Binaural Beamforming Robust to Errors in Direction of Arrival Estimates

By

Parinaz Khayeri

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of
Master of Applied Science
in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa
March 2016

Abstract

Binaural beamforming technology, which is based on the auditory perception of both ears, uses a wireless data connection to exchange data between the right-side and the left-side hearing aids. Over the years, several multichannel speech enhancement algorithms have been used in the hearing aid industry. For example, beamforming algorithms work by keeping a target signal undistorted while attenuating the noise fields (such as diffuse noise or white noise) and the interferers from different directions. Fixed and adaptive algorithms of this nature have been under active investigation by the hearing aid industry. Although binaural beamforming hearing aids designs have shown better performance than single-channel based hearing aids or bilateral hearing aids, the performance of binaural beamforming still suffers from errors in the direction of arrival estimates, i.e., errors which occur when the right set of steering vectors is used in a beamformer design but the target signal source is not located at the direction considered in the design. Therefore, this thesis is devoted to find and propose structures showing more robustness to errors in the direction of arrival estimates. The focus is mainly on the Generalized Sidelobe Canceller (GSC) structure and several binaural beamforming algorithms and configurations are proposed in this thesis as alternatives for the fixed beamformer and blocking matrix units of the GSC. The proposed algorithms show promise of providing wider notch and/or wider beam possibilities, as well as providing greater noise reduction and superior adaptive null positioning capabilities. The algorithms proposed in this thesis were simulated in MATLAB using recorded signals and data provided by a hearing aid firm, to assess their utility for improving hearing aid performance. The results demonstrated a superiority over algorithms currently in use in industry.

Acknowledgements

First and foremost, I offer my sincerest gratitude to my supervisor, Professor Martin Bouchard, who has supported me throughout my thesis with his patience and knowledge. I attribute the level of my Masters degree to his encouragement and effort. Without his thoughtful suggestions and persistent help this thesis, too, would not have been completed or written. One simply could not wish for a better or friendlier supervisor.

Moreover, I would like to thank Dr. Homayoun Kamkar-Parsi for his technical feedback and valuable support.

I would also like to thank my best friend, Neda. Thank you for sharing in my happiest moments, and for genuinely feeling the same; for listening to me in hard times, radiating compassion and empathy from wherever you are.

Moreover, I would like to thank my cute niece, Avina, and my sister, Parisa. All my life I knew that I am so lucky because I had you. Thank you for everything you`ve done for me, my dear and lovely family. I owe a debt of gratitude to my mother for her wise counsel and sympathetic ear. She is always supporting me spiritually throughout my life no matter how far I am from her and what happens. Thank you for being the only person I ever want to confide in. Finally, I would like to offer my gratitude to my father who passed away years ago. I could no longer see him and hear his kind words and receive his warm support, but all the time I feel him as a magnificent motivation in my heart to make his soul happy. My thesis is dedicated to my bestest dearest parents.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ANC	Adaptive Noise Canceller
ADMA	Adaptive Differential Microphone Array
BTE	Behind-The-Ear
CIC	Completely-In-the-Canal
DMA	Differential Microphone Array
DOA	Direction of Arrival
GSC	Generalized Sidelobe Canceller
HRTFs	Head-Related Transfer Functions
ITC	In-The-Canal
ITE	In-The-Ear
LCMV	Linearly Constrained Minimum Variance
LMS	Least Mean Square
MMSE	Minimum Mean Squared Error
MVDR	Minimum Variance Distortion Response
PSD	Power Spectral Density
RITE	Behind-the-ear hearing aid with the Receiver-In-The -Ear
SDR	Signal to Distortion Ratio
SNR	Signal to Noise Ratio
STFT	Short-Time Fourier Transform
TDOA	Time Difference of Arrival
VAD	Voice Activity Detection

Chapter 1 **Introduction**

1.1 Motivation and Previous Work

A hearing aid is an electroacoustic device designed to amplify and enhance sounds for the hearing-impaired user with the aim of making a target speaker voice more intelligible [1]. Over the years many kinds of hearing aid devices have been marketed such as Completely-In-the-Canal (CIC), In-The-Canal (ITC), In-The-Ear (ITE), Behind-The-Ear (BTE), and behind-the-ear hearing aid with the Receiver-In-The-Ear canal (RITE) [2].

Various hearing aid systems using multichannel signal processing techniques have been introduced to the market, including monaural hearing aids, bilateral hearing aids and binaural hearing aids [3]. Monaural hearing aid systems have a microphone array set on one ear which uses the signals received by the microphones to obtain the enhanced signal for the hearing-impaired user. In bilateral hearing aid systems, there is a hearing aid (i.e., a microphone array) device on both sides, but there is no wireless link or information transformation between the two sides. Binaural hearing aid systems have recently been introduced to have better performance in noise reduction and target extraction in realistic reverberant and noisy environments. The binaural hearing aid systems have a wireless connection between the left and right hearing aids. Therefore, the hearing aid on each side can use the information of the signals received on the remote side. This is equivalent to having more sensors for each hearing aid and consequently it can achieve better performance in terms of speech intelligibility and quality [4, 5].

Hearing-impaired people have decreased ability to hear a signal that rapidly follows, or is rapidly followed by, a different signal, and might also be less able to recognize the directions of arrival of the sounds that they receive. They usually face decreased audibility or decreased resolution (frequency, temporal) compared to a normal-hearing person. Decreased audibility makes most hearing-impaired people miss much of the high frequency information of speech components, because the high-frequency components of speech are weaker than the low-frequency components. Hearing-impaired people also suffer from temporal resolution loss, with intense

sounds masking weaker sounds that immediately precede or follow them. Decreased frequency resolution is another difficulty faced hearing-impaired people. It results in a difficulty of separating sounds of different frequencies so that the brain cannot separate speech from noise, even in simple noise fields where the speech component and the noise are sufficiently separated spatially [4].

To assist in mitigating the effects of hearing loss, hearing aids have been designed and used for several decades. Over the years the hearing aid industry has devoted significant research and development resources to improve speech enhancement and noise reduction. The two general types of signal processing which have been used for years in many applications such as hearing aids are single-channel and multichannel configurations [6]. Single channel speech enhancement algorithms use one microphone and attempt to reduce noise and improve quality in the output signals. Typical single channel algorithms cannot perform very well when the input SNR is low or in non-stationary noise fields or in reverberant environments. They usually create musical noise especially in high level noise environments. On the other hand, multichannel speech processing algorithms using arrays of microphones outperform the single channel algorithms in terms of de-noising and speech intelligibility enhancement [7].

A beamforming or multichannel signal processing algorithm is a kind of spatial filtering which uses a microphone array (a group of sensors located in a certain spatial position e.g., linear or nonlinear, one dimensional, two dimensional or three dimensional array) to try to extract the target signal from the various microphone signals. At the same time, sounds coming from other directions are attenuated (including the environment noise or other interfering speakers) [7]. This requires the locations of the speakers to be known in advance, for example by using the time difference of arrival (TDOA) from the sources to each microphone in the array, and inferring the locations from the distances. TDOA is the time difference between a source signal received at the reference microphone and at the other microphones of the array [7, 8].

The delay and sum beamformer is the simplest beamforming algorithm in the categories of fixed beamformers and frequency dependent algorithms. For the free field condition, the target input signals received by the sensors are first time-aligned based on their corresponding TDOA. The output of beamformer is then calculated by adding the target-aligned input signals. In other

words, the delay and sum beamformer adds coherently the time aligned target source signals, while adding incoherently the signals for the other sources [7, 8].

The Minimum Variance Distortion Response (MVDR) beamforming algorithm, a member of the filter and sum beamformer category, does not apply the same weights to the various input signals as the delay and sum beamformer algorithms does [9, 10]. The MVDR beamformer is based on a constrained minimization, where the response of the beamformer in the direction of the target signal is constrained to unity and the power of noise signals (noise and interferers coming from other directions) is minimized at the same time. Knowledge of the target steering vector is one of the requirements of MVDR design [11].

The Linearly Constrained Minimum Variance (LCMV) is another type of filter and sum beamforming algorithm which can be considered as a more general version of the MVDR. In MVDR design the minimization problem is based on only one constraint (for the target direction) whereas the LCMV solution is based on multiple constraints [11, 12]. Therefore, for the LCMV design, not only the target steering vector is required, but the steering vectors corresponding to the other direction constraints are also necessary.

The Generalized Sidelobe Canceller (GSC) is another widely used beamforming structure which can be considered as an adaptive beamforming algorithm [13]. Adaptive beamforming algorithms have the ability to adaptively get some information on the TDOA of target or noise signals and the noise field type, so that they can adaptively locate the nulls and the main lobe of a beampattern in the correct directions. In many situations the GSC is theoretically equivalent to the LCMV beamformer but it is implemented in two paths [14]. The first path utilizes a fixed beamformer with a response that has the specified gains in the directions specified in the LCMV constraints (normally including a high gain in the target direction). The second path includes a blocking matrix where nulls are positioned in each of the directions specified in the LCMV constraints (normally including a null in the target direction). The second path also includes an adaptive noise canceller (ANC) [13]. The ANC system attempts to remove the noise signal components in the fixed beamformer output, which are correlated with the outputs of the blocking matrix.

Another adaptive beamforming algorithm introduced for an array with 2 microphones is the adaptive back-to-back first-order differential microphones algorithm (ADMA) [15]. This algorithm can be used to adaptively adjust the null response of a beampattern in order to track a possibly moving noise source in the back hemisphere [15, 16]. Forward cardioid and backward cardioid signals can be formed by the subtraction of the delayed microphone signals by the factor of d/c (d is the space between microphones and c is the speed of sound). The backward cardioid output is generated by reversing the roles of the microphones compared to the forward cardioid. The ADMA output is generated by multiplying the backward-cardioid output by a factor β, and then subtracting the result from the front-cardioid output. β can be adaptively updated by using the Least Mean Square (LMS) algorithm or a normalized version of the LMS. The system is adjusted to adaptively place a null when the signal direction angle is varying from 90 to 180, where 0 degree is defined here as being collinear with the microphones, in the direction of the reference microphone. The advantage of the ADMA algorithm is that it does not require knowledge of the target or noise steering vectors [16, 17].

1.2 Objectives, Contributions and Organization

The objectives of this thesis are as follows:

- To develop binaural beamforming algorithms which are more robust to inaccuracies in the assumed direction of arrival (DOA) for the target source, for a range of about +-20 degrees around an assumed target location;
- To validate the developed algorithms by performing simulations based on real-life recordings provided by a hearing aid manufacturer;
- As a secondary objective, an attempt will also be made to design an algorithm that can improve the performance of the GSC algorithm for diffuse noise.

The main contributions of the thesis are:

- to provide a basic review of several beamforming algorithms, i.e. GSC, MVDR, LCMV constraint-based design and ADMA design in the context of binaural hearing aids

- to propose an LCMV design or constraint-based design instead of the classic Griffiths-Jim blocking matrix in the GSC, in order to obtain a wider notch in the target direction and consequently provide a more robust algorithm to cope with the DOA mismatch
- to propose an ADMA beamformer instead of the classic Griffiths-Jim blocking matrix in the GSC due, to obtain adaptive notch-positioning capacity, again to introduce a more robust algorithm to DOA mismatch
- to propose an LCMV design or constraint-based design to replace the target equalized and sum fixed beamformer in the GSC, in order to obtain a wider beam in the target direction to improve the system robustness to DOA mismatch
- to propose an MVDR design tuned for diffuse noise instead of the target equalized and sum fixed beamformer in the GSC, in order to provide better capacity in diffuse noise reduction
- to develop Matlab codes for all the above algorithms, supporting different microphone configurations and acoustic environments
- to investigate the performance of the considered algorithms in complex noisy acoustic environments including reverberation, diffuse-like noise and multiple directional interfering sources
- to investigate the performance of the considered algorithms using objective measures related to speech quality and intelligibility
- to evaluate the robustness of proposed algorithms to HRTFs mismatch and DOA mismatch.

To achieve these objectives, the thesis is organized in the following way:

- Chapter 2 describes the concept of binaural hearing aids, the different microphone configurations and some necessary definitions. In addition, it provides a review of the GSC structure, the MVDR and LCMV beamforming, and the back to back cardioid ADMA algorithms.
- Chapter 3 describes the problems with the traditional GSC and proposes to use some alternatives for the fixed beamformer and the blocking matrix in order to solve those problems and achieve a better robustness to DOA mismatch for the new models used in this thesis.

- Chapter 4 describes the experimental setup first and afterwards it introduces some speech quality and intelligibility related objective measures.
- Chapter 5 includes the simulation results for the different structures presented in Chapter 3. The beampatterns and directivity indexes of different blocking matrix and fixed beamformer designs are shown and investigated. Afterwards the performance of the considered algorithms in complex noisy acoustic environments including reverberation, diffuse-like noise, and multiple directional interfering sources are presented.
- Chapter 6 includes a conclusion as the ending part of the thesis, discussing the main results of the thesis and presenting some possible future work.

Chapter 2 Overview of Binaural Beamforming Configurations and Various Beamforming Algorithms

This chapter provides an overview of the various methods of multichannel speech enhancements. This thesis reviews these various methods as they relate to binaural beamforming. It should be noted that binaural hearing aids involve the extension of the beamformer design into two beamformer designs, whereby the hearing aid on each side has its own output. The left side output derived from the beamformer design uses the front left microphone as the reference microphone, while the right side output derived from the beamformer design uses the right side microphone as the reference microphone. Therefore, all the approaches in this section can be seen as applicable for the local side beamformer, where the local front microphone is selected as the reference microphone. Binaural hearing aids will be considered where we will have two microphones at each ear and a wireless link between them.

2.1 Binaural Beamforming

The hearing aids currently existing in the market are either monaural hearing aids, bilateral hearing aids or binaural hearing aids [3]. A monaural hearing aid only uses its own microphone inputs to generate an output. A pair of bilateral hearing aids is composed of two independent monaural hearing aids on each side, and typically little information is shared between the hearing aids from the two ears. In contrast, in binaural hearing aids, it is assumed that a wireless link allowing the exchange of information between the two sides is available, so that the outputs of both ears are obtained by processing the sensor inputs from both the left and right sides.

In this thesis, the binaural beamforming design is based on binaural Behind-The-Ear hearing aids (BTE) with two microphones at each side [2]. Different possible configurations for the binaural beamformer design in BTE hearing aid can be considered. These configurations are introduced in this section. In binaural hearing aids, it is required to produce two outputs (one output on each side). The binaural beamformer on the left side uses the front

left microphone as a reference and the binaural beamformer on the right side uses the front right microphone as a reference, and they must run in parallel.

2.1.1 2+2 Structure with Two Wireless Connections

The first beamformer design can be achieved by having a direct access to the four microphones signals, as Figure 2.1 shows.

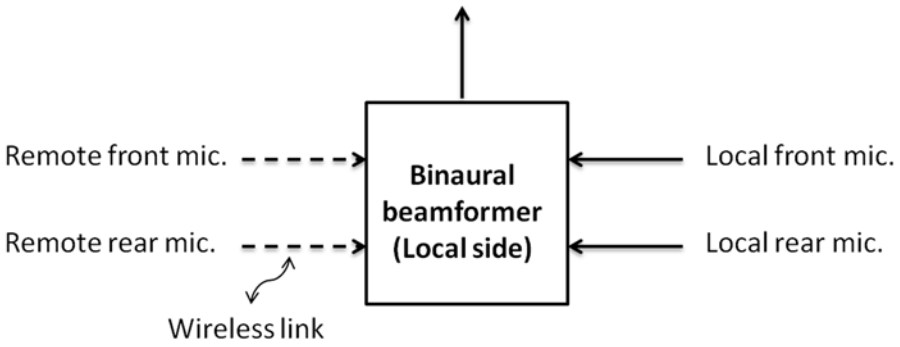


Figure 2.1 2+2 structure with two wireless links

We will refer to this configuration by 2+2 structure. As long as we have a direct access to the four microphone signals, this structure has access to all correlations between the four available signals. Therefore, it theoretically can perform more optimally than the other configurations. However, the binaural beamformer on each side needs to have access to two transmitted signals from the other side through a wireless link, which is a highly bandwidth-consuming and power-consuming operation. In this thesis, we have assumed an ideal wireless link. Nevertheless, practically it is not a realistic assumption, as the wireless link suffers from bandwidth limitations, jitter, delay, etc.

2.1.2 2+1 Structure Without Pre-Processing

Another configuration can be used for binaural beamforming by transmitting only one signal to the other side so that only three microphone inputs are involving in the design. Although, this structure does not take the advantage of having four microphones, which might lead to degradation in the performance, it reduces the bandwidth requirements and consequently the power

consumption. The structure is shown in Figure 2.2. In this configuration, which we will refer to as 2+1 structure without pre-processing, each side transmits only one signal to the other side.

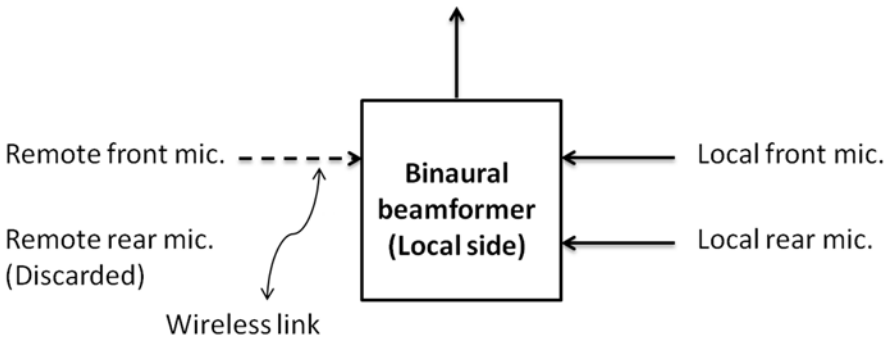


Figure 2.2 2+1 structure without pre-processing

2.1.3 2+1 Structure with Pre-processing on One Side Only

Figure 2.3 shows the configuration of 2+1 structure with pre-processing on only one side. It can be considered as a variation of the previous structure where each side only transmits one signal to the other side. In this structure, the only signal to be transmitted to the other side is obtained by a monaural beamformer applied to the inputs of the local side array; whereas, in the previous 2+1 structure without pre-processing, the signal to be transmitted to the other side is directly one of the local microphone inputs (i.e., the reference microphone, front microphone). In this way, we are taking advantage of having four available microphones in contrast to the previous 2+1 structure without pre-processing.

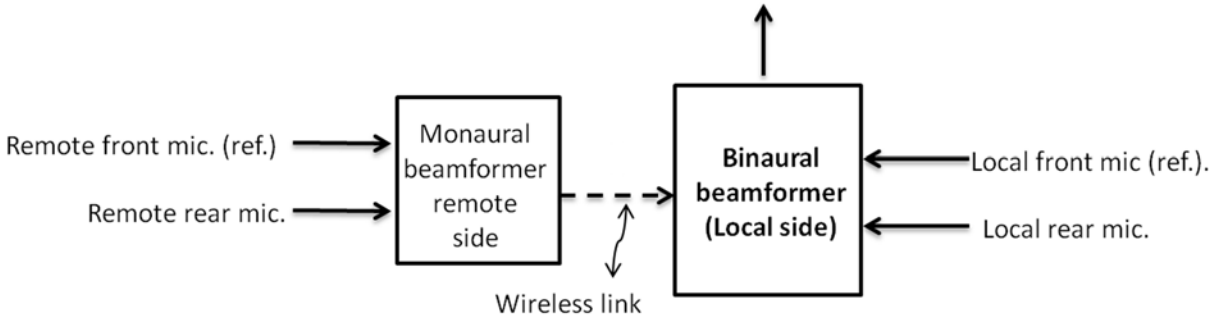


Figure 2.3 2+1 structure with pre-processing on one side only

2.1.4 1+1 Structure with Pre-processing on Both Sides

Another configuration that needs only one wireless link and takes the advantage of the four microphones is shown in Figure 2.4. We will refer to this configuration as 1+1 structure with pre-processing on both sides. In this configuration, a monaural beamformer output is computed on each side, to be transmitted to the other side. Therefore, the binaural beamformer on each side becomes simpler since there are only two input channels to the binaural beamformer. On the other hand, the previous structure in Figure 2.3 with 3 inputs for the binaural beamformer on each side may allow to consider more correlations between the signals. Note that the 1+1 configuration is a 4-microphone system, i.e., the 1+1 name just comes from the fact that the binaural beamformer receives 1 signal from each side.

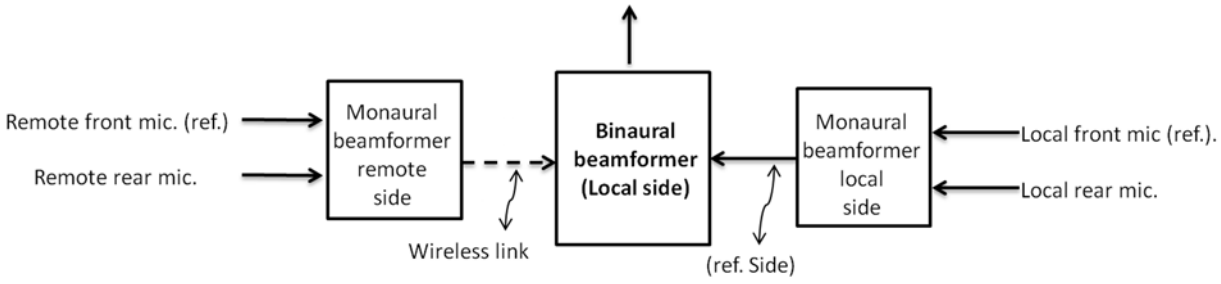


Figure 2.4 1+1 structure with pre-processing on both sides

2.2 Definitions

In this section, there is an overview on several concepts and definitions which are highly used in this thesis. The concepts are relative to speech processing and hearing aids technology.

2.2.1 Steering Vector

The steering vector or steering vector $\underline{d}$ is defined here as the frequency response from a far field source at an angle θ_s, and at a constant distance for any angle, when the free field condition is considered (i.e., an environment with plane wave far field propagation with no physical object affecting propagation, and with no head shadow effect existing) [9, 10]. The steering vector is defined in equation (2.1).

$$\underline{X}_s(f) = \underline{d}(f, \theta_s) S(f, \theta_s) \tag{2.1}$$

where $\underline{X}_s$ denotes the signal received by the microphone, S is the signal propagated from the source and f is the frequency index.

2.2.2 Head Shadow Effects/Head Related Transfer Functions (HRTFs)

A sound becomes weaker the farther away it gets from its source. If a signal first reaches the left ear, its spread to the right ear is both delayed and weakened by the head shadowing effect. This means that the right ear perceives the signal less strongly than the left one, with a delay. If this effect is impaired, it has an especially negative influence on the hearing ability. Therefore, it

affects both amplitude attenuation and complex filtering effects, which are called head shadow effects. The HRTFs describe the complex filtering effects of the sound diffraction and reflection caused by the pinnae, head, shoulders and torso before the sound reaches the ear drums [18]. They also include the interaural effects such as the interaural time differences (ITD) and interaural level differences (ILD). The HRTFs vary considerably from person to person.

