Dual Source Transfer-Function Generalized Sidelobe Canceller

Gal Reuven, Sharon Gannot and Israel Cohen

Abstract

Full duplex hands-free man/machine interface often suffers from directional nonstationary interference (such as a competing speaker or an echo signal) as well as a stationary interference (which may comprise both directional and non-directional signals). We propose a new structure for handling both interferences, based on the transfer-function generalized sidelobe canceller (TF-GSC). As in the standard GSC structure, it contains three blocks: a matched beamformer (MBF), a blocking matrix (BM) and a multi-channel adaptive noise canceller (ANC). The MBF is designed to block the competing speech signal, while maintaining the desired signal arriving from the direction of interest. The blocking matrix is designed to block both the desired and the competing speech signals. The multi-channel ANC is used to mitigate the remaining noise in the MBF output. The performance of the proposed algorithm is evaluated through a series of simulations, in reverberant environments. This structure is shown to be related to both the echo cancellation problem and to the convolutive blind source separation (BSS) problem.

I. Introduction

In many sound environments the received signals are contaminated by a nonstationary interference signal, as well as by a stationary noise. These problems might be encountered in noisy and reverberant sound enclosures in which both the desired speech signal and the competing speech signal are filtered by the room impulse response and contaminated by an additive stationary noise source before being received by the microphone array.

Two distinct approaches for treating this problem exist in the literature. In one family of solutions the scenario is treated as a convolutive blind source separation (BSS) problem. In the second family beamformer techniques are used.

The most common procedure in treating the convolutive BSS problem is to transform the received signals to the frequency domain. By doing so, the convulative mixture is translated to the simpler instantaneous mixture. However, the frequency domain solution necessitates further processing, as both permutation and gain ambiguity problems are encountered. Several approaches were taken to handle these problems. Speech signal nonstationarity is exploited by Parra and Spence [1] to obtain a nonlinear minimization problem. Both permutation and gain ambiguity problems are alleviated by imposing an FIR structure on the mixing filters. When the number of microphones is larger than the number of sources, Parra and Alvino [2] propose to circumvent the problem of permutations ambiguities by imposing geometrical constraints on the solution.
the two sources case Kurita et al. [3] and Ikram and Morgan [4] use the null location of the beam-pattern to mitigate the permutation problem. Knaak et al. [5] incorporate the geometrical constrains into the FastICA algorithm[6], based on higher order statistics (HOS).

In our contribution we take the second approach, namely the beamformer technique. Opposed to the BSS approach, which tries to separate both speech signals, beamforming treats one of the speech signals as the desired signal and the other signal as an interference.

Linely constrained minimum variance (LCMV) [7] beamformer has found numerous applications in the field of speech enhancement in the recent three decades. The most attractive implementation of the LCMV is the generalized sidelobe canceller (GSC) [8]. In this structure the constraint and the minimization are decoupled, yielding a simple but yet powerful tool for handling the problem. A comprehensive survey of beamforming methods, including the use of multiple constraints, can be found in [9]. In most speech enhancement applications the beamformer is constrained to produce a dominant response towards the assumed speech source location, while minimizing the response for all other directions. However, in reverberant environment a single direction of arrival cannot be determined and the desired signal impinge on the array from several directions due to the many reflections from objects in the room. This problem might be alleviated by using a complex acoustic transfer function (ATF) rather than just a simple delay for modelling the propagation of the speech signal in the reverberant room.

Affes and Grenier proposed in [10] a subspace method for estimating and tracking the ATFs, which relate the speech source and the microphones in a reverberant environment. In [11] they further proposed a GSC structure, for situations where two speech signals are active simultaneously (usually, denoted a double talk situation), encountered in the context of acoustic echo canceller (AEC). They presented a distortionless fixed beamformer constrained to cancel the echo, and a blocking matrix constrained to block both the desired signal and echo signal. The ATFs estimates are used to construct both the fixed beamformer and the blocking matrix.

The nonstationarity of the speech signal, contrasted with the stationarity of the noise signal, was exploited by Gannot et al. [12] for estimating the ATF ratio. This extension to the classical GSC structure, nicknamed the transfer function generalized sidelobe canceller (TF-GSC), was used for enhancing speech signal deteriorated by a single stationary interference signal in an arbitrary ATF enclosure. However, in many sound environments the interference sources might be a near-end competing speech or a far-end echo signal, rather than a nonstationary noise signal. In the presence of a nonstationary interference, the TF-GSC cannot distinguish between the desired signal and the interference signal and the proposed structure is rendered useless.

In this contribution we present a novel method, which is an extension of the method presented in [12], for dealing with two interference signals, one stationary and the second nonstationary. A preliminary, conference
version of this paper was presented in [13]. A closed-form optimal solution is first derived using constrained optimization techniques. A more efficient, recursive algorithm, is then developed to enable tracking of changes in the environment. We modify two of the conventional GSC components, namely the fixed beamformer and the blocking matrix, to allow the beamformer to suppress the nonstationary interference. The modified fixed beamformer, denoted MBF, is designed to block the interference signal while maintaining the desired speech signal. The modified BM blocks both the desired signal and the interference signal. As in the conventional GSC structure, the adaptive noise canceller uses the reference signals produced by the blocking matrix to cancel the residual noise at the MBF output. A novel method for updating the blocking matrix in double talk situations is proposed as well. The method is based on the nonstationarity of both the desired and interference speech signals. The discussion is supported by an experimental study using speech and noise signals drawn from databases and filtered by simulated impulse responses. The outcome consists of the assessment of sound sonograms, signal to noise and to interference ratio enhancement and informal subjective listening tests.

The structure of this work is as follows. In Sec. II we formulate the problem of a dual-source interference cancelling in a general acoustical transfer function (ATF) environment. The optimal solution is presented in Sec. III. Sec. IV describes the proposed algorithm. Sec. V deals with the ATF estimation procedure. Sec. VI demonstrates some experimental results in practical scenarios. The application of the method to the echo cancellation problem is addressed in Sec. VII. In Sec. VIII the work is summarized.

II. Problem formulation

Let us consider an array of sensors in a noisy and reverberant environment. We assume that the received signals include three components, a desired speech source, a directional nonstationary interference signal (e.g. competing speech) and a stationary noise signal, which can be either directional, non-directional or a combination thereof. Our goal is to reconstruct the desired speech signal from the received signals, allowing reverberation. The $m$-th microphone signal is given by

$$z_m(t) = a_m(t) * s_1(t) + b_m(t) * s_2(t) + n_m(t); \quad m = 1, \ldots, M$$

(1)

where $a_m(t)$ and $b_m(t)$ are the acoustical impulse responses between the desired speech source and the non-stationary interference source to the $m$-th microphone, respectively; $s_1(t)$ and $s_2(t)$ are the desired speech source and the nonstationary interference source, respectively. $n_m(t)$ is the (directional or nondirectional) stationary noise signal at the $m$-th microphone, and $*$ denotes convolution. The analysis frame duration $T$ is chosen such that the signal may be considered stationary over the analysis frame. Typically, the ATFs are changing slowly in time so that it may also be considered stationary over the analysis frame.
In the short time Fourier transform (STFT) domain, (1) can be approximately rewritten as:

\[
Z_m(t, e^{j\omega}) \approx A_m(e^{j\omega})S_1(t, e^{j\omega}) + B_m(e^{j\omega})S_2(t, e^{j\omega}) + N_m(t, e^{j\omega})
\]

\[m = 1, \ldots, M.\]  \(2\)

\[Z_m(t, e^{j\omega}), S_1(t, e^{j\omega}), S_2(t, e^{j\omega})\text{ and } N_m(t, e^{j\omega}\text{ are the STFT of the respective signals. } A_m(e^{j\omega}) \text{ and } B_m(e^{j\omega})\text{ are the ATFs from the desired source and interference source to the } m\text{-th microphone, respectively, assumed hereinafter to be time invariant over the observation period. The vector formulation of (2) is}
\]

\[
Z(t, e^{j\omega}) = A(e^{j\omega})S_1(t, e^{j\omega}) + B(e^{j\omega})S_2(t, e^{j\omega}) + N(t, e^{j\omega})
\]

\[3\]

where

\[
Z(t, e^{j\omega}) = \begin{bmatrix}
Z_1(t, e^{j\omega}) & Z_2(t, e^{j\omega}) & \cdots & Z_M(t, e^{j\omega})
\end{bmatrix}^T
\]

\[
A(e^{j\omega}) = \begin{bmatrix}
A_1(e^{j\omega}) & A_2(e^{j\omega}) & \cdots & A_M(e^{j\omega})
\end{bmatrix}^T
\]

\[
B(e^{j\omega}) = \begin{bmatrix}
B_1(e^{j\omega}) & B_2(e^{j\omega}) & \cdots & B_M(e^{j\omega})
\end{bmatrix}^T
\]

\[
N(t, e^{j\omega}) = \begin{bmatrix}
N_1(t, e^{j\omega}) & N_2(t, e^{j\omega}) & \cdots & N_M(t, e^{j\omega})
\end{bmatrix}^T .
\]

In the subsequent sections we will derive an optimal reconstruction for \(S_1(t, e^{j\omega})\) (or a filtered version thereof) given the noisy observations \(Z(t, e^{j\omega})\).

