Dual MVDR Architecture for Adaptive Cancellation of Dynamic Interference

Amos Schreibman, Elior Hadad, Anna Barnov, Eli Tzirkel-Hancock

General Motors Technical Center Israel
Email: amos.schreibman@gm.com

***Abstract*—Microphone arrays are widely used for speech enhancement applications. We consider the enhancement of multiple, desired and undesired, sources in a noisy environment, the approximate locations of which are assumed to be known a priori. For the given scenario, the linear constraint minimum variance (LCMV) beamformer is commonly used. Although an LCMV-based beamformer provides the optimal solution for the case of static sources, the problem becomes more challenging when the sources' locations are constantly changing. The LCMV spatial notch pointing at the interference location is usually very sharp, resulting in cancellation degradation in the case of even small movements of the interfering sources. We propose an alternative scheme to the traditional LCMV beamformer that efficiently tracks and cancels the interfering source. The scheme is presented for a dual source scenario. We prove that for the static scenario, the proposed method and the LCMV beamformers are mathematically equivalent. However, for practical uses, we demonstrate that the proposed algorithm outperforms the LCMV in terms of signal-to-interference ratio (SIR) using a simulated room environment, as well as in real recorded data.**

***Index Terms*—beamforming, LCMV, MVDR, noise reduction, RTF, adaptive filter, AIC.**

I. Introduction

New applications that are arising in the fields of hearing aids, speech recognition, and cellular communication present new challenges for speech enhancement algorithms. As technology advances and the deployment of microphone arrays becomes common, multichannel processing methods, in particular beamforming techniques, are being utilized.

Commonly used beamforming designs optimize the minimum variance distortionless response (MVDR) criterion [1], which minimizes the noise power at the output while maintaining the target speech signal undistorted. In more complicated scenarios where several sources, i.e., competing speakers, exist, it is commonly desired to mitigate the interfering sources. The multiple constraints extension of the MVDR, known as the linear constraint minimum variance (LCMV) beamformer, is a suitable solution [2].

Although theoretically the LCMV provides the optimal solution, in practice the problem becomes more challenging when the interfering source location is constantly changing. Small movements of the undesired source result in a large degradation of the interference cancellation because of the sharp characteristic of the LCMV's spatial notch. This fact calls for rapid tracking of the interfering source location and frequent updates of the LCMV solution in order to maintain high cancellation levels. High cadence updates of the LCMV beamformer are computationally expensive because of the need to re-calculate the LCMV filter weights. Furthermore, in a generalized sidelobe canceler (GSC) based LCMV implementation [3][4][5][6], rapid updates of the LCMV design cause constant re-convergence of the adaptive noise canceler (ANC) block, which can lead to poor noise cancellation and degradation of the speech quality [7].

In this paper, we propose an alternative scheme to the traditional LCMV beamformer, which efficiently tracks and cancels the interfering source. A novel dual MVDR-based beamformer is derived, which (similarly to the LCMV beamformer) is aimed to extract a distortion-less target source, to cancel an interference source, and to minimize the background noise power. The derivation is achieved in two stages. First, a spatial processor applies two beamformers, which respectively are MVDR beamformers steered toward the target and interfering sources. The output signals of the processor are denoted as the target MVDR signal and the interference MVDR signal, respectively. Then, two adaptive interference canceller (AIC) systems are applied. The first AIC is aimed at cancelling the target signal from the interference MVDR signal, such that the output of the filter is the reference of the second AIC. Any target component is eliminated from this interference signal. The second AIC is aimed at cancelling this interference signal from the target MVDR signal, such that the filter forces the output signal to resemble a target signal without the interference component. We refer to this beamformer as the AIC dual MVDR (AIC-DMVDR) beamformer.

The analytical expression of the AIC-DMVDR beamformer filter is derived. We prove that for the static scenario, the proposed AIC-DMVDR and the LCMV beamformers are mathematically equivalent. However, for practical uses, we demonstrate that the AIC-DMVDR beamformer outperforms the LCMV in terms of signal-to-interference ratio (SIR) using a simulated room environment, as well as in real recorded data.