2.2.3 Beampattern

A beampattern is the response of the microphone array, which is used to display the information about the beamformer behavior [7, 10]. It has a main lobe for the target signal direction (i.e. the beamforming system allows the target signal to pass), and some nulls (or some small gain lobes) that are evidence of undesired signals attenuation.

Assuming $S(\omega)$, $X_i(\omega)\ i=1...M$, and $Y(\omega)$ are all frequency domain signals which can be considered as the output of a short-time Fourier transform applied to an input frame of time samples. θ_s is the direction of the target signal in far-field propagation, $S(\omega,\theta_s)$ is the source signal coming from angle θ_s, $X_i(\omega)\ i=1...M$ are the microphone array received signals, $Y(\omega)$ is the monaural output of the beamformer, M is the number of microphones belonging to the array.

The beampattern, $BP(\omega,\theta)$, is defined as the squared magnitude response for a signal from a single arbitrary direction. The beampattern is defined by equation (2.2) and equation (2.3).

$$\begin{aligned} E\left[\left|Y_s(\omega)\right|^2\right] &= E\left[\left(\underline{W}^H(\omega)\underline{X}_s(\omega)\right)\left(\underline{W}^H(\omega)\underline{X}_s(\omega)\right)^H\right] \\ &= \underline{W}^H(\omega)E\left[\underline{X}_s(\omega)\underline{X}_s{}^H(\omega)\right]\underline{W}(\omega) \\ &= \underline{W}^H(\omega)\left(\underline{d}(\omega,\theta_s)\underline{d}^H(\omega,\theta_s)E\left[\left|S(\omega,\theta_s)\right|^2\right]\right)\underline{W}(\omega) \\ &= \left|\underline{W}^H(\omega)\underline{d}(\omega,\theta_s)\right|^2 E\left[\left|S(\omega,\theta_s)\right|^2\right] \end{aligned} \tag{2.2}$$

and thus we obtain:

$$BP(\omega,\theta_s) = \left|\underline{W}^H(\omega)\underline{d}(\omega,\theta_s)\right|^2 \quad (2.3)$$

The beampattern is defined here for far field propagation.

2.2.4 Diffuse Noise Field

A diffuse noise field is defined as equally distributed and uncorrelated noise source components coming from all directions [10, 19]. Diffuse noise is representative of many noise environments, for example the babble noise is quite diffuse-like.

2.3 Generalized Sidelobe Canceller (GSC)

The Generalized Sidelobe Canceller (GSC) is one of the most highly used structures [10]. The block diagram of the basic GSC system is shown in Figure 2.5. *M* denotes the number of microphones. In this thesis it is equal to 4 (2+2 configuration) or 3 (2+1 configuration). The system includes two paths [10, 13]:

1) The "standard" fixed beamformer that we will attempt to improve has a single constraint on the target direction. It is a generalized version of a Delay and Sum beamformer, which considers the head shadow effect when target-aligning or target-equalizing the different microphone signals before summing them.

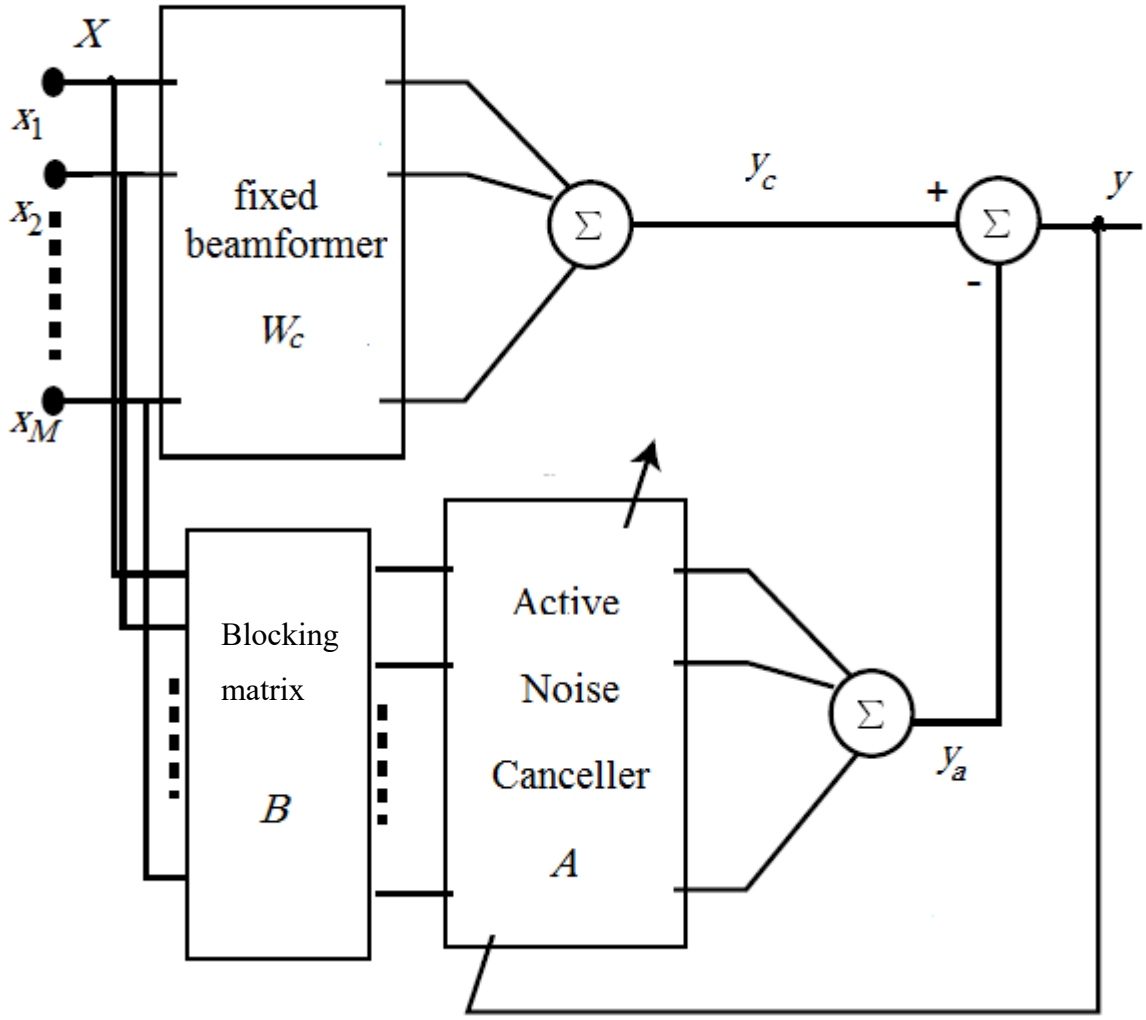


Figure 2.5: Generalized Sidelobe Canceller

2) An adaptive noise canceller (ANC) whose function is to minimize the noise power at the output. The second path also includes a blocking matrix that removes the target signal from the original noisy signal, so that the outputs of the blocking matrix can be used as noise reference signals for the linear prediction to be made by the ANC unit. The ANC system duty is to remove the undesired signal components, correlated with the outputs of the blocking matrix, from the fixed beamformer output. The ANC part is normally adaptive in practice and it can be updated using an unconstrained adaptive algorithm, such as the least-mean-square (LMS) [13]. As a result, it can deal with non-stationary interferers and background noise. As mentioned above, the blocking

matrix excludes the target signal at the input of adaptive section, so the beampattern of the blocking matrix has a null in the direction of the target signal.

The equations below describe the GSC system for a full-band frequency scenario and for certain subband frequency scenarios. The blocking matrix output is given by:

$$\underline{X'}(n) = B^H \underline{X}(n) \tag{2.4}$$

where B is M-1 by M blocking matrix, $\underline{X(n)}$ is target-aligned or target-equalized input signals and $\underline{X'}(n)$ is the outputs of the blocking matrix. As long as the inputs of the blocking matrix are already equalized input signals, the blocking of the target signal will occur when the rows of B sum to zero and the rows are linearly independent. Therefore, there are M-1 linearly independent outputs after the blocking matrix unit. The ANC system output is calculated by equation (2.5):

$$y_A(n) = \underline{A}^H . \underline{X'}(n) \tag{2.5}$$

Where $\underline{A}$ denotes the ANC unit coefficients, used to minimize the noise and interferer components at the output of the fixed beamformer, using the correlation with the noise and interference components in the M-1 outputs of the blocking matrix. The ANC coefficients can be updated by a LMS algorithm by equation (2.6):

$$\underline{A}_{k+1} = \underline{A}_k + \mu y_k \underline{X'}_k \tag{2.6}$$

and the global output of the GSC system is described by the equation below:

$$y(n) = y_c(n) - y_A(n) \tag{2.7}.$$

In practice, the GSC can cause some target signal distortion at the global output due to target leakage in the blocking matrix. This happens when the blocking matrix fails to completely remove the target signal from the lower path. It can be very problematic especially for hearing aids devices. The GSC structure is generally applicable for different configurations of binaural beamforming (for binaural 2+2, binaural 2+1, binaural 1+1, monaural 1+1 configurations). For

binaural beamforming, to obtain each local side output, the reference microphone must be the front local microphone.

2.4 Minimum Variance Distortion Response (MVDR) Design

The MVDR design is known as a very powerful methodology for suppressing a very wide range of noise fields (e.g. diffuse noise field which has different characteristic in comparison with the directional interferers) [9, 20]. All the explanations here have a general form so that the MVDR can be applicable either for monaural 1+1, binaural 1+1, binaural 2+1 or binaural 2+2 configurations.

The Minimum Variance Distortion Response (MVDR) solution belongs to the family of constrained optimization. The MVDR has been developed based on the concept of applying a minimization on the power of the noise signals [9], with a constraint on the gain in the target signal direction. In another words, the coefficients of this beamformer are designed to maximize the Array Gain (AG) given by equation (2.8) [21, 22]:

$$AG = \frac{SNR_{out}}{SNR_{in}} = \frac{G_d}{G_n} \,. \tag{2.8}$$

where G_d is the beampattern gain at the target (desired) signal direction and G_n is the average of the beampattern gain at all noise + interferers signal direction. For the free field scenario, for the target and noise signals the following definitions are given [10]:

$$\phi_{ss}(out) = \phi_{ss}(in)\left|\underline{W}^H \underline{d}\right|^2, \tag{2.9}$$

$$\varphi_{n_o n_o}(out) = \varphi_{v_o v_o}(in)\underline{W}^H \Gamma \underline{W} \,. \tag{2.10}$$

Where $\underline{W}$ denotes the MVDR coefficients, $\phi_{ss}(in)$ denotes the auto-power spectrum density (PSD) of the target signal at the reference microphone, $\varphi_{ss}(out)$ denotes the PSD of the output

target signal, $\phi_{v_o v_o}(in)$ denotes the PSD of the received noise signals on the reference microphone and Γ is the noise signals correlation matrix.

The desired target components at the *M* microphones can be written in term of the steering vector as equation (2.11) shows:

$$\underline{x}^{in} = \underline{d}\, \mathrm{s}_x \tag{2.11}$$

s_x : Target source signal at the far field plane.

$\underline{x}^{in}$: Target components at the *M* microphones

$\underline{d}$: steering vector or steering vector, frequency response between the far field target source and each microphone.

For free field conditions and 1-D uniformly distributed linear arrays, $\underline{d}$ can be defined as:

$$\underline{d} = \left[1 \quad \exp(-j2\pi f d \frac{\cos\varphi.\sin\theta}{c}) \quad \ldots \quad \exp(-j2\pi f (M-1) d \frac{\cos\varphi.\sin\theta}{c}) \right]^T, \tag{2.12}$$

The correlation matrix for the target components can be defined as equation (2.13).

$$R_x = E\{\underline{x}\underline{x}^H\} = E\{\underline{d}s_x\underline{d}^H s_x^*\} = \underline{d}\underline{d}^H E\{|\mathrm{s}_x|^2\} \tag{2.13}$$

where the superscript *H* is the complex conjugate transpose.

Similarly, the correlation matrices for the sum of the directional interference components and the background noise components, respectively, can be defined as in equations (2.14) and (2.15).

$$R_i = E\{\underline{i}\,\underline{i}^H\} \tag{2.14}$$

$$R_n = E\{\underline{n}\underline{n}^H\} \tag{2.15}$$

where $\underline{i}$ and $\underline{n}$ are respectively interferer components and noise components at the M microphones.

Assuming that the sum of directional interference components and the background noise components are uncorrelated, the correlation matrix for noise signals can be written as in equation (2.16):

$$\Gamma = R_i + R_n \tag{2.16}$$

By computing SNR_{out} and SNR_{in} using equations (2.9) and (2.10), AG can be written as in equation (2.17) [23]:

$$AG = \frac{\left|\underline{W}^H \underline{d}_0\right|^2}{\underline{W}^H \Gamma \underline{W}}. \tag{2.17}$$

As mentioned at the beginning of this section, the MVDR coefficients are designed to maximize AG. This can be specified as in equation (2.18):

$$\underline{W}_{opt} = \arg\max_{\underline{W}} \frac{\left|\underline{W}^H \underline{d}_0\right|^2}{\underline{W}^H \Gamma \underline{W}}. \tag{2.18}$$

As a solution for the constrained minimization problem in equation (2.18), a Lagrangian problem can be defined as in equation (2.19).

$$\max_{\mathbf{w}} \frac{\left|\underline{W}^H \underline{d}_0\right|^2}{\underline{W}^H (\Gamma) \underline{W}} \quad \text{subject to} \quad \underline{W}^H \underline{d}_0 = 1 \tag{2.19}$$

Therefore, a Lagrangian cost function is defined as in equation (2.20):

$$\mathrm{L}(\underline{W}, \underline{W}^H) = \underline{W}^H R_{undesired} \underline{W} + \lambda(\underline{W}^H \underline{d}_0 - 1) \tag{2.20}.$$

And as a result, the solution can be written as in equation (2.21):

$$\underline{W} = \frac{\Gamma^{-1}\underline{d}}{\underline{d}^{H}\Gamma^{-1}\underline{d}} \tag{2.21}$$

A regularization factor (i.e., diagonal loading) should be used in practical situations in order to avoid the inversion ill-conditioning. Consequently, the calculated MVDR weights would be written as in equation (2.22) [9, 10].

$$\underline{W} = \frac{(\Gamma + \mu \mathrm{I})^{-1}\underline{d}}{\underline{d}^{H}(\Gamma + \mu \mathrm{I})^{-1}\underline{d}} \tag{2.22}$$

2.4.1 Directivity index

The MVDR design is the most efficient algorithm for diffuse noise reduction [9, 11]. In this section the Directivity index (DI) is presented as a factor to measure the capability of MVDR design for the purpose of dealing with diffuse noise field [19]. The Array Gain (AG) becomes equivalent to the Directivity Index DI if the diffuse noise correlation matrix is used as Γ, which is defined as the ratio of the power gain from the target direction over the average power gain of diffuse noise (which is propagating from all directions in noise field). In the case of 3-D spherically isotropic diffuse noise in free field, the diffuse noise correlation matrix elements can be presented by equation (2.23):

$$\Gamma_{m,n}(e^{j2\pi f})\Big|_{diffuse} = \frac{\sin(2\pi f\, l_{mn}/c)}{2\pi f\, l_{mn}/c} \tag{2.23}$$

For each element in the diffuse noise correlation matrix between microphone *n* and microphone *m*, where $l_{m,n}$ is the distance between microphone *n* and microphone *m, f* is the frequency index and *c* is the speed of sound.

Note that the Directivity Index is specific to diffuse noise (and moreover limited to far field propagation, plane wave, point source model), while the Array Gain can be defined for other types of noise as well (e.g. directional, etc.) and for other types of propagation models [10, 23].

2.4.2 Choice of Correlation Matrix in MVDR Design

The noisy correlation matrix R_y under the assumption of uncorrelated target components, directional interference components, and background noise components can be written as in equation (2.24)

$$R_y = R_x + R_v + R_n \tag{2.24}$$

It has been investigated in the literature that the noisy correlation matrix R_y can be used instead of the correlation matrix of the noise signals Γ in most situations [24]. In many practical and realistic situations, it is not straightforward to estimate the correlation matrix of the noise signals. In some applications, voice activity detectors (VAD) are applied to estimate Γ, but for non-stationary noise fields or in the presence of interferers, moving or not, it will not end up with an accurate Γ estimate. However, using the noisy correlation matrix can lead to a decrease of performance (or decrease of robustness) in the case of mismatch between the estimated target steering vector (used for MVDR design) and the actual steering vector corresponding to the processed data.

2.5 Linearly Constrained Minimum Variance (LCMV) Design

In MVDR design the focus is on assigning the constraint of unity in the direction of the target signal, and minimizing the power of the noise signal at the output. If we have the opportunity to apply constraints for more than one direction, the beampattern can be more controllable to suppress certain interferers coming from other directions. It has been the motivation for the development of the Linearly Constrained Minimum Variance (LCMV) beamformer [12, 21]. The MVDR solution can be expanded for the more general case where the optimization is subjected to multiple linear constraints (including that of unity response for the target signal). The solution under such linear constraints is given by [12]:

$$\min_{\mathbf{w}} \underline{W}^H(\Gamma)\underline{W} \quad \text{subject to} \quad C^H_{N\times M}\underline{W}_{M\times 1} = \underline{g}_{N\times 1} \tag{2.25}$$

where N is the number of constraints, C is the constraint matrix which includes the steering vector of each constraint direction and $\underline{g}$ is the constraint gain vector including the gains for each constraint direction. Using the complex Lagrangian multiplier the coefficient vector is obtained by equation (2.26):

$$\underline{W} = \frac{\Gamma^{-1}C}{[C^{H}\Gamma C]^{-1}}\,\underline{g} \tag{2.26}$$

In the case of binaural beamformer coefficients, (2.26) would be computed separately for each local side.

The LCMV is similar to the MVDR as both of them are developed to extract the target signal which is coming from a specific direction while minimizing the noise signals from other directions. As mentioned before, the LCMV design considers multiple constraints for multiple directions and provides the advantage of a more flexible beampattern for dealing with interference.

Under some conditions the LCMV and the GSC algorithms are theoretically equivalent [13, 14]. As previously mentioned, the GSC is implemented with two paths. The first path of the GSC (fixed beamformer) should be distortionless to the target signal and have the appropriate gains for the different constraints specified in the LCMV. Also the second path consists of a "blocking matrix", which aims to provide the ANC unit with reference signals for the noise signals (while "blocking" the target signal). The blocking matrix has zero gains for the constraint directions specified in the LCMV. And finally, in the ANC system, the fixed beamformer output is used as a "desired signal" by a linear predictor, and the blocking matrix outputs are used as the input signals from which the linear prediction is made. The duty of the ANC system is to remove the undesired noise signal components, correlated with the outputs of the blocking matrix, from the fixed beamformer output.

If the LCMV has only one constraint in the target direction, the LCMV becomes equivalent to an MVDR, and consequently the GSC implementation becomes an adaptive equivalent to the MVDR. In this case, applying target alignment / target equalization to the different microphone signals can be used to generate the GSC fixed beamformer (summing the aligned signals, similar

to classic delay and sum) and to produce the blocking matrix (pairwise subtraction of the aligned signals, similar to classic Griffths-Jim blocking matrix).

However, the GSC does have certain advantages when compared with the LCMV/MVDR:

1. No need to estimate any correlation matrix

2. No need for matrix inversion,

3. Greater robustness when encountering non-stationary noise fields and moving directional noises or moving interference, due to the ANC system working adaptively.

2.6 Adaptive Differential Microphone Array (ADMA) Design

Adaptive first order Differential Microphone Arrays (ADMA) are able to adaptively track and attenuate a moving signal source that is located in the back hemisphere of a pair of microphones (in the direction of the non-reference microphone, interferer cancelling mode) [15, 16, 17]. This noise reduction is accomplished by adaptively steering a null for the direction of arrival of the noise source. The term first-order differential array comes from the fact that the signals from a 2 element microphone array are subtracted during the processing. With a linear one-dimensional microphone array including only two sensors separated by a distance *d*, a first order DMA can generate a directional response. The first order DMA configuration is shown in Figure 2.6 [25].

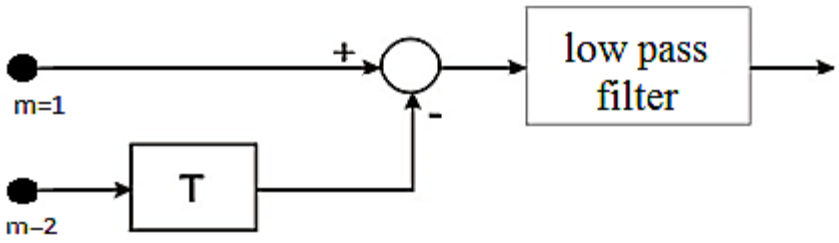


Figure 2.6 First order DMA diagram

For a far-field condition, the magnitude of the frequency response of the array is given by equation (2.27)

$$\left|H(w,\theta)\right| = \left|1 - e^{-j\omega 2(T+1/(c.\cos\theta))}\right| \tag{2.27}$$

where c is the speed of sound and θ is the target direction to the axis of the array. The delay T (propagation delay between the microphones) may be adjusted to cancel a signal from a certain direction to obtain the desired directivity response. In free field this delay T is fixed to d/c. Obviously an implementation of a non-adaptive first-order DMA would require the capability of dealing with the mismatch encountered with non-stationary noise fields. Therefore, the DMA approach is not suitable for real-time situations.

Adaptive beamformers are definitely more efficient than the fixed ones in dealing with reverberation and non-stationary conditions [15, 25]. The adaptive back-to-back first-order differential microphones algorithm (ADMA) configuration is shown in Figure 2.7. This system can be used to adaptively adjust the response of the backward facing cardioid in order to track a possibly moving noise source in the back hemisphere.