### III. Optimal solution based on constrained optimization

In this section, we derive a linearly constraint beamformer, specifically designed for suppressing undesired interference signals. We first obtain a closed-form linearly constrained minimum variance beamformer, and then derive an adaptive solution. We initially assume that the ATFs are known and in Section V we present system identification procedure, based on the nonstationarity of the speech signals.

Let \(W_m^*(t, e^{j\omega}) ; m = 1, \ldots, M\) be a set of \(M\) filters,

\[
W(t, e^{j\omega}) = \begin{bmatrix}
W_1^* (t, e^{j\omega}) & W_2^* (t, e^{j\omega}) & \cdots & W_M^* (t, e^{j\omega})
\end{bmatrix}
\]

where * denotes conjugation and \(^\dagger\) denotes conjugation transpose. A beamformer is realized by filtering each sensor output \(Z_m(t, e^{j\omega})\) by \(W_m^*(t, e^{j\omega}) ; m = 1, \ldots, M\) and summing the outputs:

\[
Y(t, e^{j\omega}) = W(t, e^{j\omega})Z(t, e^{j\omega}) = W(t, e^{j\omega})A(e^{j\omega})S_1(t, e^{j\omega})
\]

\[+ W(t, e^{j\omega})B(e^{j\omega})S_2(t, e^{j\omega}) + W(t, e^{j\omega})N(t, e^{j\omega}) \stackrel{\Delta}{=} Y_{s1}(t, e^{j\omega}) + Y_{s2}(t, e^{j\omega}) + Y_{n}(t, e^{j\omega}) \]

where \(Y_{s1}(t, e^{j\omega})\) is the desired signal part, \(Y_{s2}(t, e^{j\omega})\) is the directional interference part and \(Y_{n}(t, e^{j\omega})\) is the stationary noise part. The output power is given by:

\[
E\{Y(t, e^{j\omega})Y^*(t, e^{j\omega})\} = E\{W(t, e^{j\omega})Z(t, e^{j\omega})Z^*(t, e^{j\omega})W(t, e^{j\omega})\} = W(t, e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega})
\]
where $\Phi_{ZZ}(t, e^{j\omega}) = E\{Z(t, e^{j\omega})Z(t, e^{j\omega})\}$. We want to minimize the output power subject to the following constraints:

$$
Y_{s1}(t, e^{j\omega}) = W^\dagger(t, e^{j\omega})A(e^{j\omega})S_1(t, e^{j\omega}) = F^*(e^{j\omega})S_1(t, e^{j\omega})
$$

$$
Y_{s2}(t, e^{j\omega}) = W^\dagger(t, e^{j\omega})B(e^{j\omega})S_2(t, e^{j\omega}) = 0
$$

where $F^*(e^{j\omega})$ is an arbitrary filter response. This filter may absorb any filtering operation imposed on the output due to reverberation. We thus have the following minimization problem

$$
\min_W \left\{ W^\dagger(t, e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right\}
$$

subject to $W^\dagger(t, e^{j\omega})A(e^{j\omega}) = F^*(t, e^{j\omega})$ and $W^\dagger(t, e^{j\omega})B(e^{j\omega}) = 0.

The minimization depicted in (6) is demonstrated in Fig. 1, using a two dimensional slice modelling three sensors system. The noise space is assumed to be perpendicular to the sheet (i.e. pointing towards the reader). The tangent point of the equi-power contours with the constraint line is the optimum vector of beamformer filters. The perpendicular $F(e^{j\omega})$ from the origin to the constraint line is calculated in Section III.

Solution to the problem in (6), is obtained by minimizing the complex Lagrangian:

$$
\mathcal{L}(W) = W^\dagger(t, e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) + \lambda_1 \left[ W^\dagger(t, e^{j\omega})A(e^{j\omega}) - F^*(e^{j\omega}) \right] + \lambda_2 \left[ A^\dagger(e^{j\omega})W(t, e^{j\omega}) - F(e^{j\omega}) \right] + \lambda_2 B^\dagger(e^{j\omega})W(t, e^{j\omega})
$$

Setting the derivative with respect to $W^*$ to 0 (see for instance [14]) we obtain

$$
\nabla_W \cdot \mathcal{L}(W) = \Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) + \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) = 0
$$
and since $\Phi_{Z}(t, e^{j\omega})$ is usually invertible\(^1\), $W(t, e^{j\omega})$ can be written as:

$$W(t, e^{j\omega}) = -\Phi_{Z}^{-1}(t, e^{j\omega}) \left[ \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) \right].$$

(9)

Imposing the constraints on $W(t, e^{j\omega})$ and solving for the Lagrange multipliers yields (see Appendix -A)

$$W^{opt}(t, e^{j\omega}) = \mathcal{F}(e^{j\omega}) \Phi_{Z}^{-1}(t, e^{j\omega}) \frac{A(e^{j\omega})}{\|A(e^{j\omega})\|_{\Phi}} - \rho(e^{j\omega}) \frac{B(e^{j\omega})}{\|B(e^{j\omega})\|_{\Phi}} - |\rho(e^{j\omega})|^2$$

(10)

where

$$\|X(e^{j\omega})\|^2_{\Phi} \triangleq X^{\dagger}(e^{j\omega}) \Phi_{Z}^{-1}(t, e^{j\omega}) X(e^{j\omega})$$

(11)

denotes a weighted norm of a vector $X(e^{j\omega})$ and

$$\rho_{\Phi}(e^{j\omega}) \triangleq \frac{B^{\dagger}(e^{j\omega}) \Phi_{Z}^{-1}(t, e^{j\omega}) A(e^{j\omega})}{\sqrt{A^{\dagger}(e^{j\omega}) \Phi_{Z}^{-1}(t, e^{j\omega}) A(e^{j\omega}) \sqrt{B^{\dagger}(e^{j\omega}) \Phi_{Z}^{-1}(t, e^{j\omega}) B(e^{j\omega})}}}$$

(12)

is the cosine of the angle between the vectors $A(e^{j\omega})$ and $B(e^{j\omega})$ in a weighted inner product space.

The closed-form solution for the minimization problem, $W^{opt}(t, e^{j\omega})$, lacks the ability to track changes in the environment and is difficult to implement. Hence we replace the closed-form solution with an adaptive one.

Consider the following steepest descent, recursive algorithm, for minimizing the complex Lagrangian in (8):

$$W(t + 1, e^{j\omega}) = W(t, e^{j\omega}) - \mu \nabla_{W} \mathcal{L}(e^{j\omega}) = W(t, e^{j\omega}) - \mu \left[ \Phi_{ZZ}(t, e^{j\omega}) W(t, e^{j\omega}) + \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) \right].$$

(13)

Imposing the constraints on $W(t + 1, e^{j\omega})$ yields (see Appendix -B):  

$$W(t + 1, e^{j\omega}) = P(e^{j\omega}) W(t, e^{j\omega}) - \mu P(e^{j\omega}) \Phi_{ZZ}(t, e^{j\omega}) W(t, e^{j\omega}) + F(e^{j\omega})$$

(14)

where

$$P(e^{j\omega}) = I - \frac{\|B(e^{j\omega})\|^2 A(e^{j\omega}) A^{\dagger}(e^{j\omega}) - A^{\dagger}(e^{j\omega}) B(e^{j\omega}) B^{\dagger}(e^{j\omega}) - B(e^{j\omega}) B^{\dagger}(e^{j\omega}) A(e^{j\omega}) A^{\dagger}(e^{j\omega}) + \|A(e^{j\omega})\|^2 B(e^{j\omega}) B^{\dagger}(e^{j\omega})}{\|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2 - A^{\dagger}(e^{j\omega}) B(e^{j\omega}) B^{\dagger}(e^{j\omega}) A(e^{j\omega})}$$

$$F(e^{j\omega}) = \frac{\|B(e^{j\omega})\|^2 A(e^{j\omega}) A^{\dagger}(e^{j\omega}) - A^{\dagger}(e^{j\omega}) B(e^{j\omega}) B^{\dagger}(e^{j\omega}) A(e^{j\omega})}{\|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2 - A^{\dagger}(e^{j\omega}) B(e^{j\omega}) B^{\dagger}(e^{j\omega}) A(e^{j\omega})} \mathcal{F}(e^{j\omega}).$$

(15)

This forms the constrained recursive structure. Now, defining $\rho(e^{j\omega})$ as the coherence function (or the cosine of the angle between the vectors $A(e^{j\omega})$ and $B(e^{j\omega})$ in an inner product space).

$$\rho(e^{j\omega}) \equiv \frac{B^{\dagger}(e^{j\omega}) A(e^{j\omega})}{\|A(e^{j\omega})\| \|B(e^{j\omega})\|}$$

(16)

\(^1\)\text{As a small amount of uncorrelated sensor noise always exists, the invertibility of } \Phi_{Z}(t, e^{j\omega}) \text{ might be guaranteed in practical scenarios.}
we obtain

\[
P(e^{j\omega}) = I - \frac{A(e^{j\omega})A^\dagger(e^{j\omega})}{\|A(e^{j\omega})\|^2} - \rho^*(e^{j\omega}) \frac{A(e^{j\omega})B^\dagger(e^{j\omega})}{\|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2} - \rho(e^{j\omega}) \frac{B(e^{j\omega})A^\dagger(e^{j\omega})}{\|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2} + \frac{B(e^{j\omega})B^\dagger(e^{j\omega})}{\|B(e^{j\omega})\|^2}
\]

(17)

\[
F(e^{j\omega}) = \frac{A(e^{j\omega})}{\|A(e^{j\omega})\|^2} - \rho(e^{j\omega}) \frac{B(e^{j\omega})}{\|B(e^{j\omega})\|^2} \mathcal{F}(e^{j\omega}).
\]

The meaning of the coherence function will be discussed in the next section.