II. Configuration and notation

Consider an enclosure with 2 sources, a target source denoted s_0 and an interference source s_1. The sources signals are contaminated by additive noise comprising of any combination of coherent, diffuse and spatially white noise signals, impinging on a microphone array (MA) consisting of M microphones. The microphone signals are sampled at a sampling rate of f_s and transformed into the short time Fourier transform (STFT) using a window of length K with overlap η between frames. The transformed microphone signals are stacked into an M dimensional vector per time-frequency bin.

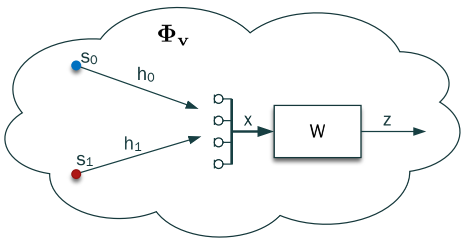


Fig. 1: Problem Description.

The signal $\mathbf{x}$ picked up by the MA can be described as:

$$\mathbf{x}(\ell, k) = \mathbf{h}_0(\ell, k)s_0(\ell, k) + \mathbf{h}_1(\ell, k)s_1(\ell, k) + \mathbf{v}(\ell, k) \quad (1)$$

where ℓ and k are time-frame and frequency bin indices respectively, $\mathbf{h}_i(\ell, k)$, $i = [0, 1]$, is a vector of the acoustic transfer functions (ATFs) between source s_i and the MA, and $\mathbf{v}(\ell, k)$ is the noise component vector with a spatial covariance matrix of:

$$\mathbf{\Phi_v}(\ell, k) = E\{\mathbf{v}(\ell, k)\mathbf{v}^H(\ell, k)\} \quad (2)$$

with operator $(\cdot)^H$ denoting conjugate-transpose, and $E\{\cdot\}$ denoting the expectation operator.

Given a-priori knowledge of $\mathbf{h}_i(\ell, k)$[1], the problem at hand is to design a beamformer $\mathbf{w}(\ell, k)$ such that the output signal

$$z(\ell, k) = \mathbf{w}^H(\ell, k)\mathbf{x}(\ell, k) \quad (3)$$

is enhanced by minimizing the output noise energy and cancelling the undesired source signal while maintaining distortion-less response towards the target source signal. The described scenario is depicted in Fig. 1.

In the following, time-frame and frequency bin indices are omitted for brevity.

III. Noise reduction technique

In Section III-A, the MVDR and LCMV beamformers are reviewed. Then, in Section III-B the proposed AIC-DMVDR beamformer is introduced.

A. MVDR and LCMV beamformers

The well-known MVDR beamformer is designed to reproduce the target source without distortion while minimizing the background noise power, i.e.,

$$\mathbf{w}_{\text{MVDR}} = \arg\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi_v} \mathbf{w} \quad \text{s.t.} \quad \mathbf{h}_0^H \mathbf{w} = 1. \quad (4)$$

The filter that solves the problem can be written as[2]

$$\mathbf{w}_{\text{MVDR}} = \frac{\mathbf{\Phi_v}^{-1}\mathbf{h}_0}{\mathbf{h}_0^H \mathbf{\Phi_v}^{-1}\mathbf{h}_0}. \quad (5)$$

To solve the problem addressed in this paper, the multi-constraints extension of the MVDR, known as the LCMV beamformer, is required [1]. The LCMV beamformer is designed to reproduce the target source component while cancelling the directional interference and minimizing the background noise power. The LCMV beamformer cost function can be written as

$$\mathbf{w}_{\text{LCMV}} = \arg\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi_v} \mathbf{w} \quad \text{s.t.} \quad \mathbf{C}^H \mathbf{w} = \mathbf{g}, \quad (6)$$

where $\mathbf{C}$ denotes the constraint matrix and $\mathbf{g}$ denotes the desired vector with

$$\mathbf{C} = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T. \quad (7)$$

Operator $(\cdot)^T$ denotes matrix transpose. The filter that solves the problem can be written as [1]

$$\mathbf{w}_{\text{LCMV}} = \mathbf{\Phi_v}^{-1}\mathbf{C}\left(\mathbf{C}^H \mathbf{\Phi_v}^{-1}\mathbf{C}\right)^{-1}\mathbf{g}. \quad (8)$$

B. Proposed beamformer

In the following, we present a novel scheme called AIC-DMVDR as an alternative to the LCMV solution.