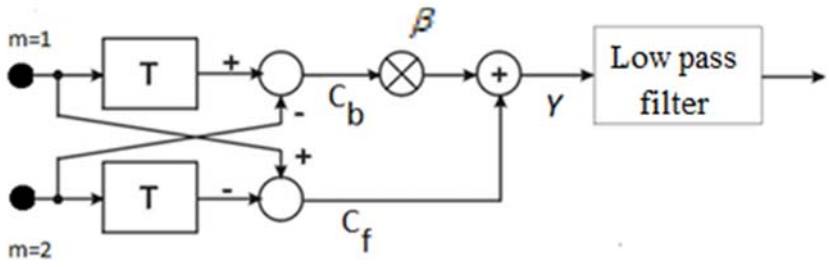


Figure 2.7 Adaptive back-to-back cardioid DMA (ADMA) configuration

In this algorithm, a back-to-back combination of arrays with cardioid responses is implemented. The cardioid response is an array system response, which has a beampattern that attenuates (or nulls) the signals coming from a certain angle. The coefficients of the ADMA system, the equation to calculate β from the desired direction angle (θ) and the equation for T (the propagation delay between the microphones) are given by the following equations [3]:

$$\underline{W}(\omega) = \begin{bmatrix} \dfrac{1+\beta e^{-j\omega T}}{1-e^{-j\omega 2T}} \\ \dfrac{-\beta - e^{-j\omega T}}{1-e^{-j\omega 2T}} \end{bmatrix} \quad \beta = \frac{1+\cos\theta}{1-\cos\theta}, \quad 90° \le \theta \le 180° \quad T = \frac{d}{c} \tag{2.28}$$

The output of the ADMA configuration is generated by multiplying the backward-cardioid signal by a factor β, and then subtracting the result from the forward-cardioid output. The factor β can be found by minimizing the mean squared error of the back-to-back configuration output using the Least Mean Square (LMS) algorithm.

For a subband domain implementation, assuming that we have a 1-D array with two microphones in the endfire direction, the forward-cardioid output is generated by delaying the rear microphone signal by $T=d/c$. After that, the delayed rear microphone signal is subtracted from the front microphone signal as the process is given by equation (2.29).

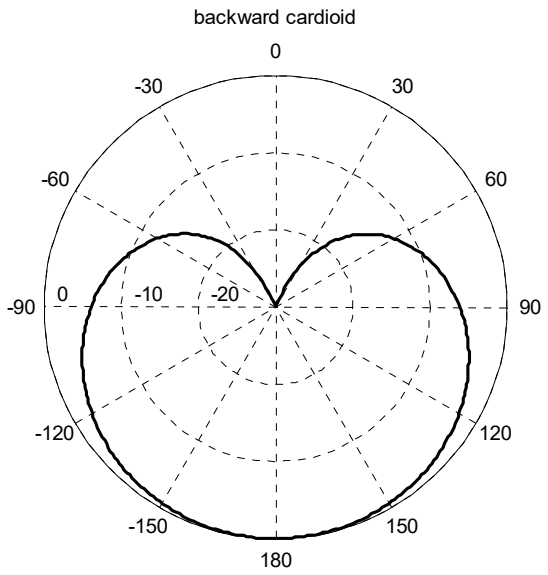
$$Cf(f,\tau) = X_1(f,\tau) - \mathrm{e}^{-j2\pi fT} X_2(f,\tau) \tag{2.29}$$

where X_1 and X_2 are a pair of microphone signals, with the (double, symmetric) null being placed in the half-hemisphere on the side of X_2, f is the center frequency of a subband, and τ is the subband time index.

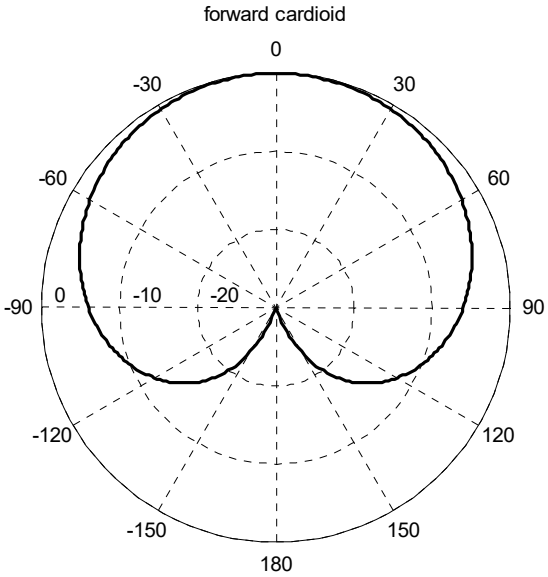
The backward-cardioid output is generated by reversing the roles of the microphones, where the front microphone signal is delayed by $T=d/c$ and then subtracted from the rear microphone signal as in equation (2.30).

$$Cb(f,\tau) = X_2(f,\tau) - \mathrm{e}^{-j2\pi fT}\, X_1(f,\tau) \tag{2.30}$$

If the desired direction in Figure 2.7 is from up to down, the beampattern for C_b is shown in Figure 2.8.a and the beampattern for C_f is shown in Figure 2.8.b.



a



b

Figure 2.8 Beampattern for a) backward-cardioid output C_b and b) forward-cardioid output C_f

The recursive equation for updating β in each frequency bin is calculated by the LMS or Normalized LMS as in equation (2.31):

$$\begin{aligned}&\beta(f,\tau)=\beta(f,\tau-1)+\mu\left(\mathrm{Re}\left\{Cf(f,\tau)Cb^{*}(f,\tau)\right\}-\beta(f,\tau-1)\left|Cb(f,\tau)\right|^{2}\right)\\&\text{with } 0<\mu<2\Big/E\left\{\left|Cb(f,\tau)\right|^{2}\right\}\\&(\text{or } 0<\mu'<2 \text{ with } \mu=\mu'\Big/E\left\{\left|Cb(f,\tau)\right|^{2}\right\})\end{aligned}\qquad(2.31)$$

where μ is the LMS step size value and μ' is the normalized LMS (NLMS) step size value. The β range should not exceed the upper and lower limits. It should be theoretically an amount between 0 ($\beta_{\min}$) and 1 ($\beta_{\max}$). The equation (2.32) shows how the limitation can be applied to β :

$$\beta(f,\tau)=\max(\beta(f,\tau),\beta_{\min})\qquad\beta(f,\tau)=\min(\beta(f,\tau),\beta_{\max})\qquad(2.32)$$

A lowpass filter is needed in the ADMA in order to avoid (relative) high frequency amplification. The global output of the system before and after the low-pass filter, respectively Y and Y' , are calculated by equation (2.33) and equation (2.34):

$$Y(f,\tau)=Cf(f,\tau)-\beta(f,\tau)Cb(f,\tau)\qquad(2.33)$$

$$Y'(f,\tau)=Y(f,\tau)\Big/\left(1-\mathrm{e}^{-j4\pi fT}+\varepsilon\right)\ \text{ or }\ Y'(f,\tau)=Y(f,\tau)\Big/\max(\left(1-\mathrm{e}^{-j4\pi fT}\right),\varepsilon)\qquad(2.34)$$

where ε is a small positive value to prevent the filtering to be ill-conditioned.

Figure 2.9 shows the beampattern when the system is adjusted to adaptively position a null with the desired direction angle is varying from 90 to 180 degrees.

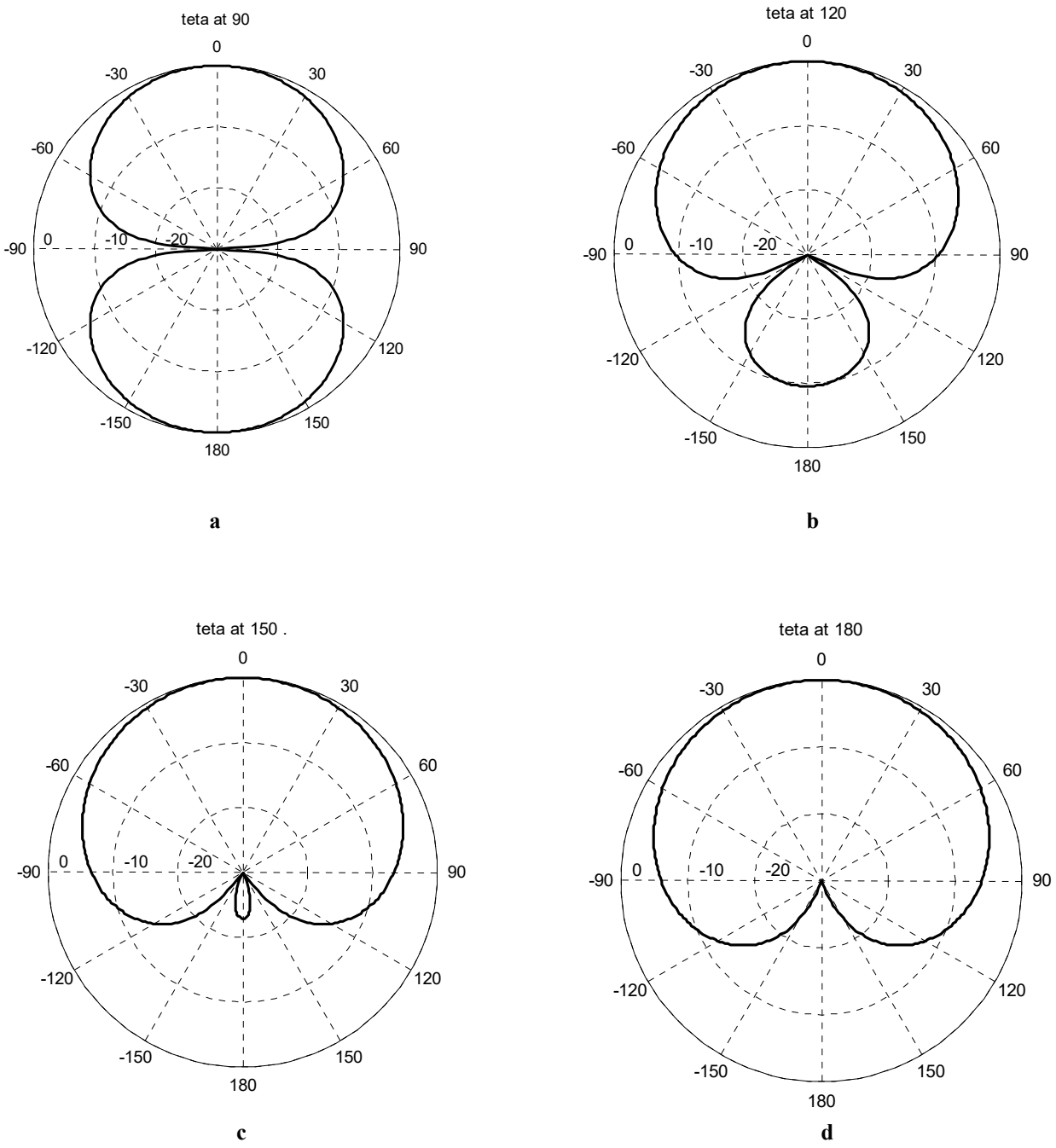


Figure 2.9 Beampattern when the system is adjusted to adaptively place a null with the desired direction angle varying from a) 90 b) 120 c) 150 and d) 180 degrees

The above smooth beampatterns are obtained under free field propagation, when the head shadow effect is not considered. Thus, in reality, ADMAs may not perform like the ones above, especially if a simple free field model is used to adjust the propagation delays T used in the ADMAs.

As discussed before, first order ADMAs are applicable for a one dimensional array consisting of only 2 microphones. Therefore, in the context of binaural hearing aids, ADMAs can be used for a pair of microphones on the same side of the head or for a pair of microphones from each side of the head. In the next chapter, we will investigate different ideas which can be implemented for binaural 2+1 and binaural 2+2 using ADMAs.

Chapter 3 **Binaural Beamforming Algorithms**

This thesis is concerned with the design of a beamformer with left and right outputs to achieve the highest interference/noise attenuation possible, when for each local output the reference microphone is the front local one. The main objective is to develop methods to improve the robustness to target DOA or target HRTF ratio estimation in binaural beamformers. The target DOA mismatch means the mismatch between the assumed target DOA (considered for the beamformer design) and the actual target DOA. The algorithms to be proposed should be robust for a range of about +-20 degrees around an assumed target location. A secondary objective is to include good target selectivity. The Generalized Sidelobe Canceller approach has been selected as the global beamforming system due to the advantages over other beamforming methodologies. As mentioned in chapter 2, the GSC methodology does not need to estimation of any correlation matrix, so it does not need any matrix inversion. Also, due to the adaptive ANC system in the second path of the GSC structure, the GSC can cope with non-stationary conditions.

In the following, several binaural beamforming architectures and several alternatives for the blocking matrix and the fixed beamformer will be suggested to deal with and control the DOA mismatch: wide-beam fixed beamformer, wide-notch blocking matrix, and adaptive null positioning for blocking matrix. A modification to improve the diffuse noise reduction of a GSC beamformer is also considered, by using a fixed beamformer tuned for diffuse noise.

3.1 GSC Structure Using the Reference Method with Target Equalization Coefficients

The basic GSC structure for the 2+1 configuration is shown in Figure 3.1:

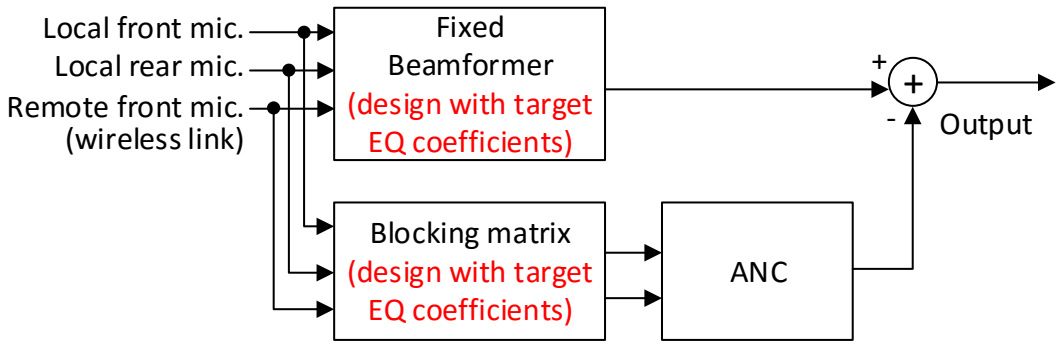


Figure 3.1 GSC structure using the reference method with target EQ coefficients (2+1 configuration shown)

Except for the fact that they can include acoustic transfer function effects (head shadow, reverberation), the fixed beamformer and the blocking matrix implemented using target equalization coefficients in the GSC are equivalent to a delay and sum fixed beamformer and a Griffiths-Jim blocking matrix (used to suppress the target signal), where only one constraint is used in the target direction. Therefore, for the fixed beamformer we use the term 'target equalized and sum' instead of delay and sum beamformer. The target equalization coefficients (EQ) design approach suffers from robustness issues for cases where the assumed target DOA (and its EQ coefficients) does not match the actual target DOA, and the performance can suffer (target leakage in blocking matrix outputs, etc.). Therefore, the idea of using a wide-notch beamformer or null-beamformer comes to mind. Designs of a wide-beam fixed beamformer and wide-notch blocking matrix beamformer(s) are considered to fulfill this purpose. Besides, considering a blocking matrix with the capacity of adaptively blocking the target signal can be beneficial to decrease the risk of DOA mismatch. Therefore, the ADMA can be also considered as a good candidate for this goal. Moreover, the 'target equalized and sum' fixed beamformer is not necessarily optimal to deal with diffuse noise, therefore we will also consider using a MVDR beamformer tuned for diffuse noise as the fixed beamformer in order attempt improved reduction of diffuse noise.

3.2 Binaural Beamforming Architectures for Hearing Aids

In this section, two different types of architectures are presented for binaural hearing aids. The first one limits the choice of beamforming methods which can be used but the second one allows more freedom in the choice of the beamformer designs. The first architecture is depicted in Figure 3.2.

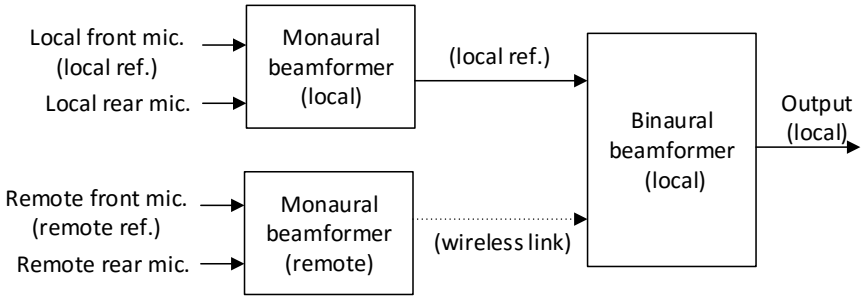


Figure 3.2 Structure of binaural beamforming preprocessed by local monaural beamformers

This structure suggests using a monaural preprocessing beamformer for each side, so that the local front microphone is the reference one and the local rear one is the second microphone to receive input signals. Therefore, two outputs are obtained which are to be considered as the new inputs of the main binaural beamformer. Then for obtaining the output using binaural beamforming for each side, the local input is considered as the reference signal and the other input transferred to the local side by the wireless link is considered as the second input of the binaural beamformer.

If the monaural beamformers are target distortion-less, the steering vector for the target direction is left unchanged for the binaural beamformer, and binaural beamformer designs which only require the steering vector of the target direction can be used (i.e., the binaural beamformer here is under just one constraint, in the desired signal direction). For example, a MVDR where the only constraint is in the target direction, or a GSC with fixed beamformer and blocking matrix based on equalization coefficients (which only depend on target steering vector), and as long as local reference target phases/magnitudes are preserved on each side with target distortionless monaural pre-processing.

But a LCMV with multiple constraints (for different directions) could not be used here for the binaural beamformer, because all the steering vectors used in the design need to be valid (for all the constraints directions, not just for one target direction), and the steering vectors of the other directions defined by the constraints are modified during monaural beamformer processing.

Alternatively, another architecture for binaural beamforming is shown in Figure 3.3, which allows more freedom in the choice of the binaural beamformer design.

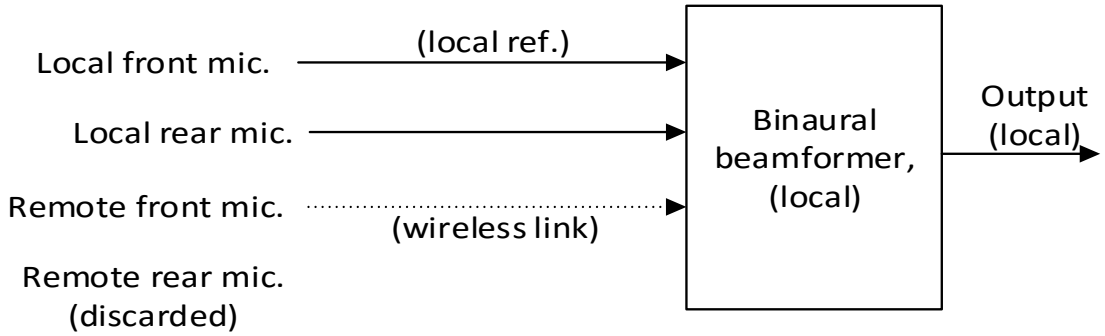


Figure 3.3 Structure of binaural beamforming without preprocessing monaural beamformers (2+1 configuration)

When no monaural beamformer pre-processing is performed, if a model for the steering vectors from different directions is known (e.g. HRTFs), then designs based on methods such as LCMV with multiple constraints become feasible. Consequently, the wide-beam fixed beamformer and wide-notch blocking matrix beamformer(s) will be designed based on the LCMV method or the constraint-based design method, either through a direct design or an indirect design using the null-space, as explained in the next sections.

3.3 Alternative Wide-Beam Fixed Beamformer and Wide-Notch Blocking Matrix Designs

The overall configuration of the proposed method for 2+1 is depicted in Figure 3.4. An offline design of wide-notch blocking matrix beamformer(s) and wide-beam fixed beamformer

will be performed based on anechoic HRTFs, whereas the beamformers will then be used for either anechoic HRTFs conditions or for reverberant recordings (more realistic conditions).

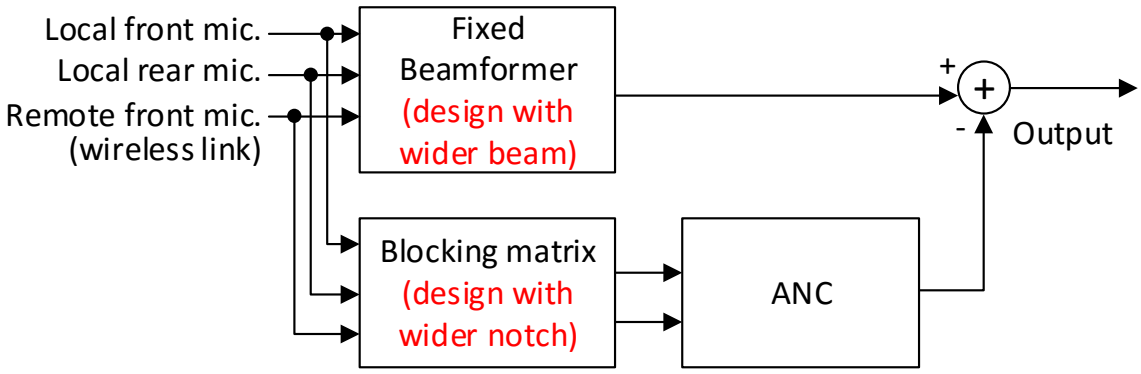


Figure 3.4 GSC structure using wide- beam fixed beamformer and wide-notch blocking matrix (2+1 configuration shown)

Achieving a wide-notch blocking matrix beamformer(s) and a wide-beam fixed beamformer would provide more robustness to mismatch between the assumed target DOA and the actual target DOA. On the other hand, there might be a potential disadvantage to this method. Another mismatch might occur, i.e., mismatch between the HRTFs used in the designs of the wide-beam and wide-notch beamformers, and the actual HRTFs (while utilizing the system in a realistic environment, which may include reverberation).

3.3.1 LCMV Design

In the general LCMV case, there can be more than one constraint, and the corresponding fixed beamformer and blocking matrix thus may have more than one constraint. The output of the fixed beamformer must give the desired response for all the constrained directions (a delay and sum beamformer cannot do this, as it is for one direction only), and the blocking matrix must give zero in ALL of its outputs for all the constrained directions (the Griffith-Jim blocking matrix cannot do this, as it is for one direction only).

Equation (3.1) describes the LCMV as:

$$\underline{W}(\omega) = (\Gamma + \mu\mathbf{I})^{-1}(\omega)C(\omega)\,(C^{H}(\omega)(\Gamma + \mu\mathbf{I})^{-1}(\omega)C(\omega))^{-1}\underline{g}(\omega) \quad (3.1)$$

Where:

$\Gamma(\omega)$ is the N x N correlation matrix of the noise, where N is the number of microphones,

$C(\omega)$ is the N x r matrix whose columns are the steering vectors of the constraint directions,

$\underline{g}(\omega)$ is the r x 1 vector containing the gain of each constraint,

$W(\omega)$ is a N x 1 vector with the beamformer coefficients,

μ is regularization factor used for diagonal loading.

For N microphones, the LCMV can have up to N-1 constraints.

The potential advantage of LCMV is wider notch in the target direction for the blocking matrix. On the other hand, there might be a mismatch between the HRTFs design and the reverberant conditions (although the mismatch may be smaller for near-frontal targets). Also, the number of blocking matrix output channels that can be generated with this approach may be limited (for example one LCMV design only produces one blocking matrix output, as opposed to the Griffiths-Jim approach which produces N-1 outputs).

