually consist of a mixture of one or more desired signals, coherent and non-coherent interferences. The received signals are processed in order to extract the desired signals, or in other words to suppress the interferences. In the last four decades many algorithms have been proposed to process the received sensor signals [1, 2].

Several researchers developed beamformers in which multiple linear constraints were imposed (e.g., Er and Cantoni [3]). These beamformers are known as linearly constrained minimum variance (LCMV) beamformers. The celebrated minimum variance distortionless response (MVDR) beamformer is a special case that uses a single constraint towards the desired source. In [4], Frost proposed an adaptive scheme of the MVDR beamformer, which is based on a constrained least mean square-type adaptation. To avoid the constrained adaptation of the MVDR beamformer, Griffiths and Jim [5] proposed the generalized sidelobe canceller (GSC) structure that separates the output power minimization and the application of the constraint. While Griffiths and Jim only considered one constraint, it was later shown in [6] that the GSC structure can also be used in the case of multiple constraints. The original GSC structure is based on the assumption that the different sensors receive a delayed version of the desired signal. The GSC structure was later re-derived in the frequency domain, and extended to deal with general acoustic transfer functions (ATFs) [7].

In the acoustic signal processing community, the LCMV beamformer has received considerably less attention compared to the MVDR. In theory, the LCMV beamformer can achieve perfect dereverberation and noise cancellation when the ATFs between all sources (including interferences) and microphones are known [8].

In this paper we study the LCMV in an acoustic environment. After formulating the problem and reviewing the LCMV beamformer we first define an objective measure to evaluate the noise reduction of the LCMV beamformer. Secondly, we analyze the noise reduction capability for one desired source and one undesired source in a homogeneous and spatially white noise field and compare with the MVDR beamformer derived in [9]. Thirdly, we modify the constraint of the LCMV beamformer in such a way that we can perfectly cancel the noise using relative transfer functions (RTFs) rather than ATFs. Finally, we evaluate the noise reduction performance of the LCMV and MVDR beamformers with respect to a displacement of the noise source.

2. PROBLEM FORMULATION

Consider the signal model in which an M-element sensor array captures K coherent source signals in some noise field. We assume that all signals are broadband, and that all source signals and the noise signals are mutually independent and zero mean. In the discrete-time Fourier transform (DTFT) domain the received signals are expressed as

$$\begin{aligned} Y_m(\omega) &= \sum_{k=1}^{K} G_{k,m}(\omega)S_k(\omega) + V_m(\omega) \qquad (1) \\ &= \sum_{k=1}^{K} X_{k,m}(\omega) + V_m(\omega),\ m = 1, 2, \dots, M, \end{aligned}$$

where $Y_m(\omega)$, $G_{k,m}(\omega)$, $S_k(\omega)$, $X_{k,m}(\omega) = G_{k,m}(\omega)S_k(\omega)$, and $V_m(\omega)$ are the DTFTs of the mth sensor signal, the impulse response from the kth source to the mth sensor, the kth source signal, the reverberant signal of the kth source received by the mth sensor, and the additive noise at sensor m, respectively, at angular frequency ω $(-\pi < \omega \leq \pi)$.

Our main objective in this paper is then to study the recovering of a mixture of $K_\mathrm{d} \leq K$ desired sources received by a reference microphone (noise reduction only). Without loss of generality, we consider the first microphone as the reference microphone. The number of undesired coherent sources is given by$K_\mathrm{u} = K - K_\mathrm{d}$.

This research was partially supported by the EU project SCENIC.