Consider the configuration depicted in Fig. 2.

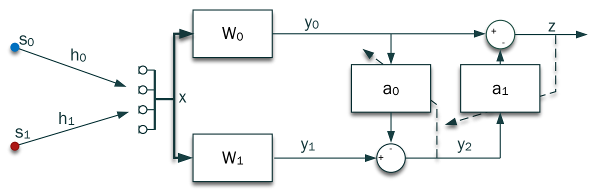


Fig. 2: Proposed solution.

The derivation of the beamformer is achieved in two stages. First, a spatial processor applies two beamformers, $\mathbf{w}_i$, $i = [0, 1]$, which respectively are MVDR beamformers steered toward the target and interfering sources. Using (5), the target MVDR signal y_0 and the interference MVDR signal y_1, can be described as

$$\begin{aligned} y_0 &= \mathbf{w}_0^H \mathbf{x} = s_0 + \frac{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0} s_1 + \frac{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}\mathbf{v} \\ y_1 &= \mathbf{w}_1^H \mathbf{x} = s_1 + \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1} s_0 + \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1}\mathbf{v}. \end{aligned} \quad (9)$$

The target MVDR signal y_0 is directly related to the target signal distortion induced at the AIC-DMVDR beamformer output. The closer $\mathbf{h}_0$ estimation is to the true ATF vector $\mathbf{h}_0$, the lower is the target signal distortion at the algorithm output. Robust beamformer design methods [9] can be employed when designing $\mathbf{w}_0$ in order to widen the beamformer's main lobe and reduce the target signal distortion.

In the second stage of the AIC-DMVDR derivation, two AIC systems are applied. The first AIC filter, a_0, is aimed at cancelling the target signal from the interference MVDR signal, thereby preventing distortion resulting from self cancellation of the target at the AIC-DMVDR beamformer output. The interference MVDR signal y_1 is the input signal to the adaptive filter, the reference of which is the target MVDR signal y_0. The output of the AIC system y_2 is the reference to the second AIC system. Any target component is eliminated from this interference signal. The second AIC filter, a_1, is aimed at cancelling this interference signal, y_2, from the target MVDR signal, y_0, such that the filter forces the AIC-DMVDR output

[1]The estimation of $\mathbf{h}_i(\ell, k)$ or alternatively its RTF representation, is widely researched [3], [8] and is not addressed in the remainder of the paper.

[2]Note that the MVDR beamformer steered toward the interfering source can be obtained by substituting $\mathbf{h}_0$ with $\mathbf{h}_1$ in (5).

signal z to resemble a target signal without the interference component.

1) Closed-form solution for static sources' location: Considering initially the static sources case, to achieve perfect cancellation of s_0, a_0 optimal weight is calculated according to the Wiener solution [10] for the noiseless case, i.e., $s_1 = 0, \mathbf{v} = 0$, and

$$a_0 = r_0^{-1} p_0 \tag{10}$$

where $r_0 = E\{|y_0|^2\}$, $p_0 = E\{y_0 y_1^*\}$, and operator $(\cdot)^*$ denotes complex conjugate. Using (9), the resulting a_0 optimum solution is given by

$$a_0 = \frac{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1} \mathbf{h}_1}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1} \mathbf{h}_1}. \tag{11}$$