3.3.2 Constraint-Based Design

The constraint-based design is similar to the LCMV but without the minimization of the trace of $\Gamma(\omega)$. For N microphones the constraint-based design can have up to N constraints.

Based on chapter 2, the condition for the optimization problem of multiple constraints design is given by equation (3.2):

$$C^{H}(\omega)\underline{W}(\omega) = \underline{g}(\omega) \quad (3.2)$$

The solution is equivalent to a pseudo-inverse solution for the fixed beamformer and a null-space search for the blocking matrix corresponding to the constraints of the fixed beamformer:

$$\underline{W}(\omega) = C(\omega)\,(C^H(\omega)C(\omega))^{-1}\,\underline{g}(\omega) \tag{3.3}$$

The advantage over the LCMV is that we can have N constraints instead of N-1. For a system with N microphones, from these equations it is theoretically possible for a blocking matrix to cancel the signals from the r constraint directions in each of its M-r outputs. So theoretically we can still generate N-r linearly independent outputs in the blocking matrix. The disadvantage is that the design of the fixed beamformer will not be very robust to white noise or diffuse noise, because there is no correlation matrix for the noise included in the design. But for the blocking matrix it is not such a serious concern.

3.3.3 Null-space Design

For both the LCMV design and the constraint-based design, it is also possible to use the null-space approach in order to design blocking matrix beamformers $V(\omega)$ given by equation (3.4):

$$[V(\omega), \lambda(\omega), U(\omega)] = SVD(\hat{C}(\omega)\hat{C}^H(\omega)) \tag{3.4}$$

where $\hat{C}(\omega)$ contains the steering vectors from $C(\omega)$ with non-zero corresponding gains in $\underline{g}(\omega)$. If the rank of $\hat{C}(\omega)$ is r (number of non-zero constraints), theoretically the last N-r columns of $V(\omega)$ are null-beamformers, i.e., beamformers with nulls in the directions of the non-zero constraints. Therefore, this approach can be used for the design of blocking matrix beamformers.

For the null-space computation, one small change (i.e., using matrix $\hat{C}(\omega)\hat{C}^H(\omega)$ instead of C in the SVD decomposition, which theoretically has the same eigenvectors as the constraint matrix C) made a significant practical improvement in the performance, due to numerical

sensitivity issues. Another factor to consider is the number of non-zero gain constraints that were used in the fixed beamformer design, since less constraints means more output signals in the blocking matrix, theoretically. Note also that the fixed beamformer to be used for the null-space design of the blocking matrix doesn't have to be the same as the fixed beamformer that we will use in the final GSC (which for example could be a "target-equalized and sum", etc.).

3.4 Alternative Fixed Beamformer Tuned for Diffuse Noise Reduction

MVDR is known to provide the optimal diffuse noise reduction [26, 27]. The GSC structure using an MVDR fixed beamformer tuned for diffuse noise is shown in Figure 3.5.

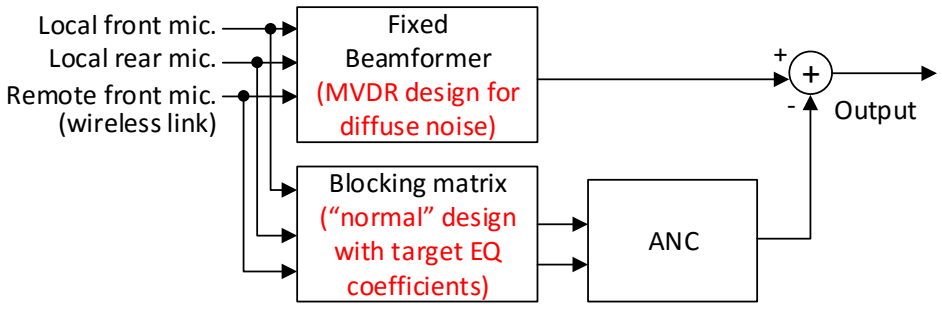


Figure 3.5 GSC structure using MVDR fixed beamformer (2+1 configuration shown)

In theory, as long as the fixed beamformer meets the constraints of a GSC/LCMV design (e.g., gain in some directions), and as long as the blocking matrix has nulls in the direction of the constraints, the choice of the fixed beamformer does not have an impact on the overall performance. However, in practice, with non-ideal fixed beamformers, blocking matrices and ANC units, it is worth investigating if a fixed beamformer using a MVDR design for diffuse noise could improve the performance in terms of diffuse-like background noise reduction.

To compute the diffuse noise correlation matrix there are three possibilities:

- It can be computed from noise+interference signals, which is not practical in many situations. However, in this case we are interested in the diffuse noise-only correlation matrix, without interferers, and it is more realistic to consider that such a correlation matrix can be estimated.
- It can be computed from noisy signals (including target components), because the MVDR design automatically attempts to maximize array gain or minimize the power of the output. However, as reported in the literature, the problem with this approach is that the performance becomes less robust in the case of target DOA mismatch.
- The diffuse noise correlation matrix can be estimated by averaging the HRTFs from all directions, however this approach can suffer from mismatch if the HRTFs used for the design don't correspond to the actual HRTFs (possibly with reverberation) found in practice.

If the MVDR noise correlation matrix is estimated based on HRTFs, the method is not applicable for a binaural beamformer with monaural pre-processing architecture (because monaural pre-processing changes the HRTFs in non-target directions). Thus, it is only applicable for the 2+1 configuration with no pre-processing. But if the diffuse noise correlation matrix is estimated from the microphone signals, the method can also be applied for configurations with monaural pre-processing. In our simulations, we used the diffuse noise field recordings (background noise) to computed the correlation matrix required in the MVDR fixed beamformer. The design was performed offline and we then loaded the diffuse-noise tuned fixed beamformer coefficients for different noise scenarios, including the presence of interferers. As for the target steering vector required in the MVDR, we can either compute them directly from the target signals (reverberant recordings), which is the ideal scenario, or we can use pre-measured HRTFs (with mismatch when reverberant conditions are found). For the section that attempts to address the effectiveness in diffuse noise reduction, we have used the target steering vector obtained directly from the target signals.

3.5 Alternative Blocking Matrix Based on ADMA

Alternatively, monaural and binaural adaptive back-to-back 1st order differential microphones (ADMAs) can be used in target cancelling mode for the blocking matrix. The GSC structure using ADMA blocking matrix is shown in figure 3.6

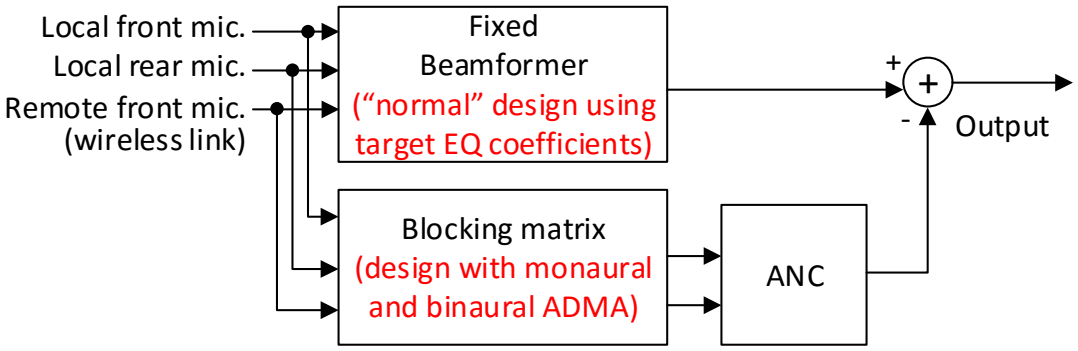


Figure 3.6 GSC structure using MVDR blocking matrix (2+1 configuration shown)

According to chapter 2, the equations below can describe the ADMA step by step:

$$Cf(f,\tau) = X_1(f,\tau) - \mathrm{e}^{-j2\pi fT}\, X_2(f,\tau) \tag{3.5}$$

$$Cb(f,\tau) = X_2(f,\tau) - \mathrm{e}^{-j2\pi fT}\, X_1(f,\tau) \tag{3.6}$$

$$Y(f,\tau) = Cf(f,\tau) - \beta(f,\tau)Cb(f,\tau) \tag{3.7}$$

$$Y'(f,\tau) = Y(f,\tau)\Big/\left(1 - \mathrm{e}^{-j4\pi fT} + \varepsilon\right) \text{ or } Y'(f,\tau) = Y(f,\tau)\Big/\max\left(\left(1 - \mathrm{e}^{-j4\pi fT}\right),\varepsilon\right) \tag{3.8}$$

where X_1 and X_2 are a pair of microphone signals, with the (double, symmetric) null being placed in the half-hemisphere on the side of X_2, f is the center frequency of a subband, τ is the subband time index, and T is the propagation delay between the microphones.

It is possible to adaptively adjust β values with an LMS or normalized LMS approach, leading to an adaptive back-to-back differential microphone array (ADMA). For the adaptation, the cost function $J(\beta)$ is $|Y(f,\tau)|^2$, minimizing the total power of $Y(f,\tau)$ as a function of $\beta(f,\tau)$. Minimizing $|Y(f,\tau)|^2$ attempts to reduce the sources in the back hemisphere. The equations below show the derivation of the recursive equation for adjusting $\beta(f,\tau)$.

$$J(\beta)=\left|Y(f,\tau)\right|^2=Y(f,\tau)Y^*(f,\tau)=(C_f(f,\tau)-\beta(f,\tau)C_b(f,\tau))(C_f^*(f,\tau)-\beta(f,\tau)C_b^*(f,\tau)) \tag{3.9}$$

$$J(\beta)=\left|C_f(f,\tau)\right|^2=-2\beta(f,\tau)\,\mathrm{Re}\left\{C_b^*(f,\tau)C_f(f,\tau)\right\}+\beta(f,\tau)^2\left|C_b(f,\tau)\right|^2 \tag{3.10}$$

$$\frac{\partial J(\beta)}{\partial\beta(f,\tau)}=-2\,\mathrm{Re}\left\{C_b^*(f,\tau)C_f(f,\tau)\right\}+2\beta(f,\tau)\left|C_b(f,\tau)\right|^2 \tag{3.11}$$

$$\beta(f,\tau+1)=\beta(f,\tau)-\mu\frac{\partial J(\beta)}{\partial\beta(f,\tau)} \tag{3.12}$$

Finally the recursive equation for adjusting β values is given by equation (3.13):

$$\begin{aligned}&\beta(f,\tau+1)=\beta(f,\tau)+2\mu\left(\mathrm{Re}\left\{Cf(f,\tau)Cb^*(f,\tau)\right\}-\beta(f,\tau)\left|Cb(f,\tau)\right|^2\right)\\&\text{with } 0<\mu<2\Big/E\left\{\left|Cb(f,\tau)\right|^2\right\}\ \ for\ LMS\quad(\text{or } 0<\mu'<2 \text{ with } \mu=\mu'\Big/E\left\{\left|Cb(f,\tau)\right|^2\right\}\ \ for\ NLMS)\\&\beta(f,\tau)=\max(\beta(f,\tau),\beta_{\min})\qquad\beta(f,\tau)=\min(\beta(f,\tau),\beta_{\max})\end{aligned} \tag{3.13}$$

Several parameters which can affect the performance of the ADMA blocking matrix need to be adjusted:

- NLMS step size (μ): reducing μ from its theoretical maximum possible value ended up with better performance

- ADMA regularization/diagonal loading (ε) when doing low-pass filter: by trial and error, increasing the regularization value was found to improve the performance.

- ADMA initial β: the monaural ADMA beamformer in target cancelling mode places a double symmetric notch in the front half-hemisphere. So for obtaining a notch in the endfire direction (for two monaural microphones on the same side of the head), we can initialize β to 0. The binaural

beamformer places a double, symmetric notch in either the left or the right half-hemisphere, so we can initialize β to 1 for a notch in the broadside direction (for two binaural microphones on each side of the head). Some broad information on the target direction is thus assumed, i.e., left or right half-hemisphere, to determine on which side the binaural ADMA should attempt to perform target cancelation.

- ADMA range of allowed β: theoretically between 0 and 1, this will be discussed in detail in chapter 5.

Using an ADMA blocking matrix avoids the previously mentioned problem of a possible HRTFs mismatch (the mismatch between the HRTFs used in design and the actual HRTFs, e.g. reverberant). Also this approach does not suffer from DOA mismatch due to the capacity of automatic and adaptive null positioning, although the range of allowed notch locations would need to be specified. However, the presence of interferers in the regions considered for positioning nulls might prevent the correct convergence of the null towards the target direction. This can possibly be mitigated by the initial conditions for the null location and the range of null locations allowed in the adaptive algorithm.

3.5.1 Monaural and Binaural ADMAs in Target Cancelling Mode

In order to increase the overall system performance, it can be highly beneficial to have more than one linearly independent blocking matrix outputs to provide more information about the noise signals to the ANC system.

For the 4-microphone 2+2 binaural configuration we ideally want to have 3 blocking matrix outputs on each side (3 independent linear combinations of the interferers). For the binaural beamformer on each side, theoretically we can choose 3 or more ADMAs from the following options:

1. Using a monaural ADMA for the right or the left side (2 options)
2. Using a binaural 1+1 ADMA for the front or the back microphones (2 options)

3. Using a binaural 1+1 ADMA for the front microphone of one side and the back microphone of the other side (2 options).

Although this is 6 different combinations in principle, several of those combinations are likely to be highly redundant (not linearly independent), and in practice the proper choice would be to use the two monaural ADMAs as well as one binaural ADMA (using the front microphones).

For the 3-microphones 2+1 binaural configuration, we normally want to have 2 blocking matrix outputs on each side (2 independent linear combinations of the interferers). For the binaural beamformer on each side, we could theoretically choose 2 or more ADMAs from the following options (they are shown in Figure 3.7 for right side binaural beamforming, with the green microphones selected for the ADMA processing):

1. Using a monaural ADMA for the local side (1 option)
2. Using a binaural 1+1 ADMA for the front microphones (1 option)
3. Using a binaural 1+1 ADMA for the local front microphone and the remote rear microphone.

However, here as well the 2nd and 3rd cases are likely to be highly redundant, and in practice the proper choice would be to use the one monaural ADMAs as well as one binaural ADMA (using the front microphones).

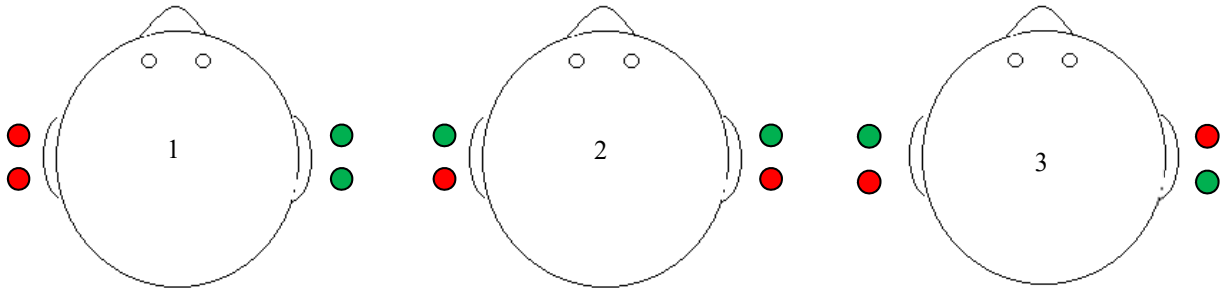


Figure 3.7 The microphones involved in binaural or monaural ADMAs for 2+1 config. (the ones in green)

3.5.2 Restriction of β Values for Monaural and Binaural Scenarios

With the ADMAs operating in target cancelling mode, if target leakage happens at the output of one of the monaural or binaural ADMA beamformers it can be troublesome. Therefore, if there is an interferer which can prevent one of the ADMAs from achieving the correct convergence towards the null in the target direction, this will degrade the final performance of the GSC beamformer. One possible solution to this issue is to restrict the β values so that the null can only be positioned near a direction where the target is estimated to be. Otherwise, for the case with more than 2 sources in the half-hemisphere (frontal half-hemisphere for monaural ADMA, left or right half-hemisphere for binaural ADMA), convergence to the target direction cannot be guaranteed, and target leakage may occur. For example, if we assume a near frontal target, a β range 0.0 to 0.2 can be used for the monaural ADMA and a β range 0.8 to 1.0 can be used for the binaural ADMA. Note that the width of the range can also be wider: for example, β range from -0.1 to 0.3 can be considered for the monaural ADMA and 0.7 to 1.1 for the binaural ADMA. If the target is not assumed to be frontal, then the range of possible β values needs to be revised accordingly.

Chapter 4 Experimental Setup and Performance Measurements

4.1 Experimental Setup

In this thesis, the algorithms are implemented using an analysis/synthesis filter bank, in order to reduce algorithmic complexity while also keeping the group delay introduced by the processing at a low level. For evaluation purposes, 10 seconds recordings were used sampled at 24 kHz sampling rate with some reverberant speech signals as target signal and interferers, as well as some background diffuse-like noise. It is assumed for this work that the signals from all the considered microphones are available for processing (no wireless link transmission limitations such as limited bandwidth, jitter, delay, signal coding and quantization noise, etc.). The algorithms have been implemented in a Matlab software environment. Some standard metrics have been used for evaluation (e.g. array SNR gain, DI (directivity index), signal distortion).

For the methods using the target equalization coefficients, filters with three coefficients per subband and one sample of causality delay were used. As previously mentioned, the target equalization coefficients were computed directly from the target signal components at the microphones. Note that each sample delay in the subband processing at the downsampled rate of 1 kHz leads to a 1 ms delay. The equalization coefficients were computed from the ratio of the cross power spectrum density between the reference microphone target signal and the other microphones' target signal, over the auto power spectrum density of the other microphones' target signal.

In the ANC system of the GSC, the fixed beamformer output was delayed by one sample causality delay. The ANC used three coefficients per subband. These ANC coefficients were estimated by a least-squares algorithm over the whole length of the signals. This is a non-adaptive implementation of the ANC, we have chosen this setup to reduce the effect of the ANC (and its

tuning) on the overall performance, so that the impact of the different fixed beamformers and blocking matrix tested could be more easily evaluated.

Also, when combining different fixed beamformers and different blocking matrix beamformers, care was taken to insure that the signals were time aligned, for example if a causality delay was used in the beamformer of one path and not in the beamformer of the other path.

4.2 Performance Measurements

Various metrics and criteria used for evaluating the proposed algorithms are covered below. They are mainly used to measure the noise reduction and speech distortion. Also we can see visually how a beamformer can steer the null or a beam in a certain direction. These criteria can be used for the performance evaluation of both the right and left side binaural beamformers.

4.2.1 Beampattern

The beampattern is a very effective and direct way to show the directivity of the beamformers. Each beamformer has a pattern of directional sensitivity, i.e., it has different sensitivities for sounds arriving from different directions. The beampattern or directivity pattern describes the sensitivity of the beamformer to a plane wave (source signal) impinging on the array from the direction φ. The beampattern concept has been explained previously in chapter 2.

4.2.2 Directivity Index

The Directivity Index (DI) is an expression in decibels of the directionality of a sound source. As described in chapter 2, a diffuse noise field is one in which noise of equal energy propagates in all directions at all times. In the case of a diffuse noise field, the array gain is also known as the DI. The DI can be considered as a performance evaluator for diffuse noise reduction. In chapter 5, the DI figures will be shown for comparing the delay and sum fixed beamformer (in free field) and the "target-align and sum" fixed beamformer (considering head shadow effect) with

MVDR tuned for diffuse noise fixed beamformer designs, in both free field and with head shadow effect.

4.2.3 Signal to Noise Ratio Gain (SNR-gain)

The Signal to Noise Ratio gain (SNR-gain) is defined as the improvement in the SNR that the system introduces. The SNR-gain is computed by taking the difference between the output SNR and the input SNR, and it is often measured in dB. The input SNR is computed by taking the ratio of the power of the target components on the reference microphone over the power of the sum of the interferers and the background (diffuse) noise components at the reference microphone signal. The output SNR is computed by taking the ratio of the power of the target components over the power of the sum of the interferers and the background noise components at the beamformer output. For the input SNR, the signal components at the front right or front left microphone can be used.

$$SNR_{in} = 10\log\left(\frac{\text{power of the target components at reference microphone}}{\text{power of sum of the interferers and the background noise components at ref. microphone}}\right)$$

$$= 10\log\left(\frac{P_{x_{ref}}(f,t)}{P_{i_{ref}+n_{ref}}(f,t)}\right) \tag{4.1}$$

$$SNR_{out} = 10\log\left(\frac{\text{power of the target components at the beamformer's output}}{\text{power of sum of the interferers and the diffuse noise components at beamformer's output}}\right)$$

$$= 10\log\left(\frac{P_{x_{out}}(f,t)}{P_{i_{out}+n_{out}}(f,t)}\right) \tag{4.2}$$

$$SNR_{gain}(dB) = SNR_{in}(dB) - SNR_{out}(dB) \tag{4.3}$$

where f is a frequency index and t is time frame index.

For simulation purposes, the SNR-gain was computed with power spectrum densities using a Welch method with a Hamming window of size 2048 with 50% overlapping and FFTs of size 8192, with a sampling rate of 24 kHz (used throughout this thesis).

4.2.4 Signal to Distortion Ratio (SDR)

The Signal to Distortion Ratio (SDR) measures how the target components in the beamformer output are distorted compared with target components at the reference microphone. The SDR is computed by taking the power of the difference between the target component at the output channel and the target component signal at the reference microphone. After that, the computed power difference is normalized by the power of the target component at the reference microphone. The SDR is often measured in dB.

The frequency dependent measurement of SDR is given in the following equations:

$$x_{distorted} = x_{out} - x_{ref} \tag{4.4}$$

$$SDR(dB) = 10\log(\frac{P_{x_{distorted}}}{P_{x_{ref}}}) \tag{4.5}$$

Where;

$P_{x_{ref}}$: Power spectrum density for the input target components at the reference microphone.

$P_{x_{distorted}}$: Power spectrum density for the difference between the output target components at the output channel and the target components at the reference microphone.

For simulation purposes, the SDR was calculated with power spectrum densities calculated using a Welch method with a Hamming window of size 2048, with 50% overlapping and FFTs of size 8192.