IV. THE DUAL SOURCE TFGSC

Following Gannot et al. [12] footsteps, we now derive an unconstrained adaptive enhancement algorithm. The unconstrained algorithm is usually advantageous due to its superior computational efficiency and the ability to use the well behaved NLMS scheme.

A. Generalized Sidelobe Canceller Interpretation

Consider the null space of \([A(e^{j\omega}) \mid B(e^{j\omega})]\), defined by

\[
\mathcal{N}(e^{j\omega}) \triangleq \left\{ W \mid \left[ A(e^{j\omega}) \mid B(e^{j\omega}) \right]^\dagger W(e^{j\omega}) = \begin{bmatrix} 0 & 0 \end{bmatrix} \right\}.
\]

Define the constraint hyperplane,

\[
\Lambda(e^{j\omega}) \triangleq \left\{ W \mid \left[ A(e^{j\omega}) \mid B(e^{j\omega}) \right]^\dagger W(e^{j\omega}) = \mathcal{F}(e^{j\omega}) \right\}
\]

which is parallel to \(\mathcal{N}(e^{j\omega})\). Furthermore, define the column space of \([A(e^{j\omega}) \mid B(e^{j\omega})]\) by

\[
\mathcal{R}(e^{j\omega}) \triangleq \{ \kappa_1 A(e^{j\omega}) + \kappa_2 B(e^{j\omega}) \mid \text{for any real } \kappa_1 \text{ and } \kappa_2 \}.
\]

Using the fundamental theorem of Linear Algebra [15], \(\mathcal{R}(e^{j\omega}) \perp \mathcal{N}(e^{j\omega})\). \(\kappa_1\) and \(\kappa_2\) can be easily identified in (17) to prove that \(F(e^{j\omega}) \in \mathcal{R}(e^{j\omega})\) and therefore is perpendicular to \(\mathcal{N}(e^{j\omega})\). Furthermore,

\[
A^\dagger(e^{j\omega})F(e^{j\omega}) = A^\dagger(e^{j\omega}) - \frac{\rho(e^{j\omega}) B(e^{j\omega})}{\|A(e^{j\omega})\| \|B(e^{j\omega})\|} \mathcal{F}(e^{j\omega}) = \mathcal{F}(e^{j\omega})
\]

and

\[
B^\dagger(e^{j\omega})F(e^{j\omega}) = B^\dagger(e^{j\omega}) - \frac{\rho(e^{j\omega}) A(e^{j\omega})}{\|A(e^{j\omega})\| \|B(e^{j\omega})\|} \mathcal{F}(e^{j\omega}) = 0.
\]

Hence, \(F(e^{j\omega}) \in \Lambda(e^{j\omega})\). Now, since \(F(e^{j\omega}) \perp \mathcal{N}(e^{j\omega})\) and \(\mathcal{N}(e^{j\omega})\) is parallel to \(\Lambda(e^{j\omega})\), \(F(e^{j\omega}) \perp \Lambda(e^{j\omega})\). This implies that \(F(e^{j\omega})\) is the perpendicular from the origin to the constraint hyperplane, \(\Lambda(e^{j\omega})\). The matrix \(P(e^{j\omega})\) is the projection matrix to the null space of \([A(e^{j\omega}) \mid B(e^{j\omega})]\), \(\mathcal{N}(e^{j\omega})\). This is easily shown by the following arguments. Using (17) we have,

\[
P(e^{j\omega})A^\dagger(e^{j\omega}) = A^\dagger(e^{j\omega}) - \frac{1}{1 - |\rho(e^{j\omega})|^2} \left[ \frac{A^\dagger(e^{j\omega})A(e^{j\omega})A^\dagger(e^{j\omega})}{\|A(e^{j\omega})\|^2} - \frac{A^\dagger(e^{j\omega})B(e^{j\omega})B^\dagger(e^{j\omega})}{\|B(e^{j\omega})\|^2} - \frac{B(e^{j\omega})A^\dagger(e^{j\omega})A^\dagger(e^{j\omega})}{\|A(e^{j\omega})\|^2} + \frac{B(e^{j\omega})B(e^{j\omega})B^\dagger(e^{j\omega})}{\|B(e^{j\omega})\|^2} \right].
\]

(18)
The term in brackets is equal to
\[
\mathbf{A}^\dagger(e^{j\omega}) - \frac{\|\mathbf{A}(e^{j\omega})\|^2}{\|\mathbf{A}(e^{j\omega})\|^2 \|\mathbf{B}(e^{j\omega})\|^2} \mathbf{A}^\dagger(e^{j\omega})\mathbf{B}(e^{j\omega}) - \frac{\|\mathbf{A}(e^{j\omega})\|^2}{\|\mathbf{A}(e^{j\omega})\|^2 \|\mathbf{B}(e^{j\omega})\|^2} \mathbf{A}^\dagger(e^{j\omega})\mathbf{B}(e^{j\omega})\mathbf{A}^\dagger(e^{j\omega}) - \mathbf{A}^\dagger(e^{j\omega})\mathbf{B}(e^{j\omega})\mathbf{B}^\dagger(e^{j\omega})
\]
and therefore \(P(e^{j\omega})\mathbf{A}^\dagger(e^{j\omega}) = 0\). In a similar manner, \(P(e^{j\omega})\mathbf{B}^\dagger(e^{j\omega}) = 0\) as well.

Now, a vector in linear space can be uniquely split into a sum of two vectors in mutually orthogonal subspaces (see for instance [15]). Hence,
\[
\mathbf{W}(t, e^{j\omega}) = \mathbf{W}_0(t, e^{j\omega}) - \mathbf{V}(t, e^{j\omega})
\]
where \(\mathbf{W}_0(t, e^{j\omega}) \in \mathcal{R}(e^{j\omega})\) and \(-\mathbf{V}(t, e^{j\omega}) \in \mathcal{N}(e^{j\omega})\). By the definition of \(\mathcal{N}(e^{j\omega})\),
\[
\mathbf{V}(t, e^{j\omega}) = \mathcal{H}(e^{j\omega})\mathbf{G}(t, e^{j\omega})
\]
where \(\mathcal{H}(e^{j\omega})\) is a full-rank \(M \times (M - 2)\) matrix, such that the columns of \(\mathcal{H}(e^{j\omega})\) span the null space of \([\mathbf{A}(e^{j\omega}) \mid \mathbf{B}(e^{j\omega})]\), i.e.
\[
\mathbf{A}^\dagger(e^{j\omega})\mathcal{H}(e^{j\omega}) = 0,
\mathbf{B}^\dagger(e^{j\omega})\mathcal{H}(e^{j\omega}) = 0.
\]
The vector \(\mathbf{G}(t, e^{j\omega})\) is an \((M - 2) \times 1\) vector of adjustable filters.

Using the geometrical interpretation of Frost’s algorithm (see Fig. 1),
\[
\mathbf{W}_0(t, e^{j\omega}) = \mathbf{F}(e^{j\omega}) = \frac{\mathbf{A}(e^{j\omega}) - \rho(e^{j\omega})\mathbf{B}(e^{j\omega})}{1 - \|\rho\|^2}\mathcal{F}(e^{j\omega})
\]
(Recall that \(\mathbf{F}(e^{j\omega})\) is the perpendicular from the origin to the constraint hyperplane, \(\mathcal{H}(e^{j\omega})\)). Now, using (4), (20) and (21) we get
\[
Y(t, e^{j\omega}) = Y_{\text{MBF}}(t, e^{j\omega}) - Y_{\text{NC}}(t, e^{j\omega})
\]
where
\[
Y_{\text{MBF}}(t, e^{j\omega}) = \mathbf{W}_0^\dagger(t, e^{j\omega})\mathbf{Z}(t, e^{j\omega}),
Y_{\text{NC}}(t, e^{j\omega}) = \mathbf{G}^\dagger(t, e^{j\omega})\mathcal{H}^\dagger(e^{j\omega})\mathbf{Z}(t, e^{j\omega}).
\]
The solution structure is similar to [12], although the constraints are different. The output of the constrained beamformer is a difference of two terms, both operating on the input signal \(\mathbf{Z}(t, e^{j\omega})\). \(\mathbf{W}_0(t, e^{j\omega})\) in our
problem, steers the beam towards the desired direction, while blocking the interference direction. In [12], \( W_0(t, e^{j\omega}) \) is only responsible for steering the beam towards the desired direction. Furthermore, \( \mathcal{H}(e^{j\omega}) \) in the current contribution blocks both directions while in [12] it only blocks the desired direction. \( G(e^{j\omega}) \) in both cases has similar functionality. However, its rank here is lower, allowing less degrees of freedom.