The output of the first AIC system y_2 can be described as

$$y_2 = y_1 - a_0^* y_0 = \sin^2(\theta) s_1 + \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1} \mathbf{T}_0^H}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1} \mathbf{h}_1} \mathbf{v}, \tag{12}$$

where

$$\sin^2(\theta) = \left(1 - \frac{|\mathbf{h}_0^H \mathbf{\Phi}_V^{-1} \mathbf{h}_1|^2}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1} \mathbf{h}_0 \mathbf{h}_1^H \mathbf{\Phi}_V^{-1} \mathbf{h}_1}\right) \tag{13}$$

and

$$\mathbf{T}_0 = \left(\mathbf{I} - \frac{\mathbf{\Phi}_V^{-1} \mathbf{h}_0 \mathbf{h}_0^H}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1} \mathbf{h}_0}\right) \tag{14}$$

is the orthogonal projection matrix w.r.t. $\mathbf{h}_0$. Since y_2 does not include an instance of s_0, there cannot exist any signal distortion resulting from self cancellation due to the second AIC filter, a_1.

The second AIC filter, a_1, is similarly derived by substituting y_0 and y_1, with y_2 and y_0, respectively, such that (9) and (12) are used to evaluate $r_1 = E\{|y_2|^2\}$ and $p_1 = E\{y_2 y_0^*\}$ for the clean case ($s_0 = 0, \mathbf{v} = 0$). The optimal AIC filter, a_1, is given by

$$a_1 = r_1^{-1} p_1 = \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1} \mathbf{h}_0}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1} \mathbf{h}_0 \sin^2(\theta)}. \tag{15}$$

Finally, the output signal is given by

$$z = y_0 - a_1^* y_2. \tag{16}$$

By rearranging terms, using (11), (12), (15), and (16), the filter for the proposed AIC-DMVDR beamformer can be written as

$$\mathbf{w}_{\text{AIC-DMVDR}} = \mathbf{w}_0 - a_1 (\mathbf{w}_1 - a_0 \mathbf{w}_0). \tag{17}$$

2) Adaptive form solution for dynamic sources location: In the dynamic sources case, the a_0 and a_1 values cannot be derived explicitly from $\mathbf{h}_0$ and $\mathbf{h}_1$. For this task, system identification techniques are required. To evaluate a_0, we consider time segments in which s_1 is inactive. Using (9), (11), and (5), the interference MVDR signal y_1, can be expressed by

$$y_1 = a_0^* s_0 + \mathbf{w}_1^H \mathbf{v} = a_0^* y_0 + \tilde{v}_1, \tag{18}$$

where

$$\tilde{v}_1 = \left(\mathbf{w}_1^H - a_0^* \mathbf{w}_0^H\right) \mathbf{v}. \tag{19}$$

Assuming slowly changing ATFs and noise statistics, the cross power spectral density (PSD) between y_1 and y_0 for a given frame ℓ can be expressed as

$$S_{y_1 y_0}^{(\ell)} = a_0^* S_{y_0 y_0}^{(\ell)} + S_{\tilde{v}_1 y_0} \tag{20}$$

As $\mathbf{v}$ is assumed to be stationary and uncorrelated to s_0, $S_{\tilde{v}_1 y_0}$ is independent of the frame index ℓ [3]. If $\tilde{v}_1$ was uncorrelated with y_0, standard adaptive filter (AF) techniques could be used to solve for a_0. Unfortunately, as (19) and (9) state, the two variables are correlated, and standard AF techniques cannot be used. Alternatively, [11] suggests harnessing the non-stationary nature of speech signals to obtain an unbiased estimation of a_0. The following describes a recursive least squares (RLS) based adaptation of the solution given in [11].