4.2.5 Signal to Interferers Ratio Gain (SIR-gain)

The Signal to Interferers Ratio is defined as a kind of SNR where the system only considers the interferers as the noise signals. The SIR-gain is computed by taking the difference between the output SIR and the input SIR (in dB). This is very similar to SNR-gain, but it uses only interferers, not the other noise signals.

$$SIR_{in} = 10\log\left(\frac{\text{power of the target components at reference microphone}}{\text{power of the sum of the interferers components at reference microphone}}\right)$$

$$= 10\log\left(\frac{P_{x_{ref}}(f,t)}{P_{i_{ref}}(f,t)}\right) \tag{4.6}$$

$$SIR_{out} = 10\log\left(\frac{\text{power of the target components at the beamformer's output}}{\text{power of the sum of the interferers components at the beamformer's output}}\right)$$

$$= 10\log\left(\frac{P_{x_{out}}(f,t)}{P_{i_{out}}(f,t)}\right) \tag{4.7}$$

$$SIR_{gain}(dB) = SIR_{in}(dB) - SIR_{out}(dB) \tag{4.8}$$

4.2.6 Background Diffuse Noise SNR- gain

It is sometimes useful to analyse separately the performance of a beamformer for the diffuse-noise component of a noise field, i.e., without the interferers' noise components. The Signal to Background Noise Ratio is defined as a kind of SNR in which only the diffuse background noise is considered as the noise signal. The SNR-gain for diffuse noise is computed by taking the difference between the background noise output SNR and input SNR (in dB).

$$SNR_{in}(backgr.\ noise) = 10\log(\frac{\text{power of the target components at reference microphone}}{\text{power of the background noise components at reference microphone}})$$

$$= 10\log(\frac{P_{x_{ref}}(f,t)}{P_{n_{ref}}(f,t)}) \tag{4.9}$$

$$SNR_{out}(background\ noise) = 10\log(\frac{\text{power of the target components at the beamformer's output}}{\text{power of the background noise components at the beamformer's output}})$$

$$= 10\log(\frac{P_{x_{out}}(f,t)}{P_{n_{out}}(f,t)}) \tag{4.10}$$

$$SNR_{gain}(background\ noise) = SNR_{in}(background\ noise) - SNR_{out}(background\ noise) \tag{4.11}$$

For hearing aids (and for hearing in general) there are frequencies which are more important than other frequencies due to the human auditory system characteristics. Therefore, having a visual overview of a frequency dependent SNR–gain, SDR, and other metrics can be beneficial to deeply evaluate and investigate the performance of the proposed algorithms. In chapter 5 some figures will thus be presented to show the frequency dependent measurements.

Chapter 5 Simulation Results for Binaural Beamformers

The objective of the simulation results in this chapter is to compare the performance of different binaural beamforming configurations. Throughout the chapter, the GSC system will be designed using different fixed beamformer and blocking matrix algorithms.

5.1 Beampatterns and Directivity Index Results

In this section the beampatterns of every algorithm considered for the blocking matrix and fixed beamformer of the GSC system are provided for some specific frequencies. The beampatterns are presented for the frequency range which is more important for the human hearing system. These beampatterns provide the necessary information to assess the widths of the beams and the nulls. It should be noted that the beamformers in this section have been designed to steer the notch or the beams for the frontal direction (0-degree azimuth). The beampatterns listed below will be shown in this section for both the free field conditions and the condition with the head shadow effect (using HRTFs measured from a dummy head with artificial ears), and for both the right side and left side.

- Beampatterns for blocking matrix:
  - target-equalized Griffiths-Jim for 1+1 binaural beamformer
  - 2+1 LCMV design (with null-space)
  - 2+1 constraint-based design (without null-space)
  - 2+1 constraint-based design (with null-space)
- Beampatterns for fixed beamformer:
  - “Target equalize and sum” for 1+1 binaural beamformer
  - 2+1 “Target equalize and sum” beamformer
  - 2+1 LCMV design
  - 2+1 constraint-based design
  - 2+1 MVDR design for diffuse noise.

5.1.1 Beampatterns for Blocking Matrix Designs

First, the beampatterns for blocking matrix designs are presented. The target equalized Griffiths-Jim blocking matrix which has been described in chapter 4 is the reference blocking matrix in the simulations. The 1+1 configuration with pre-processing on both sides is also a reference configuration, as it is currently being used in binaural hearing aids. Note that this is a 4-microphone system, i.e., the 1+1 name just comes from the fact that the binaural beamformer receives 1 signal from each side. Also, note that the Griffiths-Jim blocking matrix for the 2+1 configuration would include the same beamformer as the Griffiths-Jim blocking matrix for the 1+1 configuration (plus an additional beamformer), therefore if the Griffiths-Jim blocking matrix for the 1+1 configuration suffers from a narrow beam or target leakage, the Griffiths-Jim blocking matrix for the 2+1 configuration will also suffer from the same problem. This could also be said for the Griffiths-Jim blocking matrix with the 2+2 configuration. The polar beampatterns for the target-equalized Griffiths-Jim 1+1 binaural beamformer in free field (left side and right side), with HRTFs (left side) and HRTFs (right side) are shown respectively in Figure 5.1, Figure 5.2 and Figure 5.3.

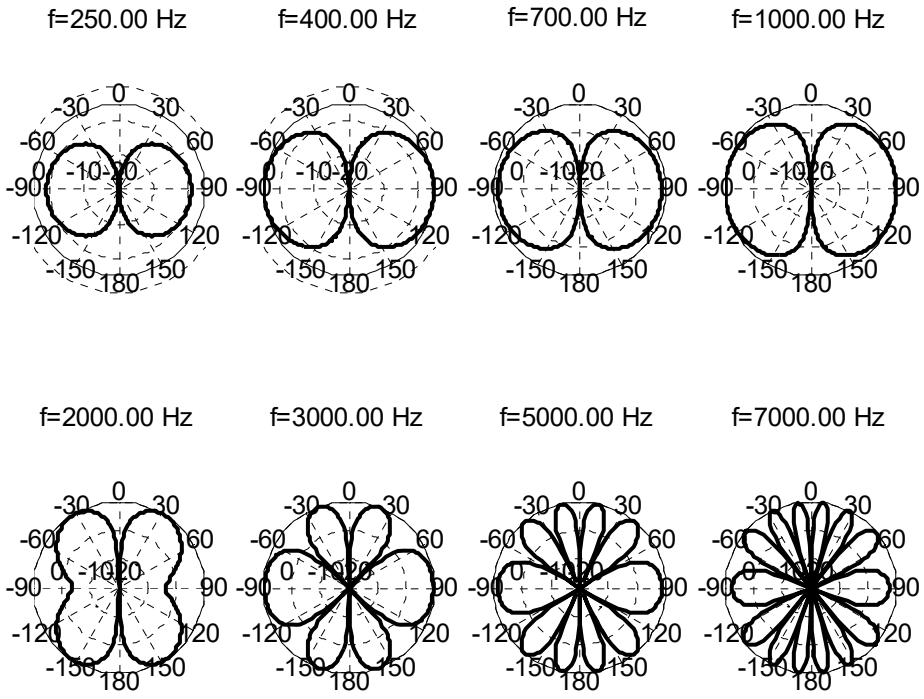


Figure 5.1 Beampatterns for blocking matrix design, target-equalized Griffiths-Jim, 1+1 binaural configuration (free field, left-side or right-side reference)

In Figure 5.1, the notch becomes more narrow starting at 1 kHz (e.g. if a 20 dB attenuation is required to avoid target leakage).

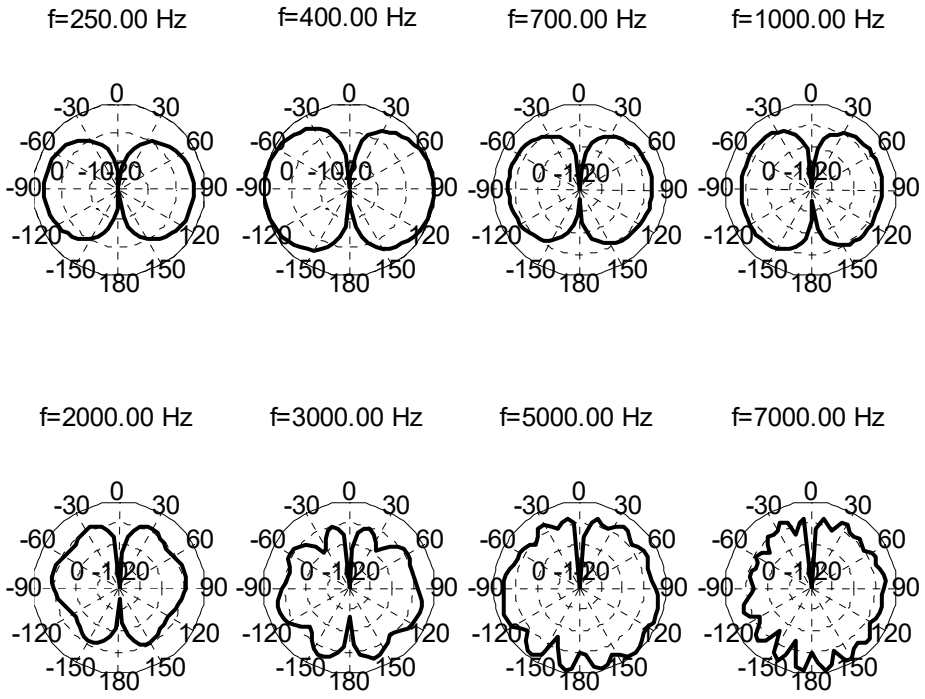


Figure 5.2 Beampatterns of blocking matrix design, target-equalized Griffiths-Jim, 1+1 binaural configuration (HRTF, left-side reference)

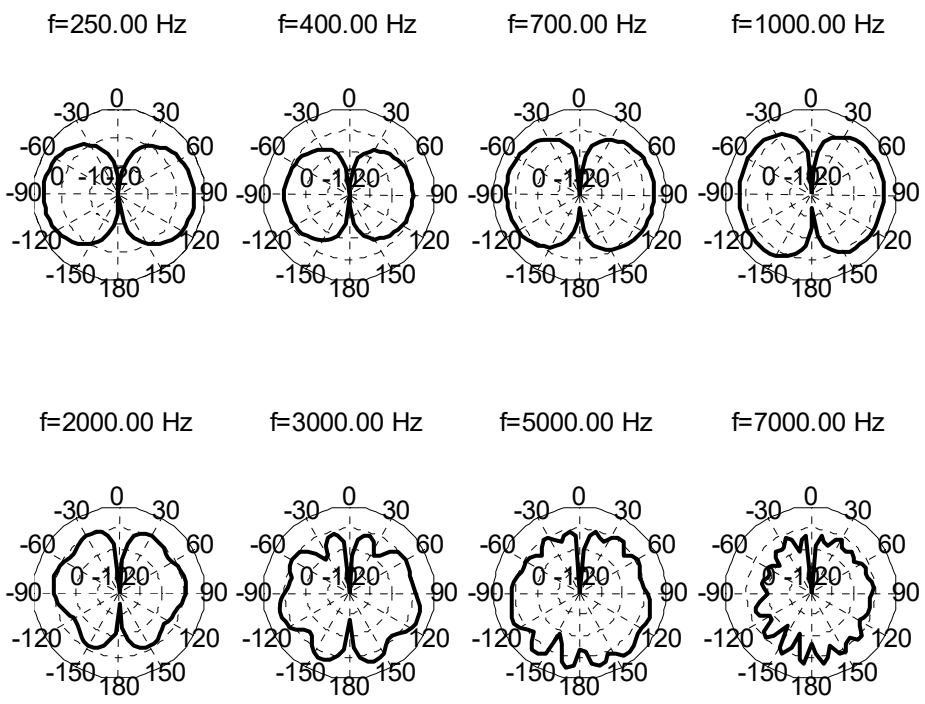


Figure 5.3 Beampatterns of blocking matrix design, target-equalized Griffiths-Jim, 1+1 binaural configuration (HRTF, right-side reference)

It can be easily seen that with the head shadow effect (using HRTFs) for either side, similar results are obtained for the 1+1 configuration in terms of the notch which becomes progressively more narrow starting around 1 kHz. Therefore, in the case of target DOA mismatch, this can create some performance degradation due to target signal leakage at the output of blocking matrix. Thus, it is worth to work on the proposed wide-notch blocking matrices with other configurations than the 1+1 configuration. For this purpose, at first, the LCMV with null-space design for the 2+1 configuration is considered to evaluate how wide the notch will be in comparison with the target-equalized Griffiths-Jim in 1+1 configuration.

Consider the following constraint vector and the directions corresponding to each constraint: g=[1.0; 1.0], theta_constraints=[10,350].

In the null-space of this design, there should be null constraints at 10 and 350 degrees, so a wider notch is expected at a minimum of 10 degrees around the frontal direction. The polar beampatterns for the LCMV blocking matrix using null-space design with 2+1 configuration binaural beamformer in free field (left side or right side), with HRTFs (left side) and HRTFs (right side) are shown respectively in Figure 5.4, Figure 5.5 and Figure 5.6.

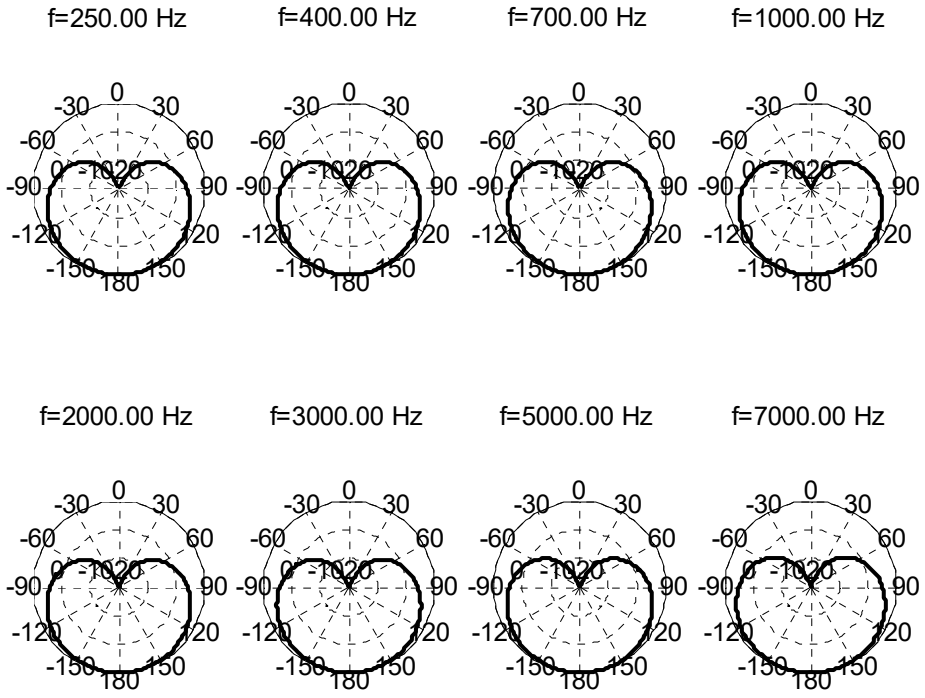


Figure 5.4 Beampatterns of blocking matrix design, LCMV with null-space, 2+1 configuration (free field, left-side or right-side reference) when g=[1.0; 1.0] and theta_constraints=[10,350]

As expected with the LCMV design, a wider notch can be achieved, even above 1 kHz, so that in the case of DOA mismatch the blocking matrix is less likely to have target leakage in the output in comparison with the target-equalized Griffiths-Jim. Therefore, the GSC system can be more robust to DOA mismatch.

Unfortunately, only designs with either narrow (like the previous case) or fairly wide notches (like here) can be achieved. It seems that either the effect of the close distance between the microphones on the same side is dominant in the response (which leads to wide notches), or the effect of the large distance between two microphones from different sides is dominant in the response (which leads to narrow notches).

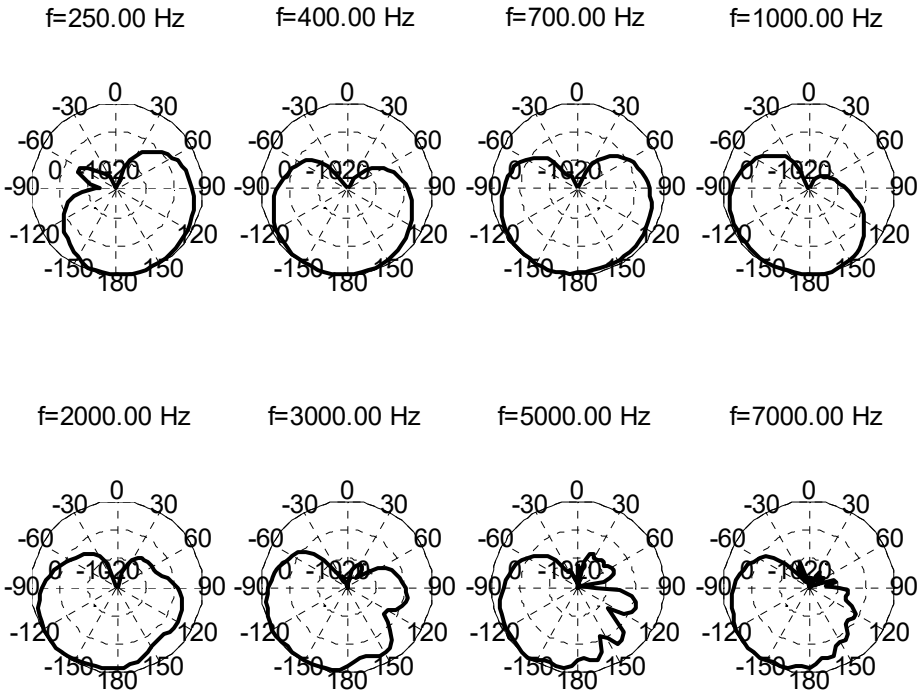


Figure 5.5 Beampatterns of blocking matrix design, LCMV with null-space, 2+1 configuration (HRTFs, left-side reference) when g=[1.0; 1.0] and theta_constraints=[10,350]

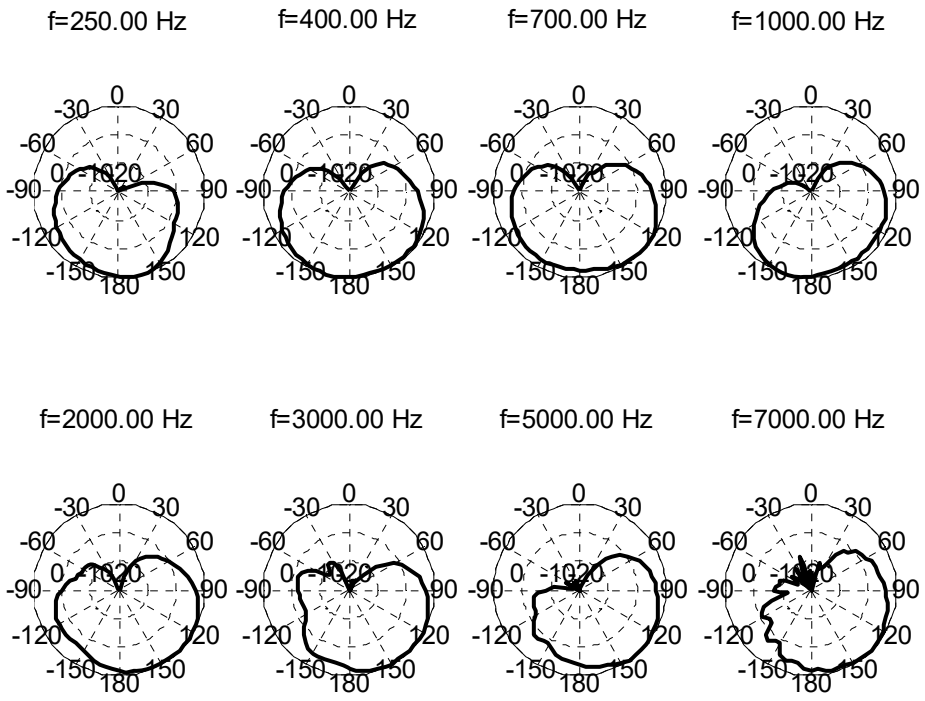


Figure 5.6 Beampatterns of blocking matrix design, LCMV with null-space, 2+1 configuration (HRTFs, right-side reference) when g=[1.0; 1.0] and theta_constraints=[10,350]

For the HRTFs conditions we obtained similar results to the free field conditions in terms of a wider notch for the LCMV with null-space design, although there was more attenuation outside the desired notch for frequencies above 3 kHz.

Another wide notch algorithm is the constraint-based design without null-space for the 2+1 configuration. Using a constraint vector g=[0.0; 0.0; 1.0] and the directions corresponding to each constraint as theta_constraints=[10,350,180], it is expected that a wider notch can be obtained at least 10 degrees around the frontal direction, as well as a gain of unity for the back direction. The polar beampatterns for the constraint-based design without null-space for 2+1 configuration binaural beamformer in free field (left side or right side), with HRTFs (left side) and HRTFs (right side) are shown respectively in Figure 5.7, Figure 5.8 and Figure 5.9.

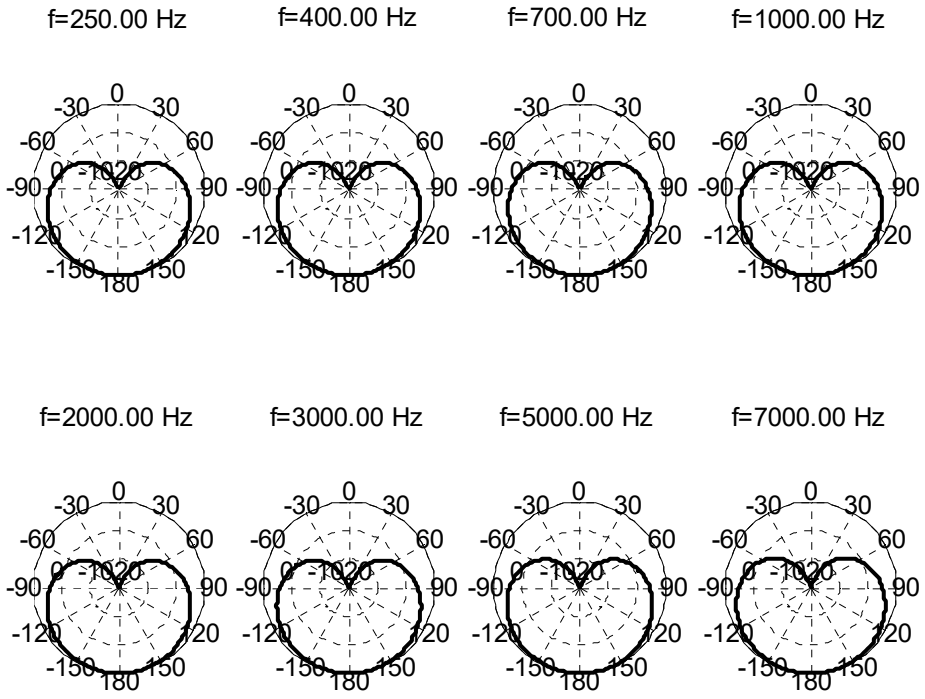


Figure 5.7 Beampatterns for blocking matrix design, constraint-based without null-space, 2+1 configuration (free field, left-side or right-side reference) when *g*=[0.0; 0.0; 1.0] and theta_constraints=[10,350,180]

For the constraint-based design without null-space in free field conditions, nearly identical results to the previous design with LCMV using null-space are obtained, i.e., the notch remains wide even at high frequencies.