The M sensor signals in the frequency domain are better summarized in a vector notation as

$$\begin{aligned} \mathbf{y}(\omega) &= \mathbf{G}(\omega)\mathbf{s}(\omega) + \mathbf{v}(\omega) \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega), \end{aligned} \tag{2}$$

where

$$\begin{aligned} \mathbf{y}(\omega) &= \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T, \\ \mathbf{G}(\omega) &= \begin{bmatrix} \mathbf{g}_1(\omega) & \mathbf{g}_2(\omega) & \cdots & \mathbf{g}_M(\omega) \end{bmatrix}^T, \\ \mathbf{g}_m(\omega) &= \begin{bmatrix} G_{1,m}(\omega) & G_{2,m}(\omega) & \cdots & G_{K,m}(\omega) \end{bmatrix}^T, \\ \mathbf{s}(\omega) &= \begin{bmatrix} S_1(\omega) & S_2(\omega) & \cdots & S_K(\omega) \end{bmatrix}^T, \\ \mathbf{v}(\omega) &= \begin{bmatrix} V_1(\omega) & V_2(\omega) & \cdots & V_M(\omega) \end{bmatrix}^T, \\ \mathbf{x}(\omega) &= \begin{bmatrix} X_1(\omega) & X_2(\omega) & \cdots & X_M(\omega) \end{bmatrix}^T, \\ X_m(\omega) &= \mathbf{g}_m^T(\omega)\mathbf{s}(\omega), \end{aligned}$$

and superscript ^{T} denotes transpose of a vector or a matrix.

The power spectral density (PSD) of the received signal at the mth sensor can be expressed as

$$\begin{aligned} \phi_{y_m}(\omega) &= \phi_{x_m}(\omega) + \phi_{v_m}(\omega) \\ &= \mathbf{g}_m^H(\omega)\mathbf{\Lambda}_s(\omega)\mathbf{g}_m(\omega) + \phi_{v_m}(\omega), \end{aligned} \tag{3}$$

for $m = 1, 2, \ldots, M$, where $\phi_{y_m}(\omega)$, $\phi_{x_m}(\omega)$, $\mathbf{\Lambda}_s(\omega) = \text{diag}\left[\phi_{s_1}(\omega), \ldots, \phi_{s_K}(\omega)\right]$, and $\phi_{v_m}(\omega)$ are the PSDs of the mth sensor signal, the mth sensor reverberant signal, the coherent signals, and the mth sensor noise signal, respectively.

The beamforming is then performed by applying a complex weight to each sensor and summing across all sensors:

$$Z(\omega) = \mathbf{h}^H(\omega)\mathbf{y}(\omega) = \mathbf{h}^H(\omega)\left[\mathbf{G}(\omega)\mathbf{s}(\omega) + \mathbf{v}(\omega)\right], \tag{4}$$

where $Z(\omega)$ is the beamformer output, $\mathbf{h}(\omega) = [H_1(\omega)\, H_2(\omega)\, \ldots\, H_M(\omega)]^T$ is the beamforming weight vector which is suitable for performing spatial filtering at frequency ω, and superscript ^{H} denotes transpose conjugation of a vector or a matrix.

The PSD of the beamformer output is given by

$$\phi_z(\omega) = \mathbf{h}^H(\omega)\mathbf{\Phi}_x(\omega)\mathbf{h}(\omega) + \mathbf{h}^H(\omega)\mathbf{\Phi}_v(\omega)\mathbf{h}(\omega), \tag{5}$$

where

$$\mathbf{\Phi}_x(\omega) = E\left[\mathbf{x}(\omega)\mathbf{x}^H(\omega)\right] = \mathbf{G}(\omega)\mathbf{\Lambda}_s(\omega)\mathbf{G}^H(\omega) \tag{6}$$

is the PSD matrix of the convolved speech signals with $E(\cdot)$ denoting mathematical expectation, and

$$\mathbf{\Phi}_v(\omega) = E\left[\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right] \tag{7}$$

is the PSD matrix of the noise field. In the rest of this paper, we assume that $\mathbf{\Phi}_v(\omega)$ is a full-rank matrix such that its inverse exists.