The estimate of the auto and cross PSDs $\hat{S}_{y_0 y_0}^{(\ell)}$ and $\hat{S}_{y_1 y_0}^{(\ell)}$ are given by

$$\begin{aligned} \hat{S}_{y_1 y_0}^{(\ell)} &= \mu \hat{S}_{y_1 y_0}^{(\ell-1)} + (1-\mu) y_1(\ell) y_0^*(\ell) \\ \hat{S}_{y_0 y_0}^{(\ell)} &= \mu \hat{S}_{y_0 y_0}^{(\ell-1)} + (1-\mu) |y_0(\ell)|^2, \end{aligned} \tag{21}$$

where μ is the estimation smoothing factor. From (20), we obtain the basis for a least squares (LS) estimation of a_0

$$\hat{S}_{y_1 y_0}^{(\ell)} = \begin{bmatrix} a_0^* & S_{\tilde{v}_1 y_0} \end{bmatrix} \phi(\ell) + \epsilon(\ell), \tag{22}$$

where $\phi(\ell) = \begin{bmatrix} \hat{S}_{y_0 y_0}^{(\ell)} & 1 \end{bmatrix}^T$ and $\epsilon(\ell)$ is the estimation error.

Instead of using a set of over determined equations to solve for a_0 in (22) as suggested by [11], we suggest using an RLS approach. Defining the deterministic autocorrelation matrix w.r.t. $\phi(\ell)$ $\mathbf{A}(\ell)$, and the deterministic cross correlation vector $\mathbf{b}(\ell)$ as

$$\begin{aligned} \mathbf{A}(\ell) &= \lambda \mathbf{A}(\ell-1) + \phi(\ell) \phi^H(\ell) \\ \mathbf{b}(\ell) &= \lambda \mathbf{b}(\ell-1) + \phi(\ell) \hat{S}_{y_1 y_0}^{(\ell)}, \end{aligned} \tag{23}$$

where λ is the RLS forgetting factor. The estimation of a_0 is given by

$$\begin{bmatrix} \hat{a}_0^* \\ \hat{S}_{\tilde{v}_1 y_0} \end{bmatrix} = \mathbf{A}^{-1}(\ell) \mathbf{b}(\ell). \tag{24}$$

A similar method is used to obtain a_1. Consider time segments where s_0 is inactive. Using (9), (5), (15), and (12), y_0 can be expressed by

$$y_0 = a_1^* \sin^2(\theta) s_1 + \mathbf{w}_0^H \mathbf{v} = a_1^* y_2 + \tilde{v}_0, \tag{25}$$

where

$$\tilde{v}_0 = \left(\mathbf{w}_0^H - a_1^* \mathbf{w}_1^H \mathbf{T}_0^H\right) \mathbf{v}. \tag{26}$$

Observing the similarity between (25) and (18), the process given by equations (21) - (24) can be used to obtain a_1, by substituting $\hat{S}_{y_1 y_0}$, $\hat{S}_{y_0 y_0}$, y_1 and $\tilde{v}_1$ with $\hat{S}_{y_0 y_2}$, $\hat{S}_{y_2 y_2}$, y_0 and $\tilde{v}_0$.

IV. Relation between the AIC-DMVDR and the LCMV beamformers

In this section, we first show that the LCMV beamformer can be decomposed to two MVDR beamformers steered toward the target and interfering sources. Then, we show the relation between the proposed AIC-DMVDR beamformer and the LCMV beamformer.

The filter of the LCMV beamformer can be calculated by substituting (7), and (13) into (8). The filter can be written as (Section III.F in [12])

$$\mathbf{w}_{\text{LCMV}} = \frac{1}{\sin^2(\theta)} \frac{\mathbf{\Phi}_V^{-1}\mathbf{h}_0}{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0} - \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}{\sin^2(\theta)\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0} \frac{\mathbf{\Phi}_V^{-1}\mathbf{h}_1}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1}. \tag{27}$$

Clearly, the LCMV beamformer is a mix of a target and an interference MVDR beamformers, i.e.,

$$\mathbf{w}_{\text{LCMV}} = \frac{1}{\sin^2(\theta)}\mathbf{w}_0 - \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}{\sin^2(\theta)\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}\mathbf{w}_1. \tag{28}$$

Note that

$$\frac{1}{\sin^2(\theta)} = \frac{\sin^2(\theta)+\cos^2(\theta)}{\sin^2(\theta)} = 1 + \frac{\cos^2(\theta)}{\sin^2(\theta)}. \tag{29}$$