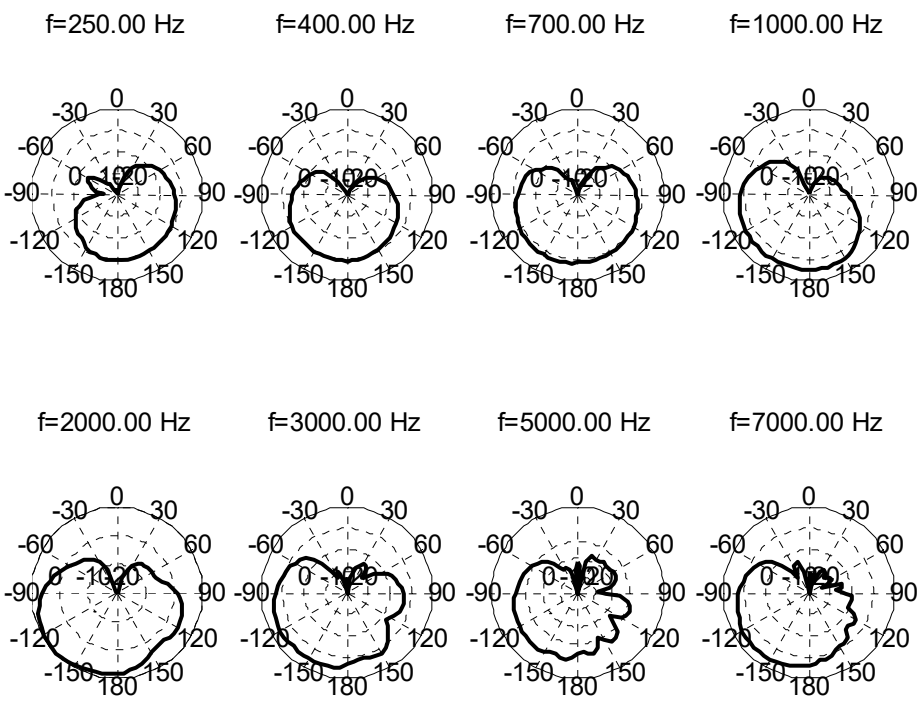


Figure 5.8 Beampatterns for blocking matrix design, constraint-based without null-space, 2+1 configuration (HRTFs, left-side reference) when *g*=[0.0; 0.0; 1.0] and theta_constraints=[10,350,180]

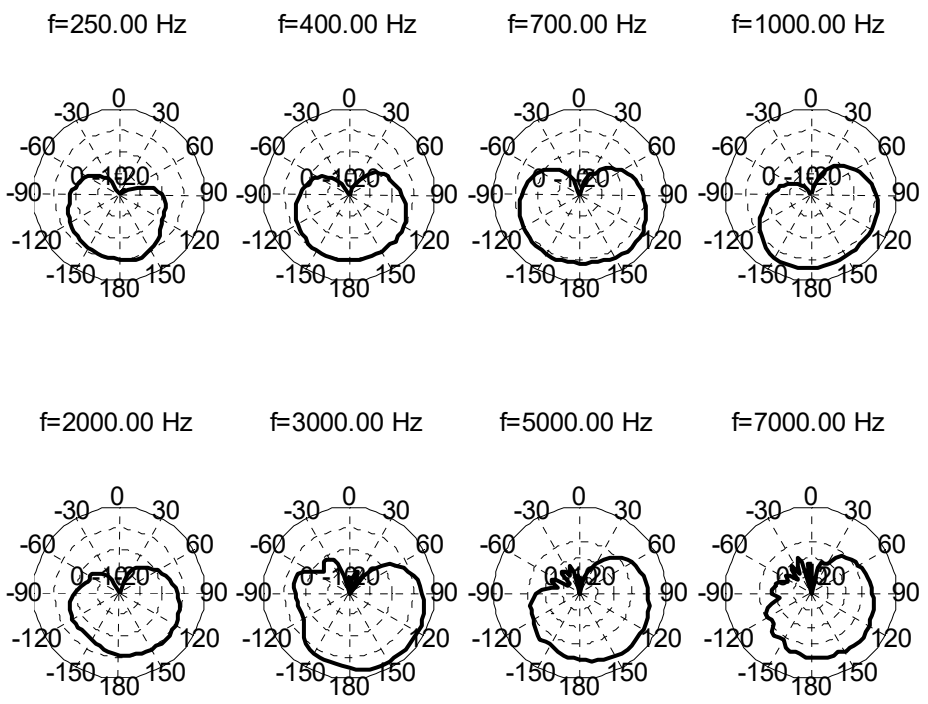


Figure 5.9 Beampatterns for blocking matrix design, constraint-based without null-space, 2+1 configuration (HRTFs, right-side reference) when *g*=[0.0; 0.0; 1.0] and theta_constraints=[10,350,180]

As the above beampatterns show, the constraint-based design without null-space for HRTFs conditions produces similar results to the previous design with LCMV and null-space, i.e., the notch remains wide at high frequencies. The notch is wide with a reasonable shape up to 3kHz, but there is more attenuation outside the desired notch for the frequencies above that.

Next we consider the constraint-based blocking matrix using null-space design for the 2+1 configuration. For the constraint vector *g*=[1.0; 1.0; 1.0] and theta_constraints equal to [0,10,350], nulls will be positioned at angles 0, 10 and 350 (-10) degrees, so it is expected that a wide notch can be obtained at least 10 degrees around the frontal direction, while ensuring that there is definitely a null in the frontal direction (in comparison with the previous case without using null-space design). The difference with the previous LCMV design and null-space is the opportunity to have one more constraint to control the beampattern.

The polar beampatterns for the constraint-based design with null-space for the 2+1 configuration binaural beamformer in free field (left side or right side), with HRTFs (left side) and HRTFs (right side) are shown respectively in Figure 5.10, Figure 5.11 and Figure 5.12.

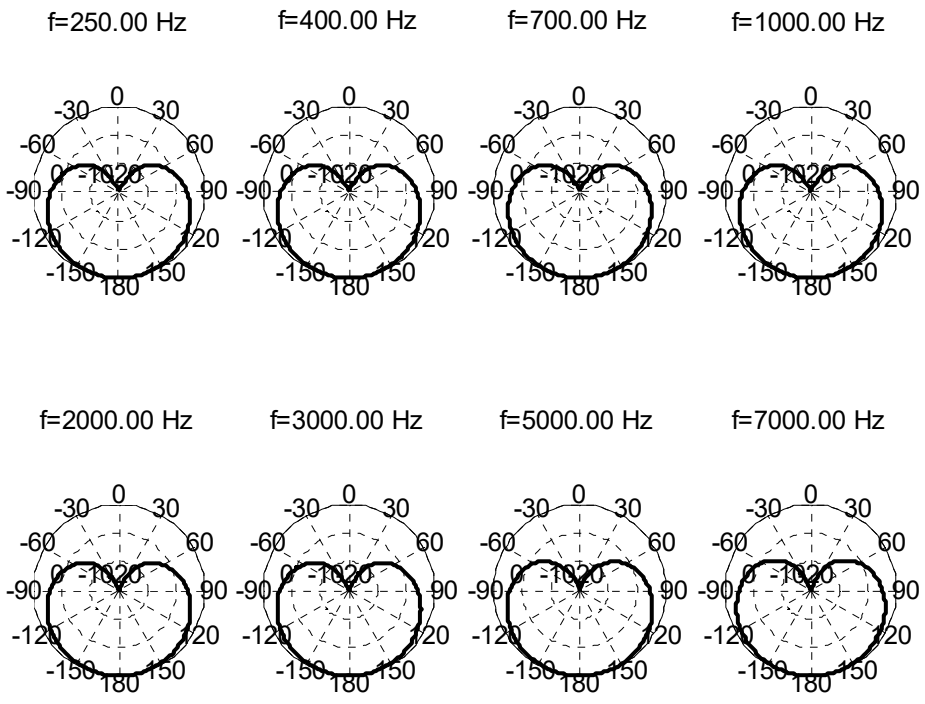


Figure 5.10 Beampatterns for blocking matrix design, constraint-based with null-space, 2+1 configuration (free field, left-side or right-side reference) when *g*=[1.0; 1.0; 1.0] and theta_constraints=[0,10,350]

The results are again nearly identical to the two previous designs (LCMV with null-space, constraint-based without null-space).

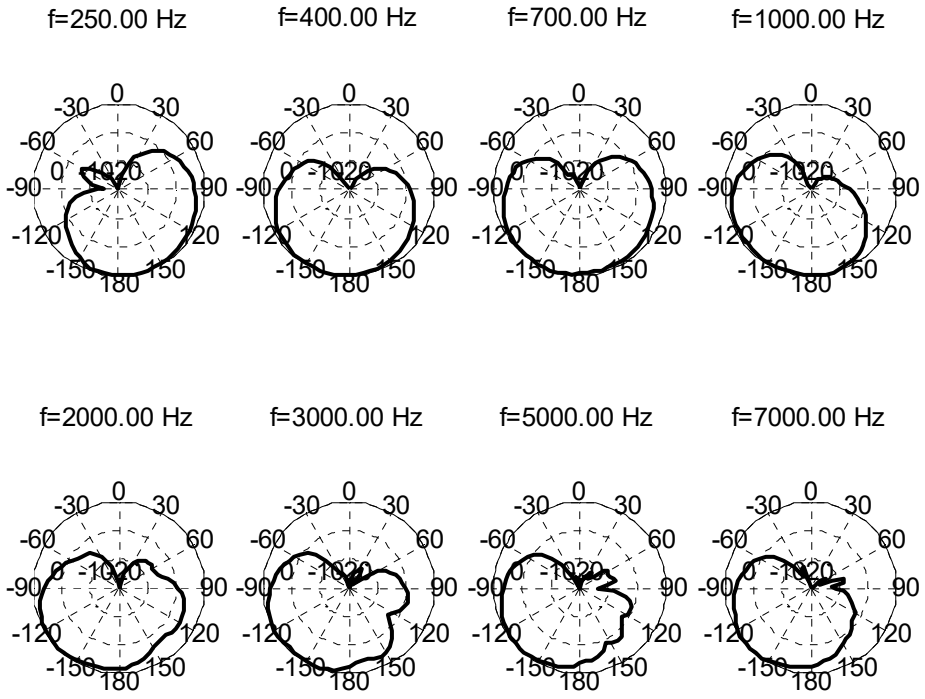


Figure 5.11 Beampatterns for blocking matrix design, constraint-based with null-space, 2+1 configuration (HRTFs, left-side reference) when *g*=[1.0; 1.0; 1.0] and theta_constraints=[0,10,350]

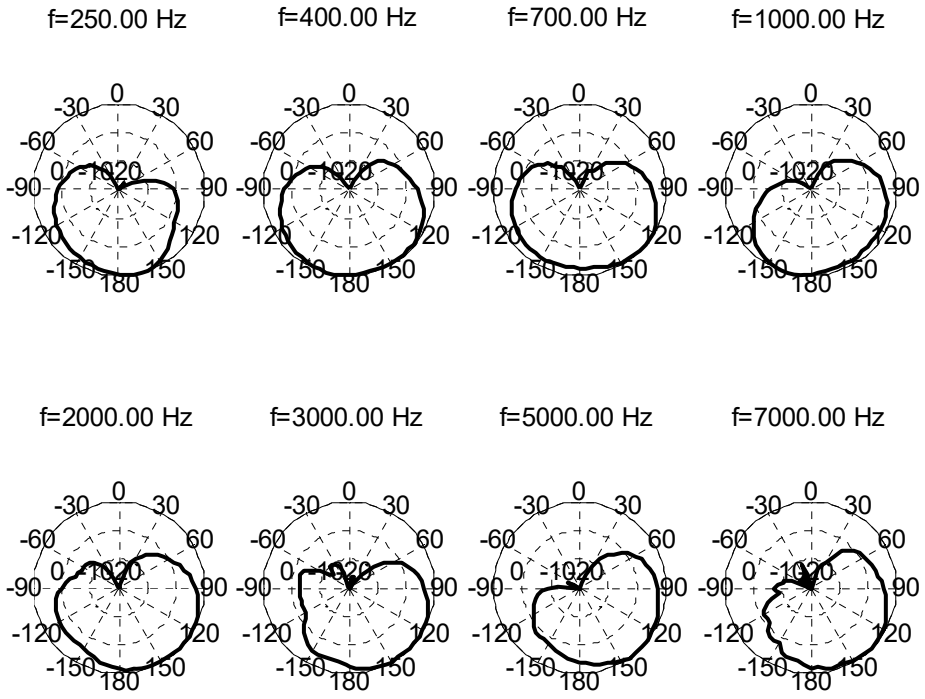


Figure 5.12 Beampatterns for blocking matrix design, constraint-based with null-space, 2+1 configuration (HRTFs, right-side reference) when *g*=[1.0; 1.0; 1.0] and theta_constraints=[0,10,350]

Here as well, the results are very similar to the two previous designs (LCMV with null-space, constraint-based without null-space) for HRTFs conditions.

5.1.2 Beampatterns for Fixed Beamformer Designs

The "reference" fixed beamformer design, i.e. "target equalized and sum", has been explained in detail in chapter 4, and is considered first. The (4-microphones) 1+1 configuration with pre-processing on both sides is again the reference configuration, as it is currently being used in binaural hearing aids. The polar beampatterns for the target equalized and sum fixed beamformer in 1+1 configuration with pre-processing for free field (left side or right side) and with HRTFs (left side or right side) are shown respectively in Figure 5.13 and Figure 5.14.

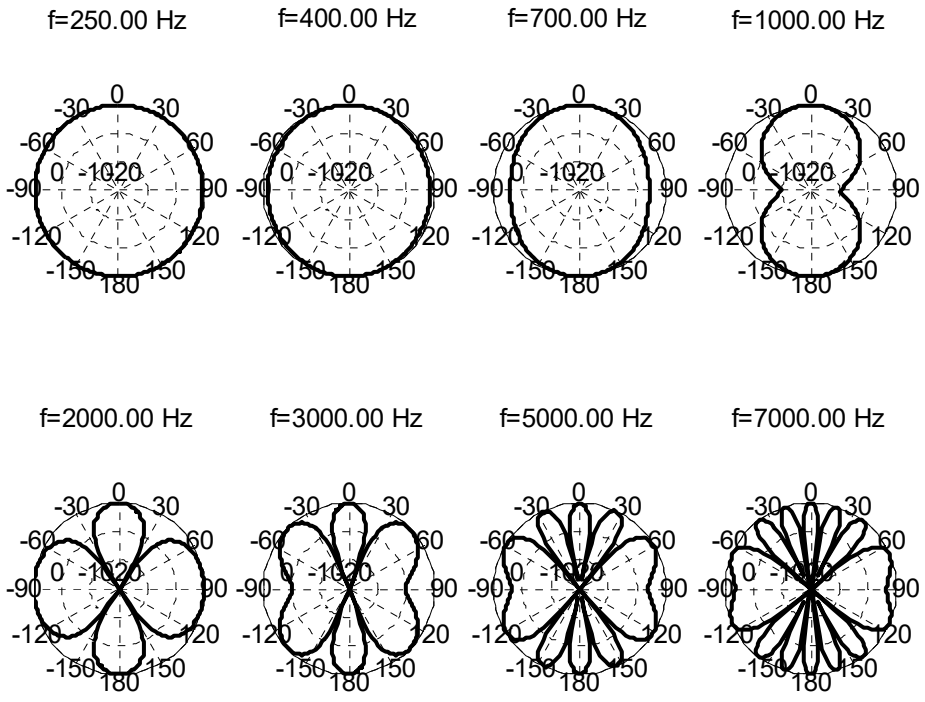


Figure 5.13 Beampatterns for fixed beamformer design, target equalize and sum, 1+1 binaural configuration (free field, left-side or right-side reference)

From Figure 5.13, the beam in the frontal direction for free field is wide and acceptable up to 2-3 kHz, but becomes more narrow starting at 2-3 kHz.

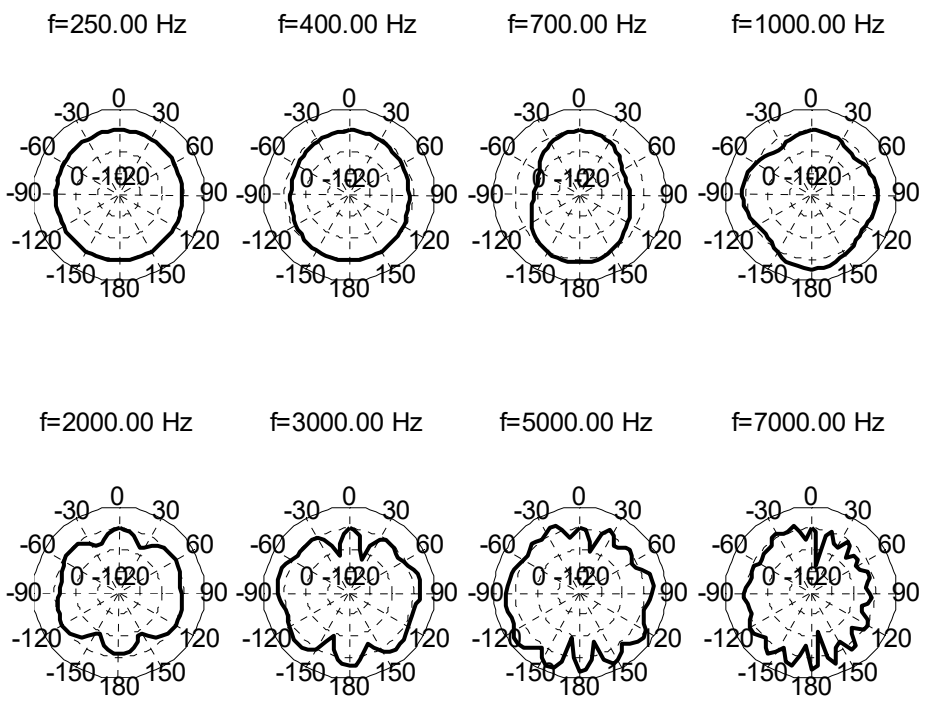


Figure 5.14 Beampatterns for fixed beamformer design, target equalize and sum, 1+1 binaural configuration (HRTFs, left-side or right-side reference)

As we can see from Figure 5.14, the results for the HRTFs conditions is overall similar to the free field conditions. The beam in the frontal direction becomes more narrow at 2-3 kHz and above. As a result, if the target source is not located exactly in the frontal direction, there might be some target signal distortion in the output of the fixed beamformer. Since the fixed beamformer output is used as the "desired signal" in the ANC system, the global output of the GSC consequently suffers from speech distortion. Thus, the idea of using a wide-beam fixed beamformer seems worthwhile to be investigated and implemented.

The Directivity Index (DI) for the target equalized and sum design with the 1+1 configuration binaural beamformer in free field (left side or right side) and with HRTFs (left side or right side) are shown respectively in Figure 5.15 and Figure 5.16. The DIs are shown for fixed beamformer algorithms to compare the effectiveness of different fixed beamformers in term of

diffuse noise suppression. The higher DI values mean a better performance in diffuse noise suppression.

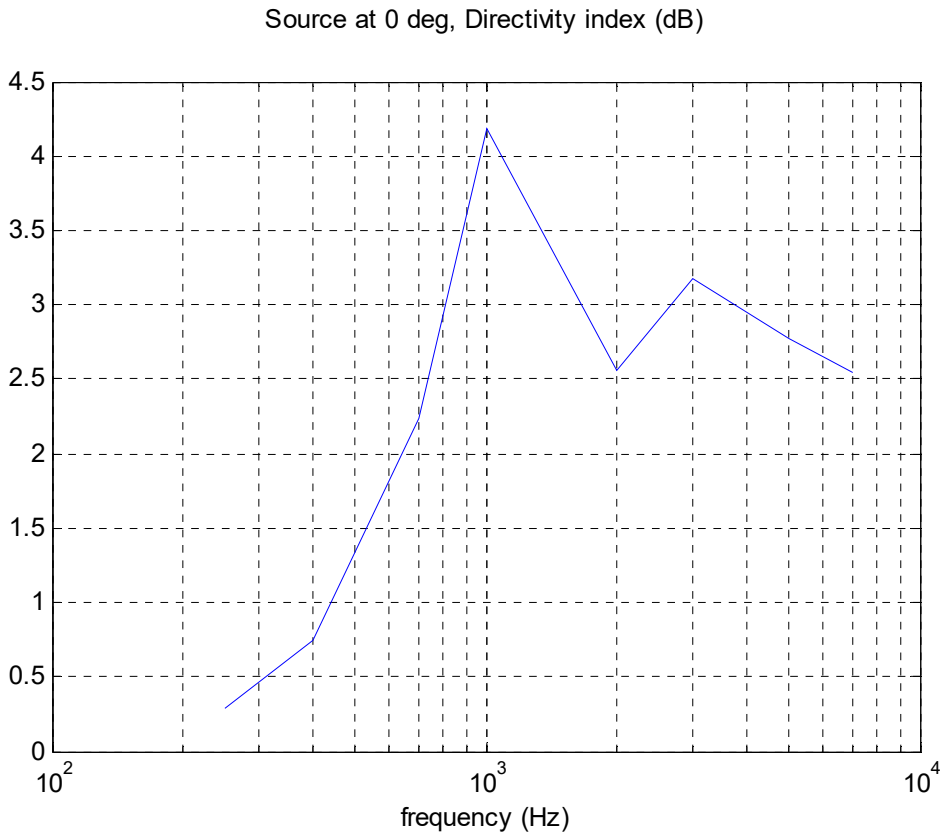


Figure 5.15 DI for fixed beamformer design, target equalize and sum, 1+1 binaural configuration (free field, left-side or right-side reference)

For the free field 1+1 with pre-processing binaural configuration, the DI for the target equalized and sum fixed beamformer varies from 2.5 to 4 dB.