3. LINEARLY CONSTRAINED MINIMUM VARIANCE BEAMFORMER

We now derive the LCMV bamformer in the context of room acoustics. The LCMV filter is conceived by providing a fixed gain, in our case modelled by a column vector $\mathbf{q}(\omega)$ of length K, to the signals while utilizing the remaining degrees of freedom to minimize the contribution of the additive noise to the array output:

$$\mathbf{h}_{\text{LCMV}}(\omega) = \arg\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega)\mathbf{\Phi}_v(\omega)\mathbf{h}(\omega) \quad \text{subject to } \mathbf{h}^H(\omega)\mathbf{G}(\omega) = \mathbf{q}^H(\omega). \tag{8}$$

In this paper we assume that the fixed gain of the K_u undesired sources is set to zero. It should be noted that the LCMV filter in constructed using only spatial information related to the undesired sources (given by $\mathbf{G}$), i.e., we do not require the PSDs of the undesired source signals. This makes the beamformer especially attractive in a senario where the undesired sources are highly non-stationary (which makes the estimation of the related PSD matrix difficult) and their spatial position is fixed or slowly time-varying.

The solution of (8) is given by [1]

$$\mathbf{h}_{\text{LCMV}}(\omega) = \mathbf{\Phi}_v^{-1}(\omega)\mathbf{G}(\omega)\left[\mathbf{G}^H(\omega)\mathbf{\Phi}_v^{-1}(\omega)\mathbf{G}(\omega)\right]^{-1}\mathbf{q}(\omega). \tag{9}$$

4. NOISE REDUCTION MEASURE

In this section we define a noise reduction measure that indicates the level of noise reduction achieved through beamforming and helps us understand how noise reduction works with the LCMV beamformer in an acoustic environment. In order to differentiate between desired and undesired sources we define a column vector $\mathbf{u}$ with K_d zeros and K_u ones, where the ones indicate the undesired sources. A diagonal matrix that only contains the PSDs of all undesired sources is then given by $\mathbf{\Lambda}_{s,\text{u}}(\omega) = \mathbf{\Lambda}_s(\omega)\,\text{diag}(\mathbf{u})$.

We define the local noise-reduction factor as the ratio of the PSD of the original noise (coherent and non-coherent) at the reference microphone over the PSD of the residual noise:

$$\xi_{\text{nr}}\left[\mathbf{h}(\omega)\right] = \frac{\mathbf{g}_1^H(\omega)\mathbf{\Lambda}_{s,\text{u}}(\omega)\mathbf{g}_1(\omega) + \phi_{v_1}(\omega)}{\mathbf{h}^H(\omega)[\mathbf{G}(\omega)\mathbf{\Lambda}_{s,\text{u}}(\omega)\mathbf{G}^H(\omega) + \mathbf{\Phi}_v(\omega)]\mathbf{h}(\omega)}. \tag{10}$$

Integrating across the entire frequency range in the numerator and denominator of (10) yields the global noise-reduction factor:

$$\xi_{\text{nr}}(\mathbf{h}) = \frac{\int_{-\pi}^{\pi}\left[\mathbf{g}_1^H(\omega)\mathbf{\Lambda}_{s,\text{u}}(\omega)\mathbf{g}_1(\omega) + \phi_{v_1}(\omega)\right]d\omega}{\int_{-\pi}^{\pi}\mathbf{h}^H(\omega)[\mathbf{G}(\omega)\mathbf{\Lambda}_{s,\text{u}}(\omega)\mathbf{G}^H(\omega) + \mathbf{\Phi}_v(\omega)]\mathbf{h}(\omega)d\omega}. \tag{11}$$

For the LCMV filter we obtain

$$\xi_{\text{nr}}\left[\mathbf{h}_{\text{LCMV}}(\omega)\right] = \frac{\mathbf{g}_1^H(\omega)\mathbf{\Lambda}_{s,\text{u}}(\omega)\mathbf{g}_1(\omega) + \phi_{v_1}(\omega)}{\mathbf{h}_{\text{LCMV}}^H(\omega)\mathbf{\Phi}_v(\omega)\mathbf{h}_{\text{LCMV}}(\omega)}. \tag{12}$$

The global noise-reduction factor $\xi_{\text{nr}}\left(\mathbf{h}_{\text{LCMV}}\right)$ can be obtained by integrating across the entire frequency range in the numerator and denominator of (12).