Rearranging the terms in (28) and using (29) and (13), the filter of the LCMV beamformer is given by

$$\mathbf{w}_{\text{LCMV}} = \mathbf{w}_0 - \frac{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}{\sin^2(\theta)\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_0}\left(\mathbf{w}_1 - \frac{\mathbf{h}_0^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1}{\mathbf{h}_1^H \mathbf{\Phi}_V^{-1}\mathbf{h}_1}\mathbf{w}_0\right). \tag{30}$$

Comparing (30) with (17) and using (11) and (15), it is shown that mathematically the LCMV beamformer is equivalent to the AIC-DMVDR beamformer. Theoretically, the proposed approach provides the same performance as a perfectly tuned LCMV beamformer. However, in practical uses, in a dynamic sources case as the a_0 and a_1 cancellation filters constantly track the sources location, the proposed AIC-DMVDR beamformer is expected to outperform the traditional LCMV beamformer.

V. Experimental Results

In this section, we present simulation results comparing the performance of the proposed beamformer and the traditional LCMV beamformer, for simulated data (Section V-A) and real-world recordings (Section V-B).

In both experiments, one target speaker and one interference speaker, contaminated by background noise, were used. The sampling frequency was 16 kHz. The signals were transformed to the STFT domain with 2048 points, 75% overlap, and a Hamming analysis window. The signals' relative transfer functions (RTFs) were estimated using the eigenvalue decomposition (EVD) method [13]. To decide on exclusive signal activity, source s_i was declared exclusive if the following directional based decision rule was fulfilled:

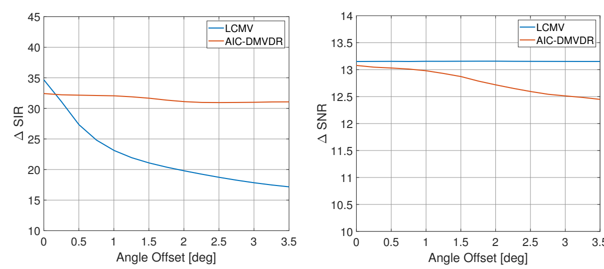
$$\frac{\tilde{\mathbf{h}}_i^H \mathbf{x}\mathbf{x}^H \tilde{\mathbf{h}}_i}{\text{Tr}\{\mathbf{x}\mathbf{x}^H\} - \tilde{\mathbf{h}}_i^H \mathbf{x}\mathbf{x}^H \tilde{\mathbf{h}}_i} > Th \frac{\tilde{\mathbf{h}}_j^H \mathbf{x}\mathbf{x}^H \tilde{\mathbf{h}}_j}{\text{Tr}\{\mathbf{x}\mathbf{x}^H\} - \tilde{\mathbf{h}}_j^H \mathbf{x}\mathbf{x}^H \tilde{\mathbf{h}}_j} \tag{31}$$

where $\text{Tr}\{\cdot\}$ denotes trace operator, $\tilde{\mathbf{h}}_i = \frac{\mathbf{h}_i}{\|\mathbf{h}_i\|}$, and $Th = 10$ db was used. For both experiments, we used a scenario with a SIR of 3 dB and signal-to-noise ratio (SNR) of 15 dB measured on the reference microphone. The proposed method AICs were estimated using the process described in Sec. III-B2. Speech input signals were taken from [14]. The two beamformers were evaluated using the SIR improvement and the SNR improvement.

A. Simulated Environment

For the simulated data, a conference room impulse responses (RIRs) were simulated with a RIR generator as described in [15]. The room dimensions were 3m × 4m × 3m with RT60 = 360 ms. A uniform line array consisting of 8 microphones and 1 cm interspacing was considered. Two active speakers were located at a distance of 0.85 m from the array center. The target source was located at 20^o and the interference source was initially located at 120^o. Diffuse noise was simulated using the noise generator described in [16]. To demonstrate the AIC-DMVDR beamformer's robustness to small movements of the interference source, the simulation was conducted several times. In each iteration, the interference source location was shifted by an additional 0.25^o without changing the beamformer's design.