The first term, \( Y_{MBF}(t, e^{j\omega}) \), is dependent on the ATFs, hence it can be regarded as a matched beamformer (MBF). We now examine the second term, \( Y_{NC}(t, e^{j\omega}) \). Define \( U(t, e^{j\omega}) \), the reference noise signals:

\[
U(t, e^{j\omega}) = H^\dagger(e^{j\omega})Z(t, e^{j\omega}) = H^\dagger(e^{j\omega})[A(e^{j\omega})S_1(t, e^{j\omega}) + B(e^{j\omega})S_2(t, e^{j\omega}) + N(t, e^{j\omega})] = H^\dagger(e^{j\omega})N(t, e^{j\omega}).
\]

The last transition is due to (22). Both desired and competing signals’ components are blocked by \( H^\dagger(e^{j\omega}) \) and therefore \( U(t, e^{j\omega}) \) contains only noise. Hence, the noise term of \( Y_{MBF}(t, e^{j\omega}) \) can be reduced by properly adjusting the filters \( G(t, e^{j\omega}) \), using the minimum output power criterion. This adjustment problem is in fact the classical multi-channel noise cancellation problem, that can be solved by using the Wiener filter. An adaptive LMS solution to the problem was proposed by Widrow [16].

Recall that the coherence function \( \rho(e^{j\omega}) \) defined in (16) is the cosine of the angle between \( A(e^{j\omega}) \) and \( B(e^{j\omega}) \). When these vectors are perpendicular \( \rho(e^{j\omega}) \) vanishes. In this case, the resulting \( F(e^{j\omega}) \) is exactly the single source MBF derived in [12]. The projection matrix in this case becomes \( P(e^{j\omega}) = I - \frac{A(e^{j\omega})A^\dagger(e^{j\omega}) - B(e^{j\omega})B^\dagger(e^{j\omega})}{\|B(e^{j\omega})\|^2} \), depicting the orthogonality of both directions.

B. Detailed Structure

The solution comprises three building blocks. The first is an MBF, which satisfies the requested constraints, i.e. the desired signal direction is kept undistorted while the nonstationary interference signal direction is blocked. The second is a BM, that produces noise-only reference signals by blocking both the desired signal and nonstationary interference signal. The third is an unconstrained LMS-type algorithm, that cancels the coherent noise in the MBF output.

B.1 Blocking Matrix

The blocking matrix should be designed to block both the desired and interference signals, and to yield noise-only components at its outputs. We propose to construct \( \mathcal{H}(e^{j\omega}) \) as a cascade of two blocking matrices, \( \mathcal{H}(e^{j\omega}) = \mathcal{H}_1(e^{j\omega})\mathcal{H}_2(e^{j\omega}) \). \( \mathcal{H}_1(e^{j\omega}) \) is simply designed to block signals arriving from the desired signal direction, while \( \mathcal{H}_2(e^{j\omega}) \) has to block the signals arriving from the interfering direction, after being rotated.
by the first matrix. As in [12], \( \mathcal{H}_1(e^{j\omega}) \) is defined by
\[
\mathcal{H}_1(e^{j\omega}) = \begin{bmatrix}
-\frac{A_2^*(e^{j\omega})}{A_1^*(e^{j\omega})} & -\frac{A_3^*(e^{j\omega})}{A_1^*(e^{j\omega})} & \cdots & -\frac{A_M^*(e^{j\omega})}{A_1^*(e^{j\omega})} \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}.
\] (27)

Regarding \( \mathcal{H}_2(e^{j\omega}) \), note that
\[
B^\dagger(e^{j\omega}) \mathcal{H}(e^{j\omega}) = B^\dagger(e^{j\omega}) \left( \mathcal{H}_1(e^{j\omega}) \mathcal{H}_2(e^{j\omega}) \right) = \left( B^\dagger(e^{j\omega}) \mathcal{H}_1(e^{j\omega}) \right) \mathcal{H}_2(e^{j\omega}).
\] (28)

Thus,
\[
B^\dagger(e^{j\omega}) \mathcal{H}_1(e^{j\omega}) = \begin{bmatrix}
B_1^*(e^{j\omega}) & B_2^*(e^{j\omega}) & \cdots & B_M^*(e^{j\omega})
\end{bmatrix} \begin{bmatrix}
-\frac{A_2^*(e^{j\omega})}{A_1^*(e^{j\omega})} & -\frac{A_3^*(e^{j\omega})}{A_1^*(e^{j\omega})} & \cdots & -\frac{A_M^*(e^{j\omega})}{A_1^*(e^{j\omega})} \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix} = \begin{bmatrix}
-\frac{B_1^*(e^{j\omega}) A_2^*}{A_1^*} + B_2^2 & -\frac{B_1^*(e^{j\omega}) A_3^*}{A_1^*} + B_3^* & \cdots & -\frac{B_1^*(e^{j\omega}) A_M^*}{A_1^*} + B_M^* 
\end{bmatrix}.
\]

This vector, multiplied by \( \mathcal{H}_2(e^{j\omega}) \), should yield an all-zero vector. Consider \( \mathcal{H}_2(e^{j\omega}) \) of the type
\[
\mathcal{H}_2(e^{j\omega}) = \begin{bmatrix}
L_3(e^{j\omega}) & L_4(e^{j\omega}) & \cdots & L_M(e^{j\omega}) \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}.
\] (29)

The following linear equation determines \( L_m(e^{j\omega}) \): \( m = 3, \ldots, M \)
\[
\left[ -\frac{B_1^*(e^{j\omega}) A_2^*}{A_1^*} + B_2^2 \right] L_m(e^{j\omega}) + \left[ -\frac{B_1^*(e^{j\omega}) A_m^*}{A_1^*} + B_m^* \right] = 0.
\] (30)

Solving (30) we obtain:
\[
L_m(e^{j\omega}) = -\frac{A_2^*(e^{j\omega})}{A_1^*(e^{j\omega})} - \frac{B_1^*(e^{j\omega})}{B_1^*(e^{j\omega})}; \quad m = 3, \ldots, M.
\] (31)
Multiplying $\mathcal{H}_1(e^{j\omega})$ by $\mathcal{H}_2(e^{j\omega})$ and rearranging terms yields

$$\mathcal{H}(e^{j\omega}) = \begin{bmatrix} Q_3(e^{j\omega}) & Q_4(e^{j\omega}) & \cdots & Q_M(e^{j\omega}) \\ L_3(e^{j\omega}) & L_4(e^{j\omega}) & \cdots & L_M(e^{j\omega}) \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

(32)

where

$$Q_m(e^{j\omega}) = -\frac{A_m^*(e^{j\omega})B_m^*(e^{j\omega})}{A_m^*(e^{j\omega})B_m^*(e^{j\omega}) - B_m^*(e^{j\omega})} - \frac{B_m^*(e^{j\omega})A_m^*(e^{j\omega})}{A_m^*(e^{j\omega})B_m^*(e^{j\omega}) - B_m^*(e^{j\omega})}; \ m = 3, \ldots, M.$$  

(33)

We will now verify that $\mathcal{H}(e^{j\omega})$ satisfies both constraints, as described in (22):

$$A^1(e^{j\omega})\mathcal{H}(e^{j\omega}) = [A_1^*(e^{j\omega})Q_3(e^{j\omega}) + A_2^*(e^{j\omega})L_3(e^{j\omega}) + A_3^*(e^{j\omega})]; \ A_1^*(e^{j\omega})Q_4(e^{j\omega}) + A_2^*(e^{j\omega})L_4(e^{j\omega}) + A_3^*(e^{j\omega}); \ \cdots; A_1^*(e^{j\omega})Q_M(e^{j\omega}) + A_2^*(e^{j\omega})L_M(e^{j\omega}) + A_M^*(e^{j\omega})].$$

(34)

Calculating the $m-$th element in the last term

$$A_1^*(e^{j\omega})Q_m(e^{j\omega}) + A_2^*(e^{j\omega})L_m(e^{j\omega}) + A_m^*(e^{j\omega}) =$$

$$\frac{1}{A_1^*(e^{j\omega})B_1^*(e^{j\omega}) - B_1^*(e^{j\omega})} \left\{-A_2^*(e^{j\omega})B_m^*(e^{j\omega})/B_1^*(e^{j\omega}) + A_m^*(e^{j\omega})B_m^*(e^{j\omega})/B_1^*(e^{j\omega}) - A_2^*(e^{j\omega})A_m^*(e^{j\omega})/A_1^*(e^{j\omega}) + A_2^*(e^{j\omega})B_m^*(e^{j\omega})/B_1^*(e^{j\omega}) + A_m^*(e^{j\omega})A_m^*(e^{j\omega})/A_1^*(e^{j\omega}) - A_m^*(e^{j\omega})B_m^*(e^{j\omega})/B_1^*(e^{j\omega}) \right\} = 0$$

we obtain the required solution. Similarly, $\mathcal{H}(e^{j\omega})$ satisfies the second constraint, $B^1(e^{j\omega})\mathcal{H}(e^{j\omega}) = 0$ and therefore is a valid blocking matrix which is suitable for generating the reference noise signals. By (26), we have

$$U_m(t, e^{j\omega}) = Q_m(e^{j\omega})Z_1(t, e^{j\omega}) + L_m(e^{j\omega})Z_2(t, e^{j\omega}) + Z_m(t, e^{j\omega})$$

$$m = 3, \ldots, M.$$  

(36)

Thus, the knowledge of either $\frac{A_m(e^{j\omega})}{A_1(e^{j\omega})}$ and $\frac{B_m(e^{j\omega})}{B_1(e^{j\omega})}$ or directly $Q_m(e^{j\omega})$ and $L_m(e^{j\omega})$ is sufficient for creating the noise reference signals.