In theory the denominator of the local and global noise-reduction factors of the LCMV filter is independent of the undesired coherent noise sources. However, in practice we employ $\hat{\mathbf{h}}_{\text{LCMV}}(\omega)$ that is computed using (8) and $\hat{\mathbf{G}}(\omega)$ rather than $\mathbf{G}(\omega)$. When $\hat{\mathbf{G}}(\omega) \neq \mathbf{G}(\omega)$ we have $\hat{\mathbf{h}}_{\text{LCMV}}(\omega)\mathbf{G}(\omega) \neq \mathbf{q}^H(\omega)$. Although $\xi_{\text{nr}}\left(\mathbf{h}_{\text{LCMV}}\right)$ is important from a theoretical point of view we will employ (11) for the performance evaluation in Section 7.

5. LOCAL NOISE REDUCTION ANALYSIS

In this section we analyze the noise reduction performance of the LCMV and MVDR filters. Let us assume one desired source and one undesired source, i.e., $K = 2, K_{\mathrm{d}} = 1, K_{\mathrm{u}} = 1$. The desired and undesired source signals are given by $\sigma_d(\omega)\,\mathbf{d}(\omega)$ and $\sigma_b(\omega)\,\mathbf{b}(\omega)$, respectively. Here $\mathbf{d}(\omega) = [D_1(\omega)\,\ldots\,D_M(\omega)]^T$ and $\mathbf{b}(\omega) = [B_1(\omega)\,\ldots\,B_M(\omega)]^T$ are column vectors of length M that contain the ATFs between the desired and undesired source and the microphones, respectively. In addition, we have a homogeneous and spatially white noise field with variance $\sigma^2_{\mathrm{nc}}(\omega)$. Our objective is to recover the desired source as received by the first microphone.

We now have $\mathbf{G}(\omega) = [\mathbf{d}(\omega)\,\mathbf{b}(\omega)]$, $\mathbf{q}(\omega) = [D_1^*(\omega)\,0]^T$, and $\mathbf{\Phi}_v(\omega) = \sigma^2_{\mathrm{nc}}(\omega)\mathbf{I}$. In this case the LCMV filter is given by

$$\mathbf{h}_{\mathrm{LCMV}}(\omega) = \mathbf{G}(\omega)\left[\mathbf{G}^H(\omega)\mathbf{G}(\omega)\right]^{-1}\mathbf{q}(\omega). \tag{13}$$

The local noise-reduction factor is obtained by substituting (13) in (12). After some manipulations we obtain

$$\xi_{\mathrm{nr}}\left[\mathbf{h}_{\mathrm{LCMV}}(\omega)\right] = C(\omega)\left(\|\mathbf{d}(\omega)\|_2^2 - \frac{|\mathbf{d}^H(\omega)\mathbf{b}(\omega)|^2}{\|\mathbf{d}(\omega)\|_2^2}\right), \tag{14}$$

where

$$C(\omega) = \frac{1}{|D_1(\omega)|^2}\left(1 + \frac{\sigma_b^2(\omega)}{\sigma^2_{\mathrm{nc}}(\omega)}\,|B_1(\omega)|^2\right). \tag{15}$$

For the MVDR filter[1], $K = 1$, we have $\tilde{\mathbf{\Phi}}_v(\omega) = \sigma_b^2(\omega)\mathbf{b}(\omega)\mathbf{b}^H(\omega) + \sigma^2_{\mathrm{nc}}(\omega)\mathbf{I}$ and $\mathbf{q}^H(\omega) = Q(\omega) = D_1(\omega)$ so that [9]

$$\xi_{\mathrm{nr}}\left[\mathbf{h}_{\mathrm{MVDR}}(\omega)\right] = C(\omega)\left(\|\mathbf{d}(\omega)\|_2^2 - \frac{|\mathbf{d}^H(\omega)\mathbf{b}(\omega)|^2}{\frac{\sigma^2_{\mathrm{nc}}(\omega)}{\sigma_b^2(\omega)} + \|\mathbf{b}(\omega)\|_2^2}\right). \tag{16}$$