The output SNR and SIR were measured for both algorithms. The results are summarized in Fig. 3. It is evident that while the two beamformers obtain comparable values of SNR improvements, regardless of the shift in the interference source location, the SIR improvement of the LCMV beamformer drops by around 10 dB for a small shift of 0.5^o. The proposed beamformer maintains the SIR levels throughout the experiment as it tracks the interference source movement. The difference in performance between the two algorithms for angle offset of 0^o is related to estimation errors of the proposed algorithm.



(a) SIR improvement   **(b)** SNR improvement

Fig. 3: SIR (a) and SNR (b) improvement, LCMV vs. proposed algorithm. Simulated environment.

B. Real-world recordings

For the second experiment, real-world recordings, using a database of RIRs obtained from the acoustic lab at Bar-Ilan University [17], were used. The room dimensions were 6m × 6m × 2.4m, and its acoustic properties could be controlled by opening and closing various panels mounted on the walls, ceiling, and floor, thereby changing their reflectivity. A room configuration with RT60 = 360 ms was evaluated. Using the recorded RIRs from [17], two alternately active speakers were located at a distance of 2 m from the array center. The target source was located at 45^o and the interference source at 135^o. To test the proposed algorithm's tracking capabilities, the interference source shifted its location to 120^o in the middle of the simulation, starting from around 20 s. The simulation noise field consisted of a directional source located at 75^o immersed in a white noise field.

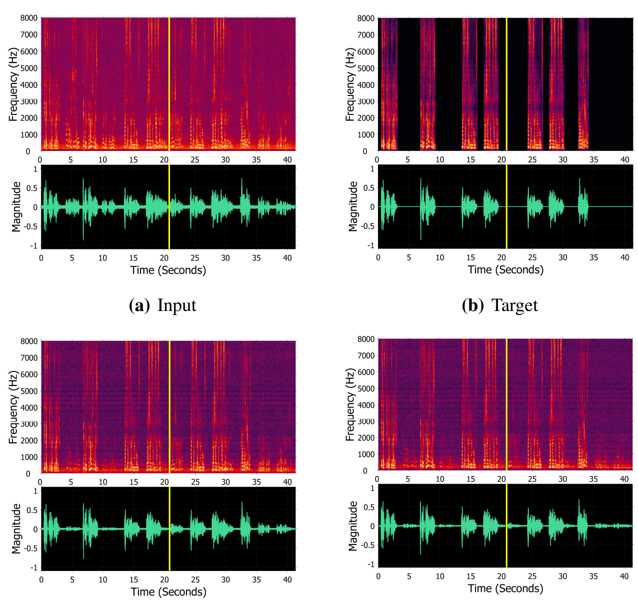
Fig. 4 depicts the sonograms of the input and the target source as sampled at the reference channel, as well as the outputs of both the LCMV and AIC-DMVDR beamformers. The time of the interference source location shift is highlighted by a yellow line.

It is evident that both the LCMV and the AIC-DMVDR beamformers significantly attenuate the stationary noise. However, from around 20 s in the simulation, as indicated by the yellow line in Fig. 4, while the interference cancellation of the proposed AIC beamformer is maintained, the cancellation of the interference source for the LCMV beamformer is degraded, corresponding to the source's movement.

In addition, the output SNR and SIR were measured for both algorithms. The results are summarized in Table I. It is evident that the two beamformers obtain comparable SNR and SIR values for the first source's position. However, while the SIR of the LCMV beamformer drops by around 10 dB for the interference source's second position, as the LCMV beamformer's spatial notch no longer points at its location, the proposed AIC-DMVDR beamformer maintains the SIR levels as it tracks the interference source's movement. The SNR improvement for both beamformers remains comparable for the second source's position.