B.2 Matched beamformer

Recalling (17), the MBF is given by:

$$W_0(e^{j\omega}) = F(e^{j\omega}) = \frac{A(e^{j\omega})}{\|A(e^{j\omega})\|^2} - \rho(e^{j\omega}) \frac{B(e^{j\omega})}{\|A(e^{j\omega})\|\|B(e^{j\omega})\|^2} F(e^{j\omega}).$$

(37)
It was verified in Sec. IV that the constraints are satisfied by $W_0(e^{j\omega})$.

Calculating $W_0(e^{j\omega})$ using $\frac{A(e^{j\omega})}{A_1(e^{j\omega})}$ and $\frac{B(e^{j\omega})}{B_1(e^{j\omega})}$ instead of $A(e^{j\omega})$ and $B(e^{j\omega})$, respectively, we obtain:

$$
A_1(e^{j\omega})W_0(e^{j\omega}) = A_1(e^{j\omega})\frac{\|B(e^{j\omega})\|^2 A(e^{j\omega}) - B(e^{j\omega}) B_1(e^{j\omega}) A(e^{j\omega})}{\|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2 A_1(e^{j\omega}) - B_1(e^{j\omega}) B_1(e^{j\omega}) A_1(e^{j\omega})} F(e^{j\omega})
$$

$$
= \frac{1}{\|B_1(e^{j\omega})\|^2 A_1(e^{j\omega})} \left[ \|B(e^{j\omega})\|^2 \|A(e^{j\omega})\|^2 - A_1(e^{j\omega}) B(e^{j\omega}) B_1(e^{j\omega}) A(e^{j\omega}) \right] F(e^{j\omega})
$$

$$
= A_1^*(e^{j\omega}) F(e^{j\omega}).
$$

Accordingly, the desired signal is only distorted by the first ATF $A_1^*(e^{j\omega})$, which can be absorbed into $F(e^{j\omega})$. In a similar way, it can be shown that the nonstationary interference is completely blocked.

It should be noticed that although the proposed MBF fulfills the constraint $A_1(e^{j\omega})W_0 = F(e^{j\omega})$, the maximum directivity can be obtained at directions other than the desired signal direction. Therefore interfering signals may be emphasized, especially in low frequencies, which deteriorates the ability of the ANC to cancel noise signals. Consider the next example, as depicted in Fig. 2. Polar plots of directivity patterns are computed for 5 and 10 microphones arrays for several frequencies in a simple delay-only ATFs system. The desired source is assumed to arrive from 90° direction while the nonstationary interference from 100° direction. It is clear that the MBF satisfies both constraints in all plots, namely the gain is 1 and 0 in the desired direction and interference direction, respectively.

The $M = 10$ microphones MBF outperforms the $M = 5$ microphones MBF in all tested frequencies, since the array is more steerable. For example, at 500Hz, using 5 microphones maximum gain of 3.3 is achieved at 45° direction, while using 10 microphones maximum gain of 1.8 is achieved at 75° direction. Furthermore, the MBF has better performance for higher frequencies. Using $M = 10$ microphones, for example, maximum gain of 1.8 is achieved at 75° direction for 500Hz, while the maximum gain of 1 is achieved at 90° for 2000Hz.

B.3 Multi channel noise canceller

Recall that our goal is to minimize the output power under constraints on the response at the desired signal direction and at the competing signal direction. By setting $W_0(t, e^{j\omega})$ according to (23), the constraints are satisfied. Hence, minimization of the output power is achieved by adjusting the filters $G(t, e^{j\omega})$. This is an unconstrained minimization, exactly as in Widrow’s classical problem [16]. It can be implemented by using the multi-channel Wiener filter. Recalling (24), our goal is to set $G(t, e^{j\omega})$ to minimize

$$
E \left\{ \|Y_{MBF}(t, e^{j\omega}) - G^t(t, e^{j\omega})U(t, e^{j\omega})\|^2 \right\}.
$$
Fig. 2. MBF directivity patterns for several scenarios: $M=5$, (a) $f=500\text{Hz}$ (c) $f=1000\text{Hz}$ (e) $f=1500\text{Hz}$ (g) $f=2000\text{Hz}$; $M=10$, (b) $f=500\text{Hz}$ (d) $f=1000\text{Hz}$ (f) $f=1500\text{Hz}$ (h) $f=2000\text{Hz}$. 
Let
\[
\Phi_{UY}(t, e^{j\omega}) = E\{U(t, e^{j\omega})Y_{\text{MBF}}^*(t, e^{j\omega})\}
\]
\[
\Phi_{UU}(t, e^{j\omega}) = E\{U(t, e^{j\omega})U^\dagger(t, e^{j\omega})\}.
\]

Then the multi-channel Wiener filter is given by [17], [18]
\[
G(t, e^{j\omega}) = \Phi_{UU}^{-1}(t, e^{j\omega})\Phi_{UY}(t, e^{j\omega}).
\] (39)

In order to be able to track changes, the signals are processed by segments. The following frequency domain LMS algorithm is used. Let the residual signal be
\[
Y(t, e^{j\omega}) = Y_{\text{MBF}}(t, e^{j\omega}) - G(t, e^{j\omega})^\dagger U(t, e^{j\omega}).
\]

Note that the residual signal is also the output of the enhancement algorithm. Using the orthogonality principle, the error is orthogonal to the measurements. Thus,
\[
E\{U(t, e^{j\omega})Y^*(t, e^{j\omega})\} = 0.
\] (40)

Following the standard Widrow procedure, the solution is given by:
\[
G(t + 1, e^{j\omega}) = G(t, e^{j\omega}) + \mu U(t, e^{j\omega})Y^*(t, e^{j\omega}).
\]

Usually, a more stable solution is obtained by using the normalized LMS (NLMS) algorithm, in which each frequency is normalized separately, yielding:
\[
G_m(t + 1, e^{j\omega}) = G_m(t, e^{j\omega}) + \frac{\mu U_m(t, e^{j\omega})Y^*(t, e^{j\omega})}{P_{\text{est}}(t, e^{j\omega})}, \quad m = 2, \ldots, M
\]
where
\[
P_{\text{est}}(t, e^{j\omega}) = \eta P_{\text{est}}(t - 1, e^{j\omega}) + (1 - \eta) \sum_m |Z_m(t, e^{j\omega})|^2
\] (41)
where \(\eta\) is a forgetting factor (typically \(0.8 < \eta < 1\))^2.

The filter update is now given by
\[
\tilde{G}_m(t + 1, e^{j\omega}) = G_m(t, e^{j\omega}) + \frac{U_m(t, e^{j\omega})Y^*(t, e^{j\omega})}{P_{\text{est}}(t, e^{j\omega})}
\]
\[
G_m(t + 1, e^{j\omega}) \xleftarrow{\text{FIR}} \tilde{G}_m(t + 1, e^{j\omega})
\] (42)
for \(m = 3, \ldots, M\). The operator \(\xleftarrow{\text{FIR}}\) includes the following three stages. First, \(\tilde{G}_m(t + 1, e^{j\omega})\) is transformed to the time domain. Second, the resulting impulse response is truncated, namely an FIR constraint is imposed. Third, the result is transformed back to the frequency domain. Performing the \(\xleftarrow{\text{FIR}}\) operator avoids cyclic convolution.
V. ATFs Estimation

The ATFs ratios $\frac{A(e^{j\omega})}{A_1(e^{j\omega})}$ and $\frac{B(e^{j\omega})}{B_1(e^{j\omega})}$ are required for calculating the MBF and the blocking matrix. Till this point, the ATFs were assumed to be known. However, in practice, it should be estimated. We assume that the ATFs ratios are slowly changing in time compared to the time variations of the desired signal and the competing speech signal. We also assume that the statistics of the noise signal is slowly changing compared with the statistics of both the desired signal and the competing speech signal.