When $\|\mathbf{d}^H(\omega)\mathbf{b}(\omega)\|_2^2 = 0$, the local noise-reduction factors of the MVDR and LCMV filters are equal. For $\|\mathbf{d}^H(\omega)\mathbf{b}(\omega)\|_2^2 \neq 0$ and $\|\mathbf{b}(\omega)\|_2^2 \approx \|\mathbf{d}(\omega)\|_2^2$, the local noise-reduction factor depends on the ratio $\frac{\sigma^2_{\mathrm{nc}}(\omega)}{\sigma_b^2(\omega)}$. For $\sigma_b^2(\omega) \gg \sigma^2_{\mathrm{nc}}(\omega)$, the local noise-reduction factors of the MVDR and LCMV filters are similar. However, when $\frac{\sigma^2_{\mathrm{nc}}(\omega)}{\sigma_b^2(\omega)} + \|\mathbf{b}(\omega)\|_2^2 < \|\mathbf{d}(\omega)\|_2^2$, the local noise-reduction factor of the LCMV filter is larger than the local noise-reduction factor of the MVDR filter.

6. MODIFYING THE CONSTRAINT

In this section different forms of the constraint are investigated. Let us assume that we are interested in extracting the sum of the reverberant signals $\{X_{1,1}(\omega), \ldots, X_{K_{\mathrm{d}},1}(\omega)\}$, while the other K_{u} coherent source signals are considered interfering signals. The corresponding desired response vector is given by $\mathbf{q}(\omega) = [G^*_{1,1}(\omega)\,\ldots\,G^*_{K_{\mathrm{d}},1}(\omega)\,0\,\ldots\,0]^T$, where superscript * denotes complex conjugation. In this special case we can modify the K constraints given by $\mathbf{h}^H(\omega)\mathbf{G}(\omega) = \mathbf{q}^H(\omega)$ by dividing the kth constraint ($1 \leq k \leq K$) by $G_{k,1}(\omega)$. We then obtain a modified constraint

$$\mathbf{h}^H(\omega)\dot{\mathbf{G}}(\omega) = \dot{\mathbf{q}}^H(\omega), \tag{17}$$

where

$$\dot{\mathbf{G}}(\omega) = \begin{bmatrix} 1 & \cdots & 1 \\ \frac{G_{1,2}(\omega)}{G_{1,1}(\omega)} & \cdots & \frac{G_{K,2}(\omega)}{G_{K,1}(\omega)} \\ \vdots & \ddots & \vdots \\ \frac{G_{1,M}(\omega)}{G_{1,1}(\omega)} & \cdots & \frac{G_{K,M}(\omega)}{G_{K,1}(\omega)} \end{bmatrix} \tag{18}$$

and

$$\dot{\mathbf{q}}(\omega) = [\underbrace{1\,\ldots\,1}_{K_{\mathrm{d}}}\underbrace{0\,\ldots\,0}_{K_{\mathrm{u}}}]^T. \tag{19}$$

It should be noted that (18) consists only of RTFs. A major advantage of the RTFs is that they can be estimated non-blindly while the ATFs can only be estimated blindly. Furthermore, the original and modified constraint provide the exact same LCMV filter. Therefore, the local and global noise-reduction factors do not change.

Since the desired response of each source is either one or zero, we can further simplify (17) and reduce the dimensions of $\dot{\mathbf{G}}(\omega)$ to $M \times K'$ with $2 \leq K' < K$ rather than $M \times K$. The latter modification reduces the number of computations required to compute the LCMV filter. For $K' = 2$ we obtain the constraint matrix

$$\ddot{\mathbf{G}}(\omega) = \begin{bmatrix} 1 & 1 \\ \sum_{k=1}^{K_{\mathrm{d}}} \frac{G_{k,2}(\omega)}{G_{k,1}(\omega)} & \sum_{k=K_{\mathrm{d}}+1}^{K} \frac{G_{k,2}(\omega)}{G_{k,1}(\omega)} \\ \vdots & \vdots \\ \sum_{k=1}^{K_{\mathrm{d}}} \frac{G_{k,M}(\omega)}{G_{k,1}(\omega)} & \sum_{k=K_{\mathrm{d}}+1}^{K} \frac{G_{k,M}(\omega)}{G_{k,1}(\omega)} \end{bmatrix} \tag{20}$$

and $\ddot{\mathbf{q}}(\omega) = [1\,0]^T$.