TABLE I: Recorded environment: Comparison of LCMV and proposed algorithms

		SNR [dB]	SIR [dB]
Input		15.0	3.8
LCMV	Position 1	19.2	19.8
	Position 2	20.0	**9.6**
Proposed	Position 1	21.6	18.9
	Position 2	19.2	**17.8**



(a) Input **(b)** Target

(c) LCMV Output **(d)** AIC-DMVDR output

Fig. 4: Sonograms of the input signal (a), target signal (b), and the LCMV (c) and AIC-DMVDR (d) output signals.

VI. Summary

In this paper, we proposed a novel MVDR-based beamformer that is designed to extract the target source and efficiently track and cancel an interference source in a noisy and reverberant environment. The proposed beamformer was theoretically analyzed and proven to be mathematically equivalent to a perfectly tuned LCMV beamformer. The proposed method was tested in both synthetic simulation and real life scenarios and was shown to outperform the traditional LCMV approach for a dynamic interference source. Examining the robustness of the proposed method to estimation errors in dynamic scenarios is left for future work.

References

[1] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *ASSP Magazine, IEEE*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[2] S. Markovich, S. Gannot, and I. Cohen, "Multichannel eigenspace beamforming in a reverberant environment with multiple interfering speech signals," *IEEE Trans. Audio, Speech and Language Processing*, vol. 17, no. 6, pp. 1071–1086, Aug. 2009.

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

[4] A. Barnov, A. Gendelman, A. Schreibman, E. Tzirkel-Hancock, and S. Gannot, "A robust RLS implementation of the ANC block in GSC structures," in *2021 29th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 261–265.

[5] W. Herbordt and W. Kellermann, "Computationally efficient frequency-domain robust generalized sidelobe canceller," in *Proc. Int. Workshop on Acoustic Echo and Noise Control*, 2001, pp. 51–55.

[6] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters," *IEEE Transactions on signal processing*, vol. 47, no. 10, pp. 2677–2684, 1999.

[7] A. Barnov, V. B. Bracha, and S. Markovich-Golan, "QRD based MVDR beamforming for fast tracking of speech and noise dynamics," in *2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2017, pp. 369–373.

[8] I. Cohen, "Relative transfer function identification using speech signals," *IEEE Transactions on Speech and Audio Processing*, vol. 12, no. 5, pp. 451–459, 2004.

[9] C.-Y. Chen and P. P. Vaidyanathan, "Quadratically constrained beamforming robust against direction-of-arrival mismatch," *IEEE Transactions on signal processing*, vol. 55, no. 8, pp. 4139–4150, 2007.

[10] S. Haykin, *Adaptive filter theory*. Prentice-Hall, Inc., 1991.

[11] O. Shalvi and E. Weinstein, "System identification using nonstationary signals," *IEEE transactions on signal processing*, vol. 44, no. 8, pp. 2055–2063, 1996.

[12] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE Trans. Audio, Speech and Language Processing*, vol. 24, no. 3, pp. 543–558, March 2016.

[13] R. Varzandeh, M. Taseska, and E. A. Habets, "An iterative multichannel subspace-based covariance subtraction method for relative transfer function estimation," in *2017 Hands-free Speech Communications and Microphone Arrays (HSCMA)*. IEEE, 2017, pp. 11–15.

[14] ITU-T Recommendation, "ITU-T P.501, Test signals for use in telephony and other speech-based applications," 2017.

[15] E. A. Habets, "Room impulse response generator," *Technische Universiteit Eindhoven, Tech. Rep*, vol. 2, no. 2.4, p. 1, 2006.

[16] E. A. Habets, I. Cohen, and S. Gannot, "Generating nonstationary multisensor signals under a spatial coherence constraint," *The Journal of the Acoustical Society of America*, vol. 124, no. 5, pp. 2911–2917, 2008.

[17] E. Hadad, F. Heese, P. Vary, and S. Gannot, "Multichannel audio database in various acoustic environments," in *14th International Workshop on Acoustic Signal Enhancement (IWAENC), Aachen, Germany, Sep.*, 2014, pp. 313–317.