A. MBF Estimate

Estimation of the MBF is done in two steps. First, the ATFs ratios $\frac{A(e^{j\omega})}{A_1(e^{j\omega})}$ and $\frac{B(e^{j\omega})}{B_1(e^{j\omega})}$ are estimated separately, using the system identification procedure described in [12]. In the second step, $W_0(t, e^{j\omega})$ is estimated

Another possibility is to calculate $P_{est}$ using the $|U_m(t, e^{j\omega})|^2$ instead of $|Z_m(t, e^{j\omega})|^2$. However, in that case an energy detector is required, so that $G(t, e^{j\omega})$ is updated only when there is no active signal. If on the other hand, we calculate $P_{est}$ using the input sensor signals, $Z_m(t, e^{j\omega})$, as indicated in (41), then an energy detector may be avoided. This is due to the fact that the adaptation term becomes relatively small during periods of active input signal.
1) Fixed beamformer:
\[ Y_{\text{MBF}}(t, e^{j\omega}) = W_0^\dagger(e^{j\omega})Z(t, e^{j\omega}) \]

2) Noise reference signals:
\[ U(t, e^{j\omega}) = H^\dagger(e^{j\omega})Z(t, e^{j\omega}) \]

3) Output signal:
\[ Y(t, e^{j\omega}) = Y_{\text{MBF}}(t, e^{j\omega}) - G^\dagger(t, e^{j\omega})U(t, e^{j\omega}) \]

4) Filters update, for \( m = 1, \ldots, M - 1 \):
\[
\begin{align*}
\tilde{G}_m(t+1, e^{j\omega}) &= G_m(t, e^{j\omega}) + \mu \frac{U_m(t, e^{j\omega})Y^*(t, e^{j\omega})}{P_{\text{est}}(t, e^{j\omega})} \\
G_m(t+1, e^{j\omega}) &= \text{FIR} \tilde{G}_m(t+1, e^{j\omega})
\end{align*}
\]
where, \( P_{\text{est}}(t, e^{j\omega}) = \rho P_{\text{est}}(t-1, e^{j\omega}) + (1 - \rho) \sum_m |Z_m(t, e^{j\omega})|^2 \)

5) keep only non-aliased samples.

(note: \( W_0(e^{j\omega}) \) is defined in (23). \( H(e^{j\omega}) \) is defined in (32)).

Fig. 4. Summary of the DTF-GSC algorithm.

using (15), where the ATFs ratios are used instead of the real ATFs. Since the system identification algorithm is designed for estimating a single system at a time, the two ratios cannot be estimated simultaneously. Therefore only frames in which both signals are not simultaneously active are used.

We will now briefly describe the system identification algorithm. The observation period is divided into frames such that the desired or the competing speech signals may be considered stationary during each \( k \)-th frame. Define \( H_m(e^{j\omega}) \triangleq \frac{A_m(e^{j\omega})}{A_0(e^{j\omega})} \). The estimates are obtained by replacing expectations with averages. An unbiased estimate of \( H_m \) is obtained by applying the least squares criterion to the following set of overdetermined equations

\[
\begin{bmatrix}
\hat{\Phi}_{z_{m1}}(e^{j\omega}) \\
\hat{\Phi}_{z_{m2}}(e^{j\omega}) \\
\vdots \\
\hat{\Phi}_{z_{mK}}(e^{j\omega})
\end{bmatrix} = \begin{bmatrix}
\hat{\Phi}_{z_{11}}(e^{j\omega}) & 1 \\
\hat{\Phi}_{z_{21}}(e^{j\omega}) & 1 \\
\vdots & \vdots \\
\hat{\Phi}_{z_{K1}}(e^{j\omega}) & 1
\end{bmatrix} \begin{bmatrix}
H_m(e^{j\omega}) \\
\varepsilon_{m1}(e^{j\omega}) \\
\varepsilon_{m2}(e^{j\omega}) \\
\vdots \\
\varepsilon_{mK}(e^{j\omega})
\end{bmatrix} + \begin{bmatrix}
\varepsilon_{m1}(e^{j\omega}) \\
\varepsilon_{m2}(e^{j\omega}) \\
\vdots \\
\varepsilon_{mK}(e^{j\omega})
\end{bmatrix}
\]

(a separate set of equations is used for \( m = 2, \ldots, M \)), where \( \hat{\Phi}_{z_{11}}(e^{j\omega}), \hat{\Phi}_{z_{m1}}(e^{j\omega}) \) and \( \hat{\Phi}_{u_{m1}}(e^{j\omega}) \) are estimates of \( \Phi_{z_{11}}(e^{j\omega}), \Phi_{z_{m1}}(e^{j\omega}) \) and \( \Phi_{u_{m1}}(e^{j\omega}) \), respectively, and \( U(t, e^{j\omega}) \) are the noise reference signals as defined in [12]. \( \frac{B(e^{j\omega})}{B_0(e^{j\omega})} \) is estimated in a similar manner. Note, that, although estimated in distinct periods, it is still assumed that the two ratios estimates are valid.

As shown earlier, by using the ATFs ratios rather than the real ATFs in (17), the desired signal component in \( Y_{\text{MBF}}(t, e^{j\omega}) \) is distorted by \( A_1(e^{j\omega}) \), namely \( W_0^\dagger(e^{j\omega})A(e^{j\omega}) = A_1(e^{j\omega})F^*(e^{j\omega}) \).
B. Blocking Matrix Estimate

In Sec. V-A we presented a method for estimating the ATFs ratios. Inspecting (31) and (33), we note that the filters $Q_m(e^{j\omega})$ and $L_m(e^{j\omega})$ can be estimated by using these estimates, in a similar manner to $W_0(e^{j\omega})$ estimation.

However, for the blocking matrix, a direct estimate of $Q_m(e^{j\omega})$ and $L_m(e^{j\omega})$ can be obtained. We propose now a direct estimation method, which is applicable for double talk situations. This novel method permits adaptation of the BM components whenever the desired speech is active, regardless of the activity pattern of the interfering speech.

Choose observation periods in which both the desired and competing speech signals are active simultaneously. Rearranging terms in (36) yields

$$Z_m(t, e^{j\omega}) = -Q_m(e^{j\omega})Z_1(t, e^{j\omega}) - L_m(e^{j\omega})Z_2(t, e^{j\omega}) + U_m(t, e^{j\omega}).$$  \hspace{1cm} \text{(44)}

An analysis interval, in which $A(e^{j\omega})$, $B(e^{j\omega})$ are assumed to be time-invariant and the noise signal is assumed to be stationary is chosen. It is then divided into frames such that the desired and directional interference signals may be considered stationary during each $k$–th frame. Using (44) a system identification procedure can be obtained,

$$\hat{\Phi}^{(k)}_{z_m z_1}(e^{j\omega}) = -Q_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_1 z_1}(e^{j\omega}) - L_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_2 z_1}(e^{j\omega}) + \Phi_{u_m z_1}(e^{j\omega}); \ k = 1, \ldots, K$$  \hspace{1cm} \text{(45)}

where $K$ is the number of frames in the interval, and $\Phi^{(k)}_{z_i z_j}(e^{j\omega})$ is the cross power spectral density (cross-PSD) between $z_i$ and $z_j$ during the $k$–th frame. $\Phi_{u_m z_1}(e^{j\omega})$ is the cross-PSD between $u_m$ and $z_1$. Recalling that $H(e^{j\omega})$ blocks both the desired and interference directions and using (3) and (36) we obtain:

$$U_m(t, e^{j\omega}) = Q_m(e^{j\omega})N_1(t, e^{j\omega}) + L_m(e^{j\omega})N_2(t, e^{j\omega}) + N_m(t, e^{j\omega}).$$  \hspace{1cm} \text{(46)}

Since $N_m(t, e^{j\omega})$, $m = 1, \ldots, M$ are assumed stationary over the analysis interval and since $S_1(t, e^{j\omega})$ and $S_2(t, e^{j\omega})$ are independent of $N_m(e^{j\omega})$, it is clear that $\Phi_{u_m z_1}(e^{j\omega})$ is independent of the frame index $k$.

By replacing in (45) real PSD values with their estimates, calculated using time-averages, and by defining $\varepsilon^{(k)}_m(e^{j\omega}) = \hat{\Phi}^{(k)}_{u_m z_1}(e^{j\omega}) - \Phi_{u_m z_1}(e^{j\omega})$ as the estimation error of the cross-PSD between $z_1$ and $u_m$ in the $k$–th frame, the following set of equations are obtained:

$$\hat{\Phi}^{(k)}_{z_m z_1}(e^{j\omega}) = \hat{\Phi}_{u_m z_1}(e^{j\omega}) - Q_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_1 z_1}(e^{j\omega}) - L_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_2 z_1}(e^{j\omega}) = \Phi_{u_m z_1}(e^{j\omega}) + \varepsilon^{(k)}_m(e^{j\omega}) - Q_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_1 z_1}(e^{j\omega}) - L_m(e^{j\omega})\hat{\Phi}^{(k)}_{z_2 z_1}(e^{j\omega}); \ k = 1, \ldots, K.$$  \hspace{1cm} \text{(47)}

An unbiased estimate of $Q_m(e^{j\omega})$ and $L_m(e^{j\omega})$ is obtained by solving the following over-determined set of
equations, using the LS criterion:

\[
\begin{bmatrix}
\hat{\Phi}^{(1)}_{z_m z_1}(e^{j\omega}) \\
\hat{\Phi}^{(2)}_{z_m z_1}(e^{j\omega}) \\
\vdots \\
\hat{\Phi}^{(K)}_{z_m z_1}(e^{j\omega})
\end{bmatrix}
= 
\begin{bmatrix}
\hat{\Phi}^{(1)}_{z_1 z_1}(e^{j\omega}) & \hat{\Phi}^{(1)}_{z_2 z_1}(e^{j\omega}) & 1 \\
\hat{\Phi}^{(2)}_{z_1 z_1}(e^{j\omega}) & \hat{\Phi}^{(2)}_{z_2 z_1}(e^{j\omega}) & 1 \\
\vdots & \vdots & \vdots \\
\hat{\Phi}^{(K)}_{z_1 z_1}(e^{j\omega}) & \hat{\Phi}^{(K)}_{z_2 z_1}(e^{j\omega}) & 1
\end{bmatrix}
\times 
\begin{bmatrix}
-Q_m(e^{j\omega}) \\
-L_m(e^{j\omega}) \\
\Phi_{u_m z_1}(e^{j\omega}) \\
\vdots \\
\varepsilon^{(1)}_m(e^{j\omega}) \\
\varepsilon^{(2)}_m(e^{j\omega}) \\
\vdots \\
\varepsilon^{(K)}_m(e^{j\omega})
\end{bmatrix}
\]  

(a separate set of equations is used for \( m = 3, \ldots, M \)).