The first and second column of $\ddot{\mathbf{G}}(\omega)$ can be constructed by combing RTFs related to the individual desired and undesired sources, respectively.

7. PERFORMANCE EVALUATION

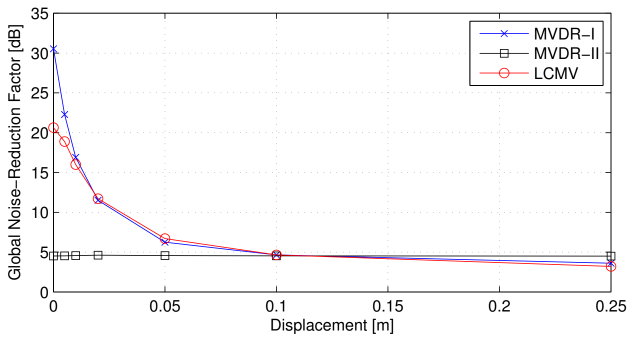
We now evaluate the noise-reduction performance of the LCMV and MVDR beamformers in an acoustic environment. We consider a scenario with one desired source and one undesired source in a homogenous and spatially white noise field. Specifically, we evaluate the global noise-reduction factor in case the position of the undesired noise source changes.

A linear microphone array was used with 4 microphones and an inter-microphone distance of 5 cm. The room size is $5 \times 4 \times 6$ m (length$\times$width$\times$height) with a reverberation time of 300 ms. All room impulse responses are generated using an efficient implementation of the image-method [10]. The desired source consists of speech colored noise. The undesired source consists of an AR(1) process (autoregressive process of order one) that was created by filtering a stationary zero-mean Gaussian sequences with a linear time-invariant filter. All signals were analyzed using discrete Fourier transforms of length 8192.

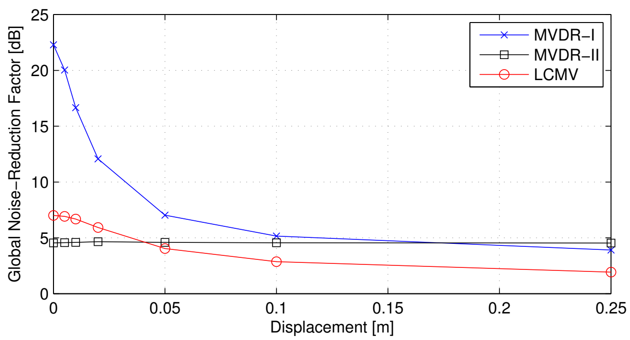
Three different filters are employed: the first filter is the LCMV beamformer as defined in (13) with $\dot{\mathbf{G}}(\omega)$ rather than $\mathbf{G}(\omega)$. The second filter, denoted by MVDR-I, is an MVDR beamformer with $\tilde{\mathbf{\Phi}}_v(\omega) = \sigma_b^2(\omega)\mathbf{b}(\omega)\mathbf{b}(\omega)^H + \sigma^2_{\mathrm{nc}}(\omega)\mathbf{I}$. The third filter, denoted by MVDR-II, is an MVDR beamformer with $\tilde{\mathbf{\Phi}}_v(\omega) = \sigma^2_{\mathrm{nc}}(\omega)\mathbf{I}$.

The constraints used to compute the LCMV and MVDR beamformers are based on the RTFs. Both beamformers require the RTFs related to the desired source. However, the LCMV beamformer requires the RTFs related to the undesired source while

[1]The MVDR filter is computed using the PSD matrix $\tilde{\mathbf{\Phi}}_v(\omega)$ of the coherent plus non-coherent noise field.



(a) $\mathrm{SNR}_{\mathrm{nc}} = 30$ dB and $\mathrm{SNR}_b = -10$ dB.



(b) $\mathrm{SNR}_{\mathrm{nc}} = 30$ dB and $\mathrm{SNR}_b = 5$ dB.

Figure 1: Global noise-reduction factor as a function the noise source displacement for two scenarios.

the MVDR-I requires the PSD matrix of the undesired source plus non-coherent noise. The MVDR-II beamformer only uses the RTFs related to the desired source and the PSD matrix of the non-coherent noise.