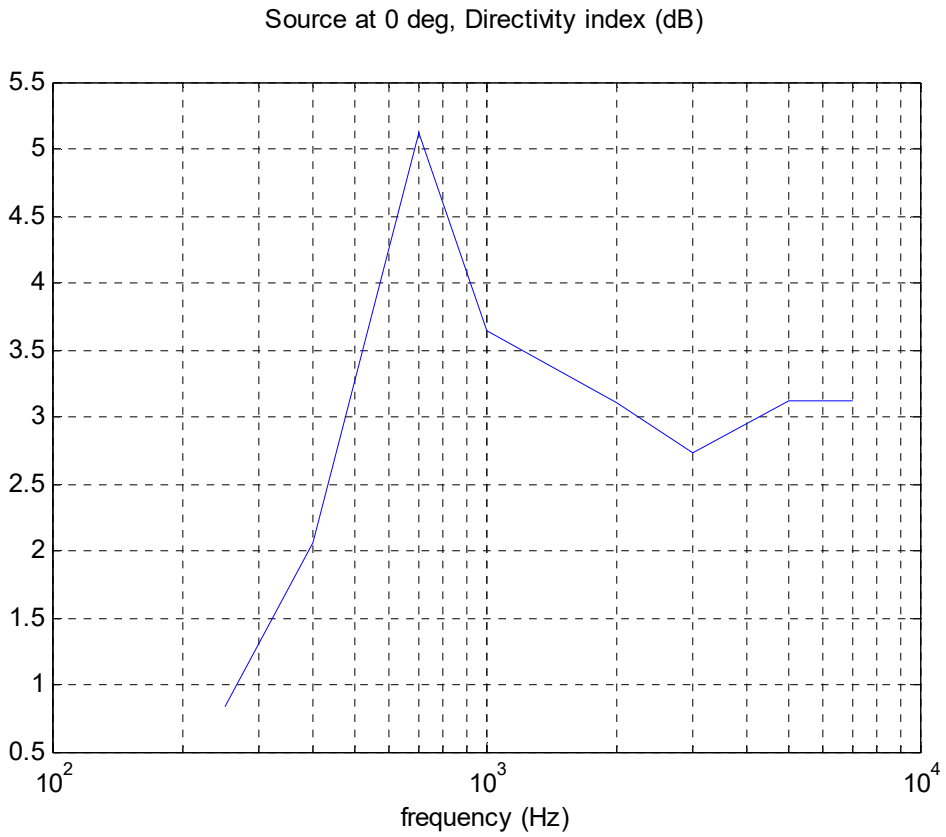


Figure 5.16 DI for fixed beamformer design, target equalized and sum, 1+1 binaural configuration (HRTFs, left-side or right-side reference)

For the 1+1 with pre-processing binaural configuration with HRTFs, the DI for the target equalized and sum fixed beamformer varies from 3 to 5 dB.

The target equalized and sum fixed beamformer for the 2+1 configuration is next considered to experimentally see the influence of using a 2+1 configuration instead of the 1+1 with pre-processing configuration, when steering the beam in the direction of the target signal. The polar beampatterns for the target equalized and sum fixed beamformer design for the 2+1 configuration binaural beamformer in free field (left side), in free field (right side), with HRTFs (left side) and with HRTFs (right side) are shown respectively in Figure 5.17, Figure 5.18, Figure 5.19 and Figure 5.20.

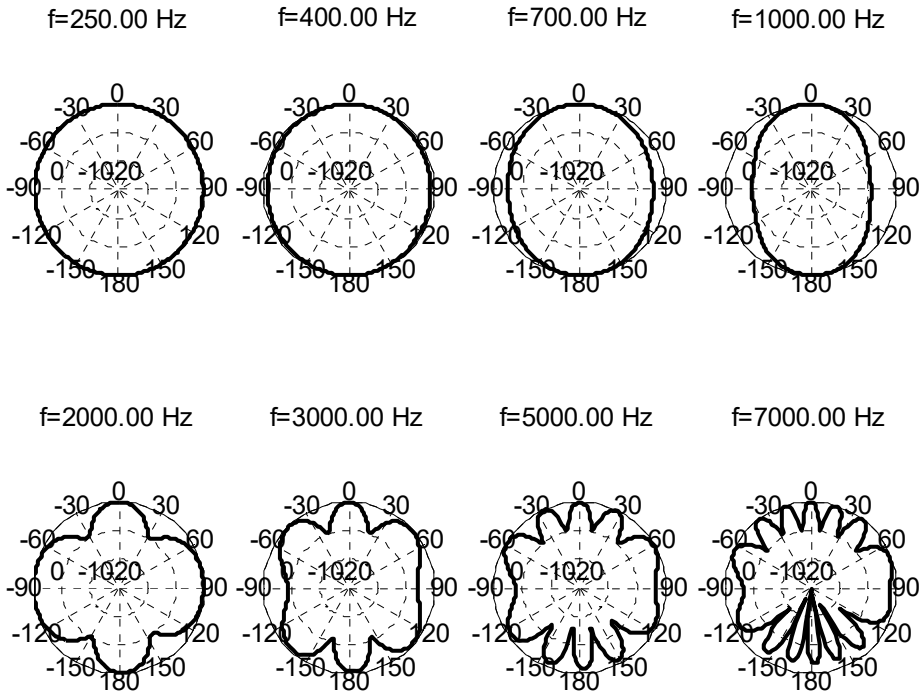


Figure 5.17 Beampatterns for fixed beamformer design, target equalized and sum, 2+1 binaural configuration (free field, left-side reference)

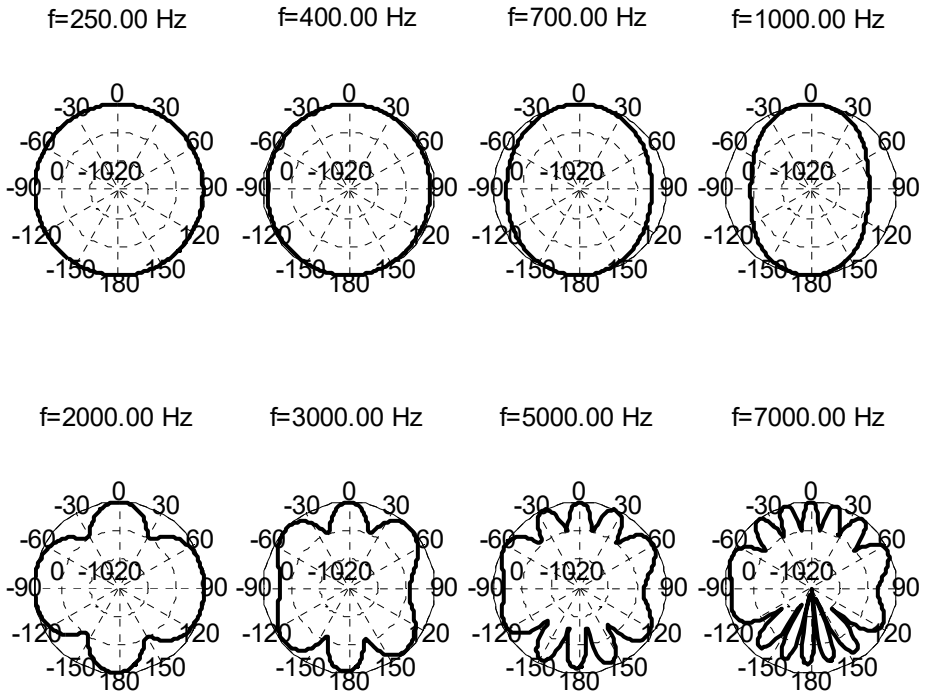


Figure 5.18 Beampatterns for fixed beamformer design, target equalize and sum, 2+1 binaural configuration (free field, right-side reference)

It can be concluded from Figure 5.17 and 5.18 that for the target equalize and sum in free field 2+1 binaural configuration, the beam becomes more narrow at 2-3 kHz but slightly less so than for the 1+1 target equalize and sum (Figure 5.13).

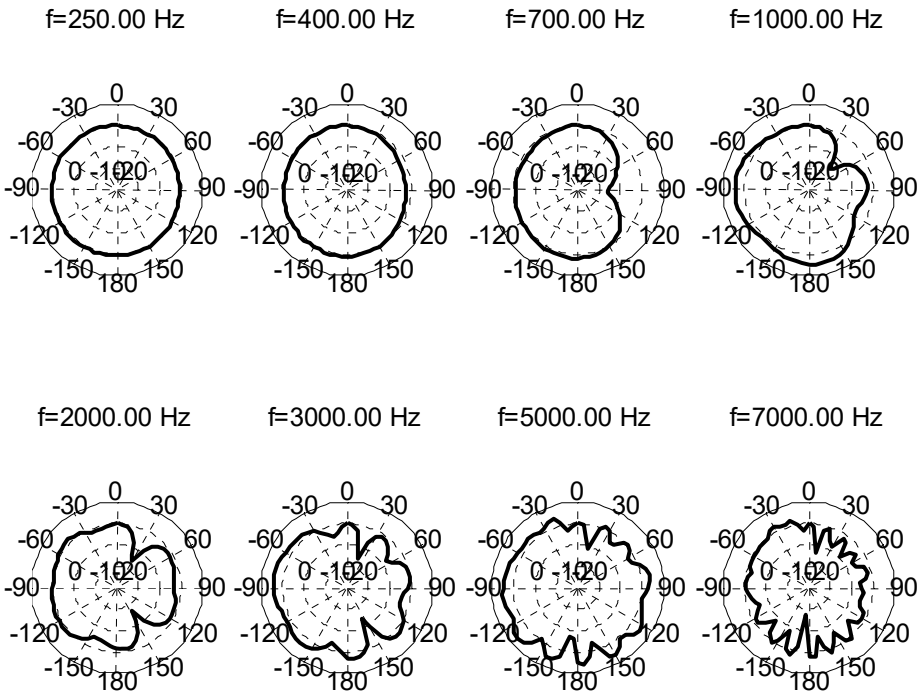


Figure 5.19 Beampatterns for fixed beamformer design, target equalize and sum, 2+1 binaural configuration (HRTFs, left-side reference)

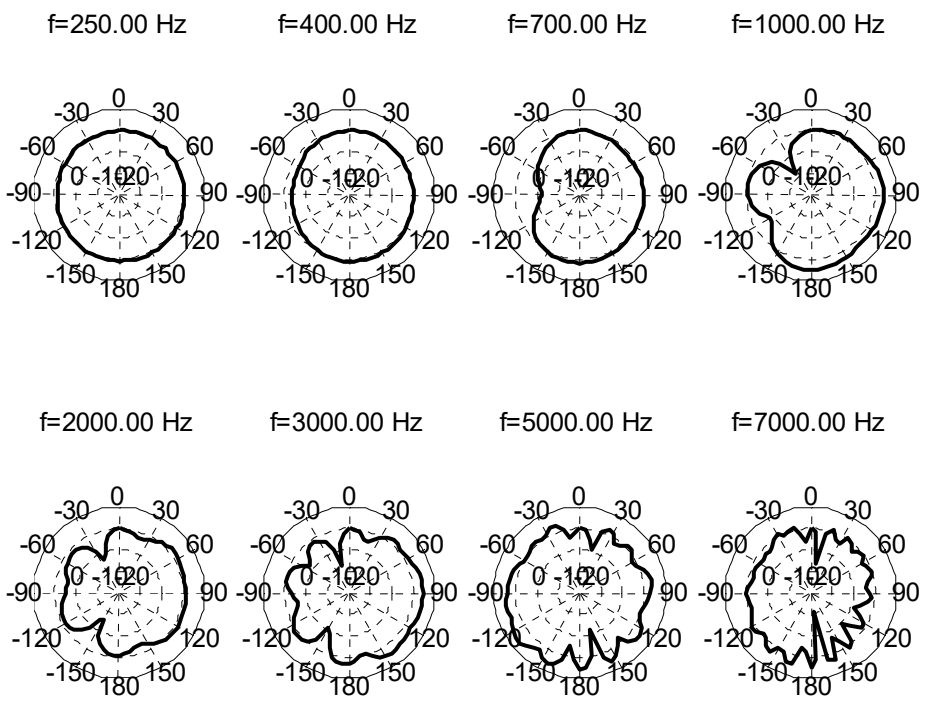


Figure 5.20 Beampatterns for fixed beamformer design, target equalize and sum, 2+1 binaural configuration (HRTFs, right-side reference)

It can be concluded from Figure 5.19 and 5.20 that for the target equalize and sum with HRTFs 2+1 binaural configuration, the beam becomes more narrow starting at 2-3 kHz (but perhaps slightly less so than for the 1+1 target equalize and sum in Figure 5.14). But overall, when using the target equalize and sum fixed beamformer there is no significant benefit in using the 2+1 binaural configuration instead of the 1+1 binaural configuration if the beam width is the factor being considered.

The directivity index for the target equalized and sum design with the 2+1 configuration binaural beamformer in free field (left side) and with HRTFs (left side) are shown respectively in Figure 5.21 and Figure 5.22.

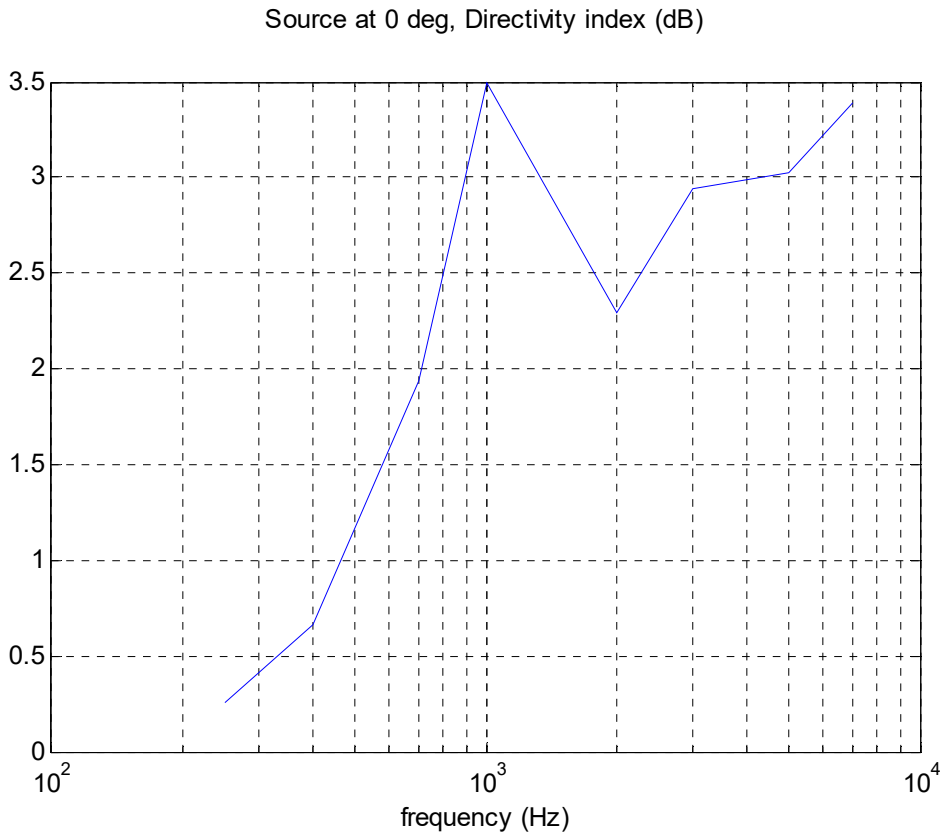


Figure 5.21 DI for fixed beamformer design, target equalized and sum, 2+1 binaural configuration (free field, left-side reference)

For the free field 2+1 binaural configuration, the DI for the target equalized and sum fixed beamformer varies from 2.5 to 3.5 dB, which is very similar to the DI for the target equalized and sum in the 1+1 binaural configuration.

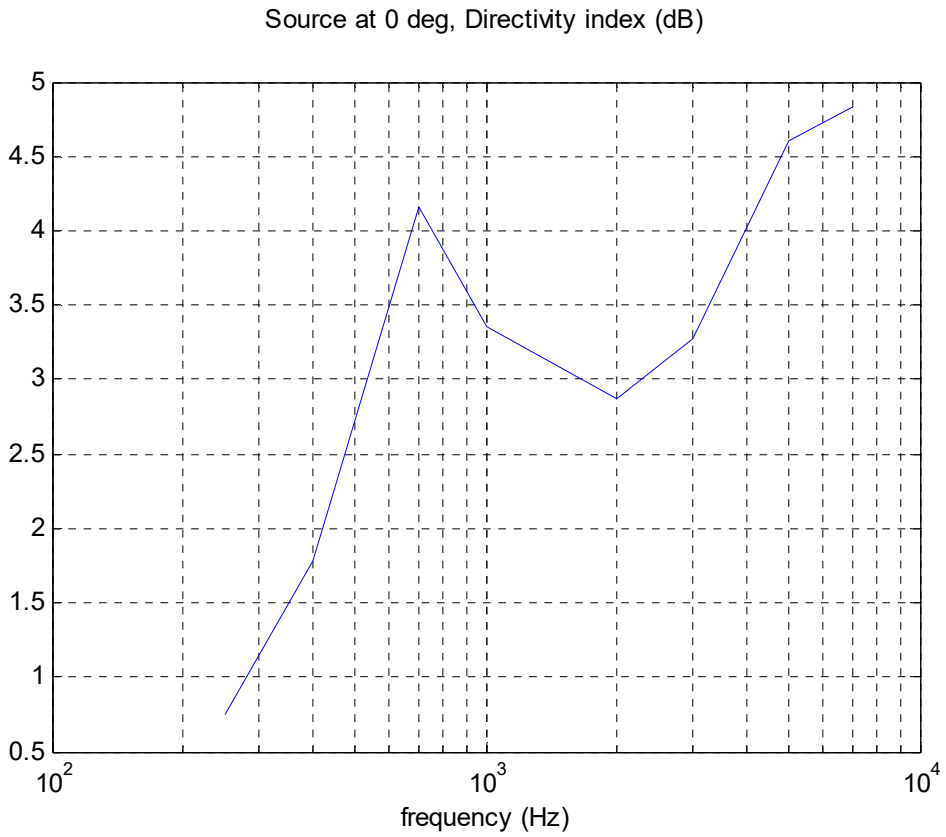


Figure 5.22 Beampatterns for fixed beamformer design, target equalize and sum, 2+1 binaural configuration (HRTFs, left-side reference)

With HRTFs the target equalized and sum fixed beamformer for 2+1 binaural configuration produces a DI which varies from 3 to 5 dB, which again is very similar to the DI for the target equalized and sum in the 1+1 binaural configuration.

The next fixed beamformer algorithm considered is the LCMV, with the expectation that a wider beam can be obtained in the target direction. The LCMV for the 2+1 configuration with the constraint vector g=[1.0; 1.0] and theta_constraints=[10,350] has been implemented. The polar beampatterns for the LCMV fixed beamformer with for 2+1 configuration binaural beamformer for free field (left side or right side), with HRTFs (left side) and HRTFs (right side) are depicted respectively in Figure 5.23, Figure 5.24 and Figure 5.25.

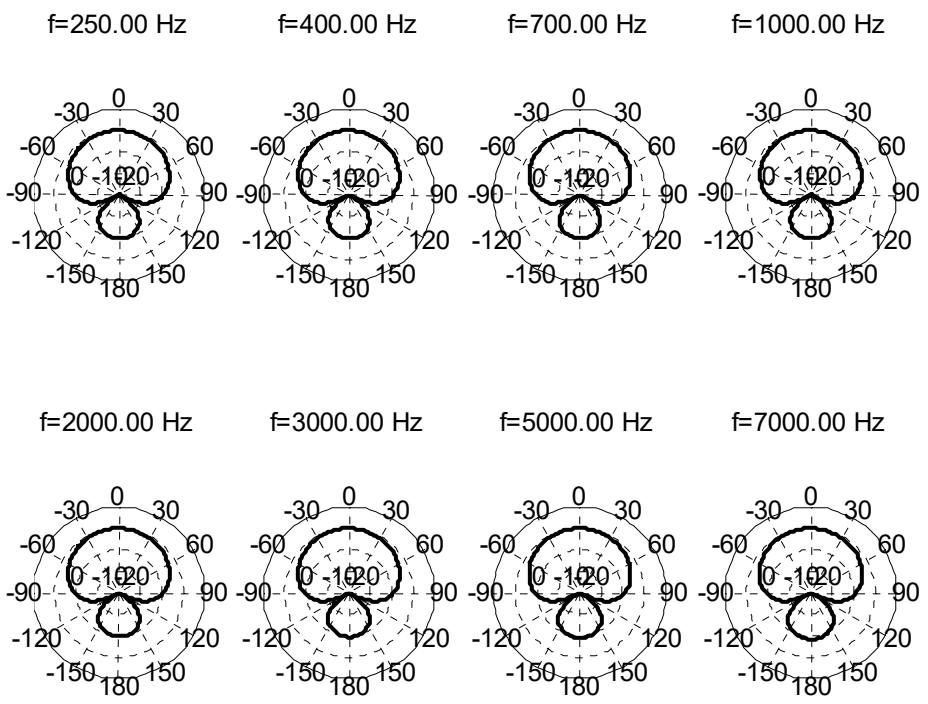


Figure 5.23 Beampatterns for fixed beamformer design, LCMV, 2+1 configuration (free field, left-side or right-side reference) when *g*=[1.0; 1.0] and theta_constraints=[10,350]

A wider beam in the target direction is achieved for all frequencies with the LCMV design for the 2+1 configuration in free field. It should be noted here that the noise correlation matrix used in the LCMV design is a constant diffuse noise correlation matrix, i.e., it is not an angle-dependent directional interference correlation matrix. The LCMV coefficients are consequently constant for all interferer directions.