Note, that two schemes for estimating the blocking matrix were presented. During periods in which only the desired signal is active, we can update the estimate of \( A(e^{j\omega}) \) and \( B(e^{j\omega}) \), and use it in conjunction with the relevant estimate of \( \frac{B(e^{j\omega})}{A(e^{j\omega})} \) to obtain the blocking matrix estimate. In double talk frames we can continue the BM update using the novel scheme to estimate \( Q_m(e^{j\omega}) \) and \( L_m(e^{j\omega}) \) directly.

When desired speech is inactive, it is not necessary to update neither the BM nor the MBF, since the interference and noise signals can be completely eliminated from the output of the algorithm. This procedure necessitate the use of a voice activity detector (VAD). Although applicable, we did not use any VAD decisions in the experimental study presented in Sec. VI.

VI. Experimental results

The proposed algorithm was evaluated using the following test scenario. The desired and competing speech signals were drawn from the TIMIT [19] database, while a simulated speech-like noise, drawn from NOISEX-92 [20] database, was used as a stationary noise source. All three signals were filtered by simulated room impulse responses, resulting in directional signals, which are received by \( M = 10 \) microphones. Allen and Berkley’s image method [21] was used to simulate the ATFs with reverberation time set to \( T_{60} = 40\text{ms} \) (see Fig. 5 for a typical impulse and frequency responses of the acoustical path). In another test, a simulated diffused noise field signals were used as the stationary noise. The diffused noise source was generated by simulating an omnidirectional emittance of a white noise signal [22].

Segments of 1024 samples were used to implement the overlap and save procedure. The sampling frequency is 8KHz. The desired speech signal and the competing signal had the same level, with average SNR of 5 dB. For the directional noise field scenario, the length of the filters in the MBF, the blocking matrix, and the interference cancellers are set to 250, 250 and 500 taps, respectively. 250, 150 and 200 taps were used for the diffused noise field scenario.

Figure 6 shows the waveforms for the directional and diffused noise fields. In (a) and (b) a segment of the desired and competing speech signals are illustrated, respectively. Double talk situation is clearly observed. The signal measured at microphone \# 1 (comprises the desired speech, the competing speech and a directional noise signal) is depicted in (c). The respective enhanced signal, after the algorithm has adapted, is depicted in the (d). The noisy measurement at microphone \# 1 and the enhanced signal for the diffused noise case
(for the same time interval) are depicted in (e) and (f), respectively.

The SNR improvement for different stages of the GSC structure during the tested time interval, is given in Table I, for both the directional and the diffused noise fields. $S_{1NR}$ and $S_{2NR}$ represents the signal to noise ratios for signals $s_1(t)$ and $s_2(t)$, respectively. $S_1S_2R$ represents the power of $s_1(t)$ to the power of $s_2(t)$ ratio. The results for the input signal at the first microphone, the outputs of the MBF and BM blocks, and total algorithm output are depicted in the Table. It is clearly seen that the MBF significantly improves the desired signal to interference ratio for both noise fields, as well as achieving minor improvement in the desired signal to noise ratio. As expected, the BM deteriorates $S_{1NR}$ and $S_{2NR}$ for both noise fields, thus reducing the amount of desired signal leakage and allowing for the ANC to use almost noise-only reference signals.

It is clearly seen that the noise level is significantly reduced for the directional noise field case. The large amount of noise reduction is due to the use of a directional noise source. The desired signal to noise level during double talk situation is increased by 23.3 dB. The desired signal to competing signal ratio level is increased by 10.4 dB. The noise reduction performance for the diffused noise field case is clearly inferior compared to the directional noise field case. The desired signal to noise level during double talk situation is improved by only 8.2 dB. We will elaborate on this issue in the sequel. Note however, that in this particular case, the desired speech to the competing speech power ratio is increased by 20.3 dB, compared to only 10.4 dB in the directional noise field. This phenomenon is mainly contributed to the MBF performance in the particular microphone and sources constellation and is not necessarily a diffused noise field attribute.

Figure 7 shows sonograms of the data depicted in Figure 6. It can be seen that for the directional noise field, both noise and interference signals are well suppressed, especially in frequencies above 500 Hz. Moreover, no self cancellation or other distortion can be noticed during the double talk situation. Low noise reduction performance is obtained in the case of diffused noise field, while the nonstationary interference is well suppressed. The poor noise reduction performance in high frequencies (above 2500Hz) is due to the low coherence between
Fig. 6. Speech waveforms: (a) Desired signal, (b) Nonstationary interference; Directional noise field: (c) Mic. #1 signal, (d) Enhanced signal; Diffused noise field: (e) Mic. #1 signal, (f) Enhanced signal.
the noise components in the received signals.

Informal listening tests confirm that the perceptual quality of the desired speech signal (for the directional noise field case) is retained in the enhanced signal, while the stationary and nonstationary interferences are well suppressed. For audio sample files see also [23].

<table>
<thead>
<tr>
<th></th>
<th>Input</th>
<th>Output of MBF</th>
<th>Output of BM</th>
<th>Output of DTFGSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1NR$</td>
<td>$S_1S_2R$</td>
<td>$S_1NR$</td>
<td>$S_1S_2R$</td>
<td>$S_1NR$</td>
</tr>
<tr>
<td>11.3</td>
<td>2.3</td>
<td>13.8</td>
<td>16.9</td>
<td>-3.9</td>
</tr>
<tr>
<td>12.7</td>
<td>2.3</td>
<td>17.4</td>
<td>25</td>
<td>-3.8</td>
</tr>
</tbody>
</table>

**TABLE I**

**Noise and interference reduction in directional (top) and diffused (bottom) noise fields**

VII. Application to echo cancellation

Our problem is closely related to the echo cancellation problem as well. In these problems a joint effort of mitigating the echo signal and reducing the noise level is required. However, the two tasks generally contradict each other [24], especially in double talk situation. However, it should be stressed that in the echo cancellation problem a separate measurement of the interference signal is available and can be used to improve the performance of the overall system.

In [24] two frequency domain schemes for joint echo cancellation and noise reduction are presented. Both contain the TF-GSC beamformer proposed in [12] and a block least mean square (LMS) acoustic echo canceller (AEC). Following Kellermann [25], the first scheme comprises multi-channel AEC followed by a beamformer, while the second comprises a beamformer followed by a single channel AEC as a post-filter. A series of simulations using real speech recordings showed that the first scheme outperforms the second one. Two additional schemes for integrated joint noise reduction and echo cancellation are proposed by Doclo and Moonen [26] and Rombouts and Moonen [27]. In [26] the far-end echo signal is incorporated into a generalized singular value decomposition (GSVD) beamformer. In [27] a more efficient, QRD-based least squares lattice (QRD-LSL) algorithm is used instead of the GSVD algorithm. It is shown that the performance of the integrated scheme is superior to the performance of traditional (cascading) schemes, while complexity is kept at an affordable level.

In our contribution, the echo reference signal can be incorporated as an additional reference signal to the ANC. It should be noticed that in our scheme the BM continues to adapt even during double talk frames (recall Sec. V-B).
Fig. 7. Sonograms: (a) Desired microphone #1, (b) Nonstationary interference microphone #1; directional noise field: (c) Noisy microphone #1, (d) Enhanced signal; diffused noise field: (e) Noisy microphone #1 and (f) Enhanced signal.
VIII. Conclusions

We presented an extension of the TF-GSC, which is capable of removing both stationary and nonstationary interference signals. The proposed method is shown to have a GSC structure with several components modifications. The MBF block is now designed for blocking the interference signal while the desired signal direction is maintained. The blocking matrix is modified to block both the desired and the interference signals. Consequently, the ANC block has the same role as in the conventional GSC, i.e. eliminating the residual noise components in the MBF output.

A new system identification method was derived for estimating the blocking matrix terms directly, using double talk frames. Hence, an update of the BM can be conducted whenever the desired speech signal is active, allowing for a more accurate estimate.

The proposed system may be applied to many interesting problems. One possible application is the BSS problem with convolutive mixtures and additive noise, in which two signals of interest should be extracted. The two sources can be separated by exchanging the roles of the desired and competing speech signals. As opposed to conventional frequency-domain BSS methods, no permutation and gain ambiguities are encountered in the proposed method. Another application is the joint echo cancellation and noise reduction problem, obtained by replacing the competing speech with an echo signal. Note, however, that in this case the input echo signal is available and should be used to improve the obtained performance.

An experimental study shows the applicability of the proposed method to the problems at hand.