We define the original and displaced position of the undesired source as $\mathbf{p} = [x, y, z]$ and $\tilde{\mathbf{p}}$, respectively. The displacement of the source is given by $\Delta = ||\mathbf{p} - \tilde{\mathbf{p}}||_2$. In the following we assume that the displaced position is given by $\tilde{\mathbf{p}} = \mathbf{p} + \Delta\,[\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi)]$, where ϕ and θ determine the direction in which the source is displaced.

The noise-reduction factor depends on the direction and displacement of the noise source. Therefore, we consider for a fixed displacement Δ a spatially averaged global noise reduction that is obtained by averaging the global noise reduction across different directions. Specifically, for each displacement value Δ we chose 32 uniformly distributed points on a sphere centered at position $\mathbf{p}$ with a radius Δ. For each point the PSD matrix of the coherent noise was computed and subsequently, for each beamformer, the global noise reduction using (11). Finally, the 32 values were averaged to obtain the spatially averaged global noise reduction for Δ.

In Fig.1 the spatially averaged global noise-reduction factor is plotted as a function of the displacement $\Delta = \{0, 0.5, 1, 2, 5, 10, 25\}$ cm. The signal to noise ratios $\mathrm{SNR}_{\mathrm{nc}}$ and $\mathrm{SNR}_{\mathrm{b}}$ are defined for the first microphone as the ratio of the power of the desired signal over the power of the non-coherent and coherent noise, respectively. Similar to the analysis of the local noise reduction factor we see that $\xi_{\mathrm{nr}}(\mathbf{h}_{\mathrm{MVDR}}) \approx \xi_{\mathrm{nr}}(\mathbf{h}_{\mathrm{LCMV}})$ when $\sigma_b^2 \gg \sigma_{\mathrm{nc}}^2$ (Fig.1a). However, when the difference between σ_{nc}^2 and σ_b^2 becomes smaller the MVDR outperforms the LCMV beamformer in terms of the global noise suppression (Fig.1b). In both cases we can see that the global noise reduction is decreasing monotonically for increasing Δ. Interestingly we see that the MVDR-II outperforms the MVDR-I and LCMV beamformers for large values of Δ.

8. CONCLUSIONS

In this paper we described a general framework to analyze the noise reduction performance of various beamformer structures. A general noise-reduction factor was defined and used to analyze the performance of the LCMV beamformer. A performance evaluation was conducted in which we studied the spatially averaged global noise-reduction factor as a function of the displacement of the undesired coherent source. For the considered senario the MVDR beamformer achieved significantly better noise reduction compared to the LCMV beamformer. A more comprehensive study for different scenarios is a subject for ongoing research. In addition, it was shown that the constraint of the LCMV beamformer can be modified in such a way that only the RTFs rather than the ATFs are required.

9. REFERENCES

[1] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.

[2] S. Gannot and I. Cohen, "Adaptive beamforming and postfiltering," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds. Springer-Verlag, 2007, ch. 48.

[3] M. Er and A. Cantoni, "Derivative constraints for broadband element space antenna array processors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 6, pp. 1378–1393, Dec. 1983.

[4] O. L. Frost, III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[5] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propag.*, vol. 30, no. 1, pp. 27–34, Jan. 1982.

[6] K. M. Buckley, "Broad-band beamforming and the generalized sidelobe canceller," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 1322–1323, Oct. 1986.

[7] S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

[8] J. Benesty, J. Chen, Y. Huang, and J. Dmochowski, "On microphone-array beamforming from a MIMO acoustic signal processing perspective," *IEEE Trans. Speech Audio Process.*, vol. 15, no. 3, pp. 1053–1065, Mar. 2007.

[9] E. Habets, J. Benesty, I. Cohen, S. Gannot, and J. Dmochowski, "New insights into the MVDR beamformer in room acoustics," *accepted for publication in the IEEE Trans. Acoust., Speech, Signal Process.*, 2009.

[10] E. A. P. Habets. (2008, May) Room impulse response (RIR) generator. [Online]. Available: http://home.tiscali.nl/ehabets/rirgenerator.html