APPENDIX

A. Proof of Eq. (10)

Imposing the constraints defined in (6) on (9) yields

\[ -A^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) \left[ \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) \right] = F(e^{j\omega}) \]

\[ -B^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) \left[ \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) \right] = 0. \]  

(49)

Solving for the Lagrange multipliers yields:

\[
\begin{bmatrix}
\lambda_1 \\
\lambda_2
\end{bmatrix} = C^{-1} \begin{bmatrix}
-F(e^{j\omega}) \\
0
\end{bmatrix}
\]  

(50)

where \( C \) is defined as

\[
C \equiv \left[ \begin{array}{c}
A^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) A(e^{j\omega}) \\
B^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) B(e^{j\omega})
\end{array} \right].
\]  

(51)

The solution for the linear equations is then given by

\[
\lambda_1 = \frac{-B^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) B(e^{j\omega}) F(e^{j\omega})}{A^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) A(e^{j\omega}) B^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) B(e^{j\omega}) - B(e^{j\omega}) B^\dagger(e^{j\omega})\Phi_{ZZ}^{-1}(t,e^{j\omega}) A(e^{j\omega})}
\]  

(52)
\[
\lambda_2 = \frac{B(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega})F(e^{j\omega})}{A(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})[A(t, e^{j\omega})B(t, e^{j\omega}) - B(t, e^{j\omega})B'(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega})]} \tag{53}
\]
and therefore the optimal solution is:
\[
W_{opt}(t, e^{j\omega}) = -\Phi_Z^{-1}(t, e^{j\omega})(\lambda_1 A(t, e^{j\omega}) + \lambda_2 B(t, e^{j\omega})) = F(e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega}) \times
\]
\[
\frac{B'(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega}) - B(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega})B(t, e^{j\omega})}{A'(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})[A(t, e^{j\omega})B(t, e^{j\omega}) - B(t, e^{j\omega})B'(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega})]}. \tag{54}
\]
Dividing the nominator and denominator of last term by \((A(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})A(t, e^{j\omega}))(B'(t, e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega})B(t, e^{j\omega}))\) yields
\[
W_{opt}(t, e^{j\omega}) = F(e^{j\omega})\Phi_Z^{-1}(t, e^{j\omega}) \frac{A(t, e^{j\omega})}{(A(t, e^{j\omega}))^2 - \mu B(t, e^{j\omega})B'(t, e^{j\omega})A(t, e^{j\omega})}.
\]
Rearranging terms and using definitions (11) and (12) yields (10).

B. Proof of Eq. (14)

Imposing our constraints on (13) yields the following linear equations:
\[
A'(t, e^{j\omega})[W(t, e^{j\omega}) - \mu (\Phi_Z(t, e^{j\omega})W(t, e^{j\omega}) + \lambda_1 A(t, e^{j\omega}) + \lambda_2 B(t, e^{j\omega}))] = F(e^{j\omega})
\]
\[
B'(t, e^{j\omega})[W(t, e^{j\omega}) - \mu (\Phi_Z(t, e^{j\omega})W(t, e^{j\omega}) + \lambda_1 A(t, e^{j\omega}) + \lambda_2 B(t, e^{j\omega}))] = 0. \tag{55}
\]
Rearranging terms in (55) yields
\[
\lambda_1 \mu \|A(t, e^{j\omega})\|^2 + \lambda_2 \mu \|B(t, e^{j\omega})\|^2 = A'(t, e^{j\omega})W(t, e^{j\omega}) - \mu A'(t, e^{j\omega})\Phi_Z(t, e^{j\omega})W(t, e^{j\omega}) - F(t, e^{j\omega})
\]
and therefore
\[
\begin{bmatrix}
\|A(t, e^{j\omega})\|^2 & A'(t, e^{j\omega})B(t, e^{j\omega}) \\
B'(t, e^{j\omega})A(t, e^{j\omega}) & \|B(t, e^{j\omega})\|^2
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\
\lambda_2
\end{bmatrix}
= \frac{1}{\mu}
\begin{bmatrix}
A'(t, e^{j\omega})W(t, e^{j\omega}) - \mu A'(t, e^{j\omega})\Phi_Z(t, e^{j\omega})W(t, e^{j\omega}) - F(t, e^{j\omega}) \\
B'(t, e^{j\omega})W(t, e^{j\omega}) - \mu B'(t, e^{j\omega})\Phi_Z(t, e^{j\omega})W(t, e^{j\omega})
\end{bmatrix}. \tag{57}
\]
Solving the linear equations yields:
\[
\begin{bmatrix}
\lambda_1 \\
\lambda_2
\end{bmatrix}
= \frac{1}{\mu}
\begin{bmatrix}
\|B(t, e^{j\omega})\|^2 & -A'(t, e^{j\omega})B(t, e^{j\omega}) \\
-B'(t, e^{j\omega})A(t, e^{j\omega}) & \|A(t, e^{j\omega})\|^2
\end{bmatrix}
\begin{bmatrix}
\|B(t, e^{j\omega})\|^2 & -A'(t, e^{j\omega})B(t, e^{j\omega}) \\
-B'(t, e^{j\omega})A(t, e^{j\omega}) & \|A(t, e^{j\omega})\|^2
\end{bmatrix}
\begin{bmatrix}
A'(t, e^{j\omega})W(t, e^{j\omega}) - \mu A'(t, e^{j\omega})\Phi_Z(t, e^{j\omega})W(t, e^{j\omega}) - F(t, e^{j\omega}) \\
B'(t, e^{j\omega})W(t, e^{j\omega}) - \mu B'(t, e^{j\omega})\Phi_Z(t, e^{j\omega})W(t, e^{j\omega})
\end{bmatrix}. \tag{58}
\]
Define
\[ \alpha \equiv \|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2 - A^\dagger(e^{j\omega})B(e^{j\omega})A(e^{j\omega}) = \left(1 - \|\rho(e^{j\omega})\|^2\right) \|A(e^{j\omega})\|^2 \|B(e^{j\omega})\|^2 \] (60)

then
\[
\lambda_1 = (\alpha \mu)^{-1} \cdot \left[ \|B(e^{j\omega})\|^2 \left( A^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu A^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) - F(e^{j\omega}) \right) - A^\dagger(e^{j\omega})B(e^{j\omega}) \left( B^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu B^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right) \right] (61)
\]

\[
\lambda_2 = (\alpha \mu)^{-1} \cdot \left[ -B^\dagger(e^{j\omega})A(e^{j\omega}) \left( A^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu A^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) - F(e^{j\omega}) \right) \right.
\]
\[
\left. + \|A(e^{j\omega})\|^2 \left( B^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu B^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right) \right]. (62)
\]

The next expression is utilized in calculating \(W(t + 1, e^{j\omega})\):
\[
\mu \left( \lambda_1 A(e^{j\omega}) + \lambda_2 B(e^{j\omega}) \right)
\]
\[
= \alpha \|B(e^{j\omega})\|^2 \left[ A^\dagger(e^{j\omega})A^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu A^\dagger(e^{j\omega})A^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) - A(e^{j\omega})F(e^{j\omega}) \right]
\]
\[
- \alpha \left[ A(e^{j\omega})A^\dagger(e^{j\omega})B(e^{j\omega})B^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu A(e^{j\omega})A^\dagger(e^{j\omega})B(e^{j\omega})B^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right]
\]
\[
- \alpha \left[ B(e^{j\omega})B^\dagger(e^{j\omega})A^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu B(e^{j\omega})B^\dagger(e^{j\omega})A^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right]
\]
\[
+ \alpha \left[ B(e^{j\omega})B^\dagger(e^{j\omega})A(e^{j\omega})F(e^{j\omega}) + \|A(e^{j\omega})\|^2 \right]
\]
\[
\cdot \left( B(e^{j\omega})B^\dagger(e^{j\omega})W(t, e^{j\omega}) - \mu B(e^{j\omega})B^\dagger(e^{j\omega})\Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) \right). (63)
\]

Rearranging terms in (63) yields
\[
= \alpha \left[ A(e^{j\omega})A^\dagger(e^{j\omega}) \left( \|B(e^{j\omega})\|^2 I - B(e^{j\omega})B^\dagger(e^{j\omega}) \right) - B(e^{j\omega})B^\dagger(e^{j\omega}) \left( A(e^{j\omega})A^\dagger(e^{j\omega}) - \|A(e^{j\omega})\|^2 I \right) \right] W(t, e^{j\omega})
\]
\[
- \alpha \mu \left[ A(e^{j\omega})A^\dagger(e^{j\omega}) \left( \|B(e^{j\omega})\|^2 I - B(e^{j\omega})B^\dagger(e^{j\omega}) \right) - B(e^{j\omega})B^\dagger(e^{j\omega}) \left( A(e^{j\omega})A^\dagger(e^{j\omega}) - \|A(e^{j\omega})\|^2 I \right) \right]
\]
\[
\cdot \Phi_{ZZ}(t, e^{j\omega})W(t, e^{j\omega}) + \alpha \left[ -\|B(e^{j\omega})\|^2 A(e^{j\omega}) + B(e^{j\omega})B^\dagger(e^{j\omega})A(e^{j\omega}) \right] F(e^{j\omega}). (64)
\]

Substituting (64) into (13) we obtain (14).

REFERENCES