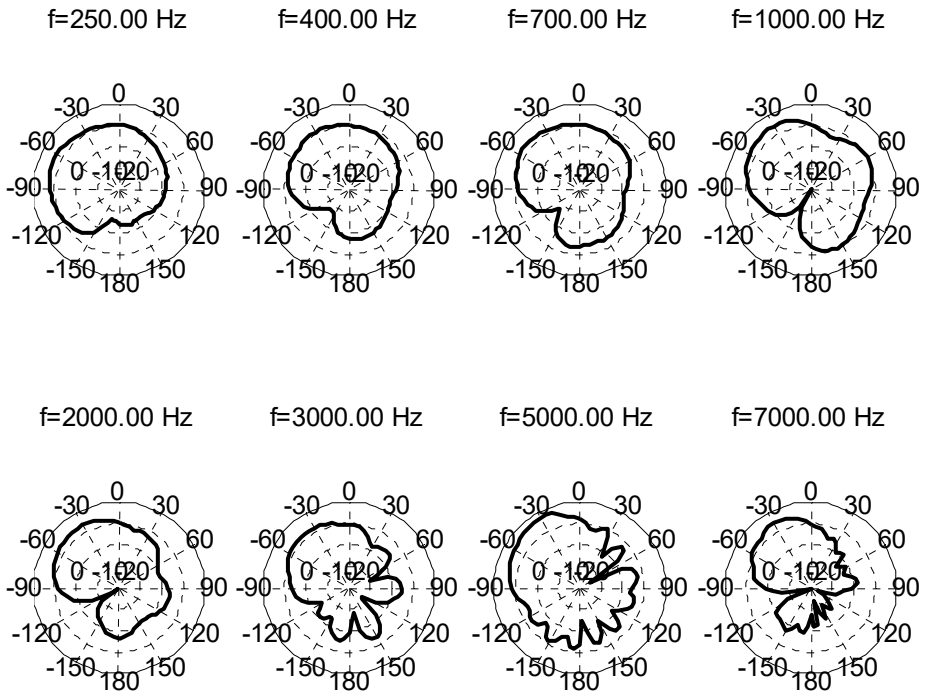


Figure 5.24 Beampatterns for fixed beamformer design, LCMV, 2+1 configuration (HRTFs, left-side reference) when *g*=[1.0; 1.0] and theta_constraints=[10,350]

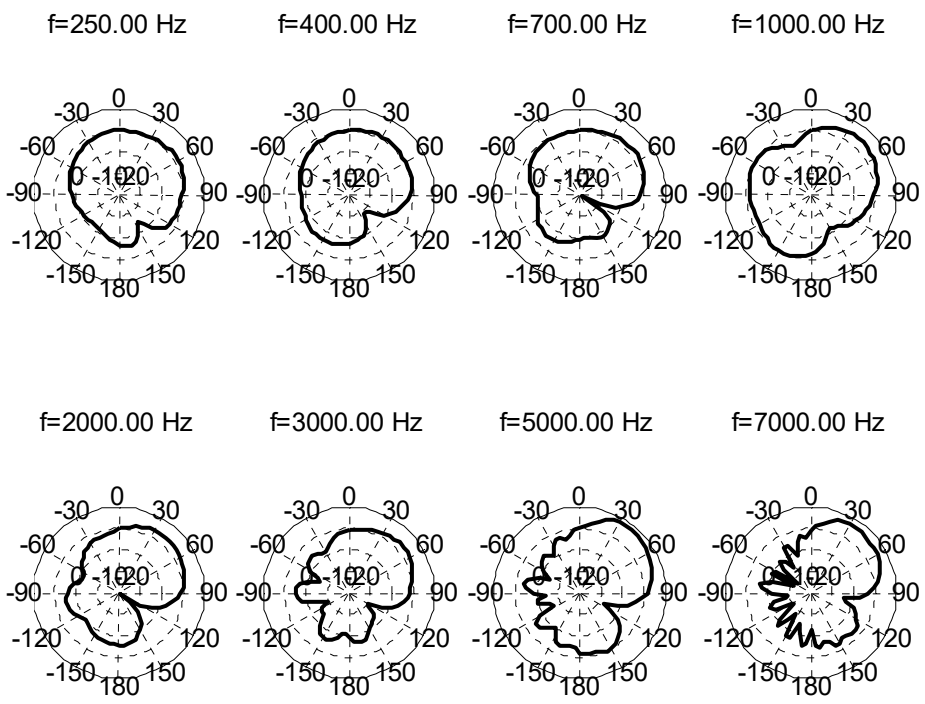


Figure 5.25 Beampatterns for fixed beamformer design, LCMV, 2+1 configuration (HRTFs, right-side reference) when *g*=[1.0; 1.0] and theta_constraints=[10,350]

It can be seen from Figure 5.24 and 5.25 that with HRTFs the LCMV with the 2+1 binaural configuration can produce a wide beam in the target direction for all frequencies but there are some angle-dependent fluctuations in the beampattern above 1 kHz.

The directivity index of the LCMV fixed beamformer for the 2+1 configuration binaural beamformer in free field (left side or right side) and with HRTFs (left side) are shown respectively in Figure 5.26 and Figure 5.27.

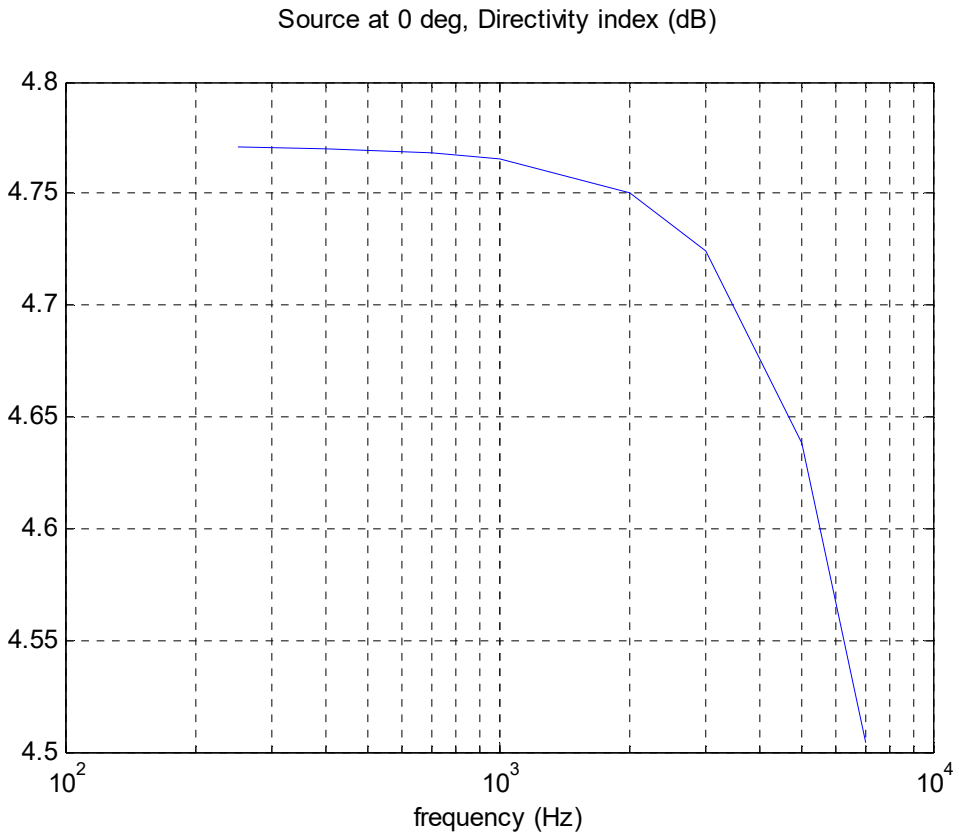


Figure 5.26 DI for fixed beamformer design, LCMV, 2+1 configuration (free field, left-side or right-side reference)

The directivity index in the free field condition is up to 4.8 dB, which is a better performance than the 2+1 target equalized and sum fixed beamformer (Figure 5.21, 2.5-3.5 dB).

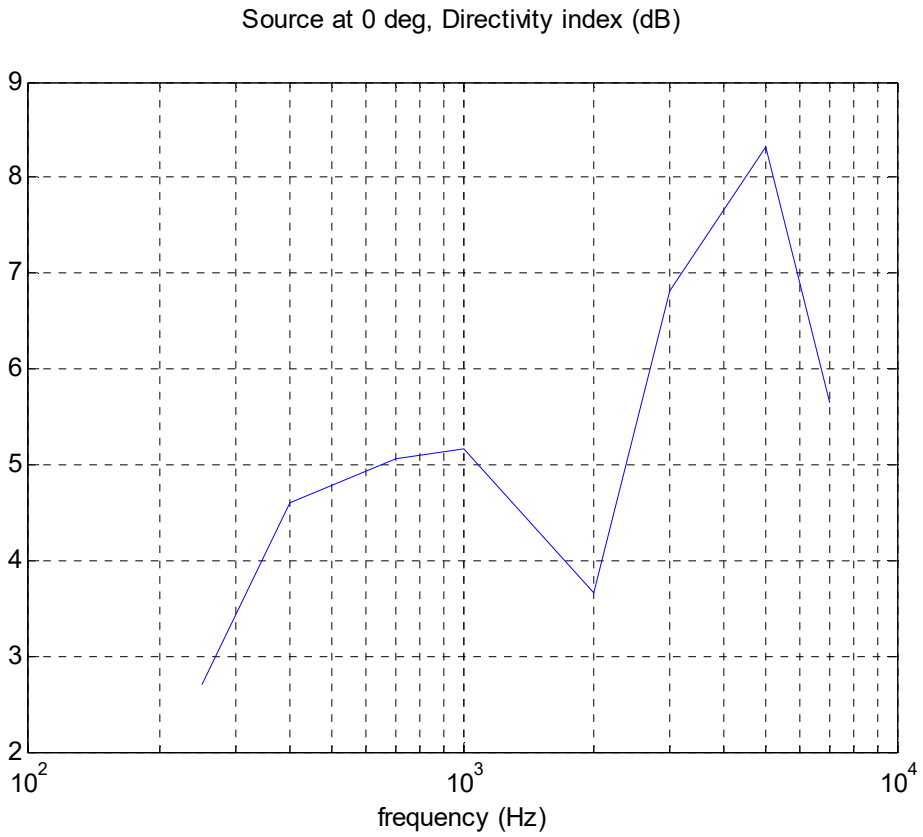


Figure 5.27 DI for fixed beamformer design, LCMV, 2+1 configuration (HRTFs, left-side reference):

In comparison with the 2+1 target equalized and sum fixed beamformer with HRTFs (Figure 5.22, 3-5 dB DI), the LCMV fixed beamformer with wider beam design produces a better performance: DI between 4-8 dB. Moreover, the GSC can take the advantage of its better robustness to DOA mismatch.

Another wide-beam fixed beamformer design described in chapter 3 is next considered: the constraint-based design. The constraint-based design fixed beamformer for the 2+1 configuration with the constraint gain vector g=[1.0; 1.0; 0.0] and theta_constraints=[10,350,180] has been implemented. The zero constraint at 180 degrees is necessary in order to avoid getting an omnidirectional response from the design. The constraint-based design has the capability to use

one more constraint compared to the LCMV (due to not using a noise correlation matrix and the minimization of the matrix trace in the design).

The polar beampatterns of the constraint-based design for the 2+1 configuration binaural beamformer in free field (left side or right side), with HRTFs (left side) and with HRTFs (right side) are shown respectively in Figure 5.28, Figure 5.29 and Figure 5.30.

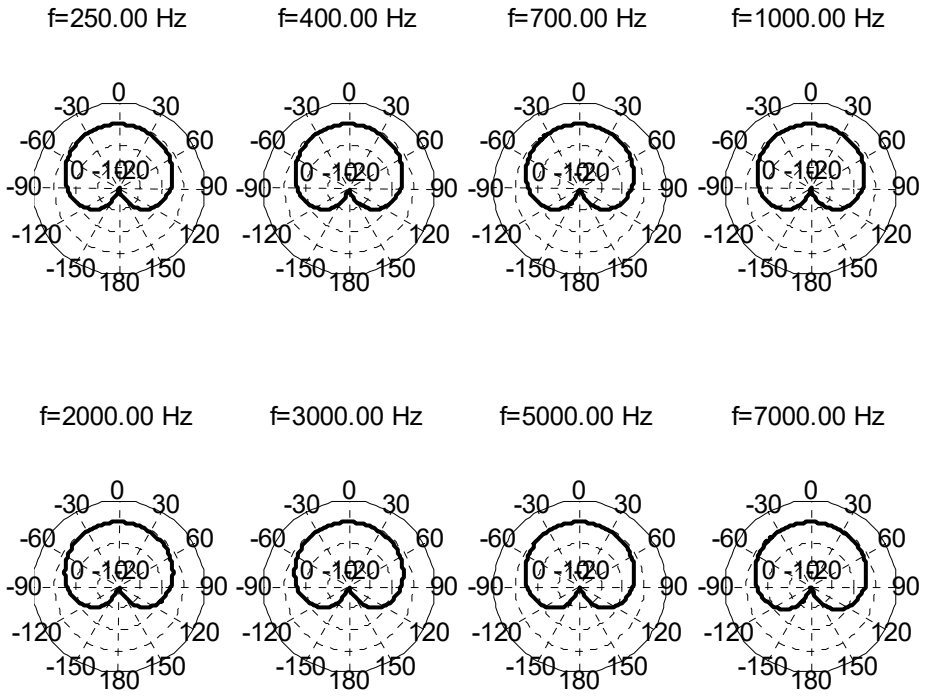


Figure 5.28 Beampatterns for fixed beamformer design, constraint-based, 2+1 configuration (free field, left-side or right-side reference)

The results show a good capability of steering a wide beam in the frontal target direction for all frequencies in free field. This is similar to the previous LCMV design (Figure 5.23).

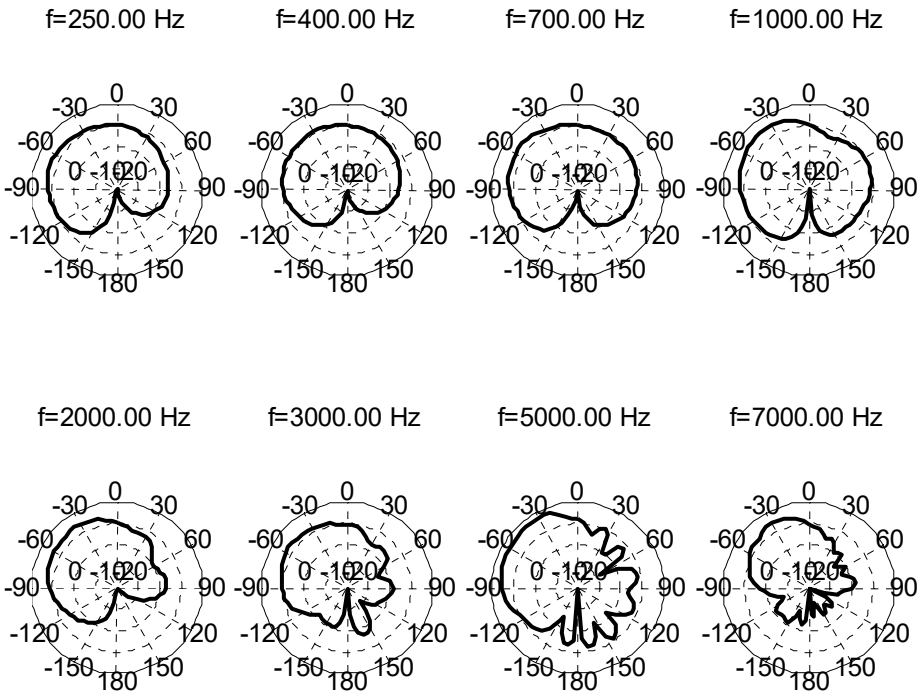


Figure 5.29 Beampatterns for fixed beamformer design, constraint-based, 2+1 configuration (HRTFs, left-side reference)

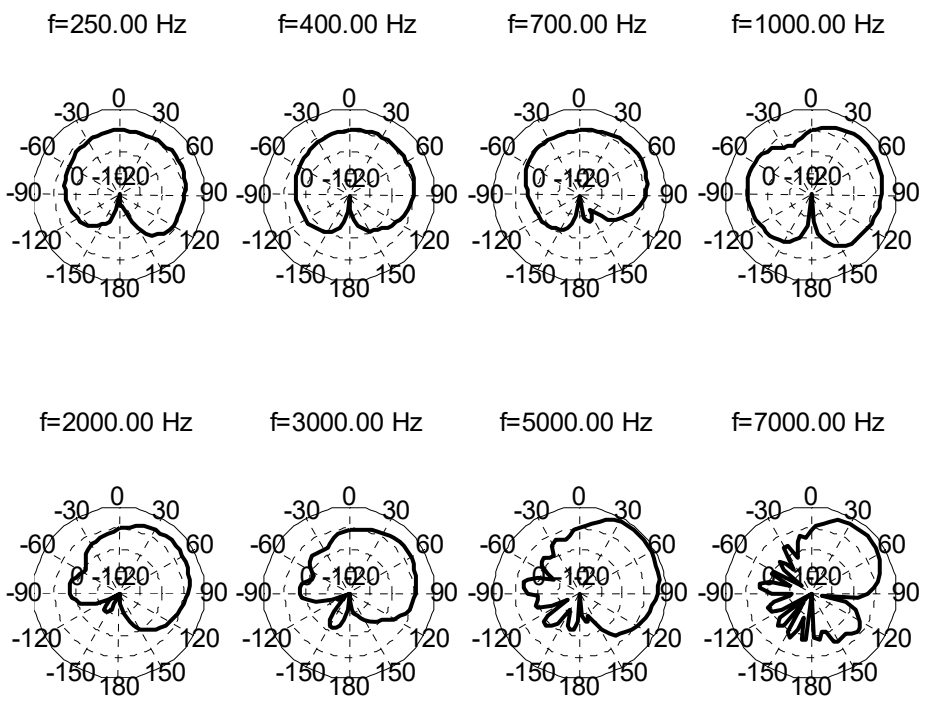


Figure 5.30 Beampatterns for fixed beamformer design, constraint-based, 2+1 configuration (HRTFs, right-side reference)

With HRTFs the results of the constraint-based design for the 2+1 configuration are similar to the previous LCMV design (Figure 5.24 and 5.25), i.e., a wide beam in the frontal target direction is achieved for all frequencies but there are some angle-dependent fluctuations in the beam gain above 1 kHz.

The directivity index for the constraint-based fixed beamformer with the 2+1 configuration binaural beamformer in free field (left side or right side) and with HRTFs (left side) are shown respectively in Figure 5.31 and Figure 5.32.

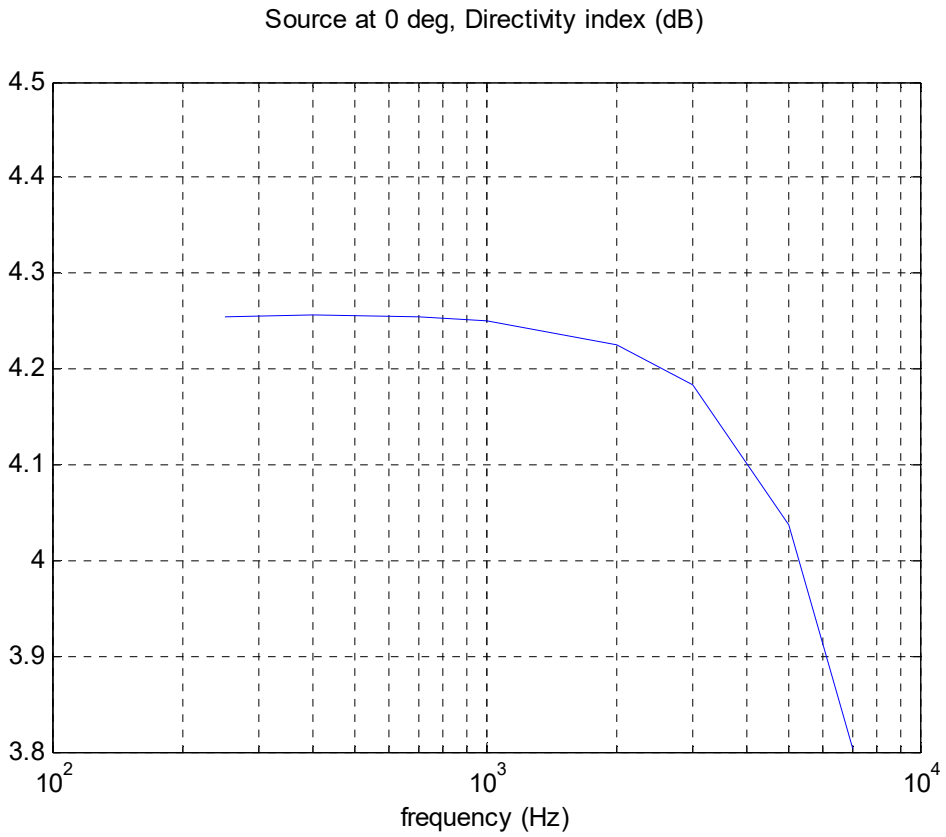


Figure 5.31 DI for fixed beamformer design, constraint-based, 2+1 configuration (free field, left-side or right-side reference)

The directivity index for the 2+1 fixed beamformer constraint-based design (approximately up to 4.25 dB in Figure 5.31) is slightly less than with the previous design using LCMV (which was up to 4.80 dB in Figure 5.26). This was predictable since the LCMV design used a diffuse noise correlation matrix in the design, therefore leading to a better diffuse noise reduction performance and a better directivity index.

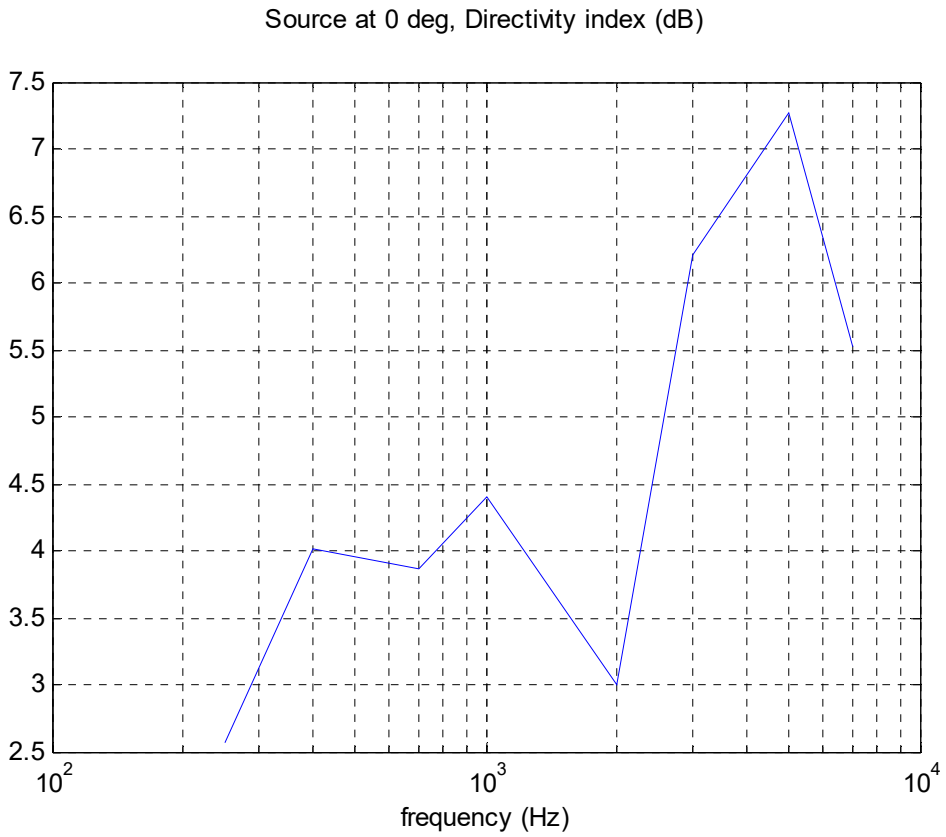


Figure 5.32 DI for fixed beamformer design, constraint-based, 2+1 configuration (HRTFs, left-side reference)

With HRTFs, the constraint-based design for the 2+1 configuration produces a directivity index which is similar to (or just slightly worse than) with the previous design using LCMV (Figure 5.27), i.e., 4-7 dB vs. 4-8 dB of directivity index. So overall, considering both the beampatterns and the directivity index curves, the LCMV design and the constraint-based design for the 2+1 fixed beamformer lead to comparable results, with a slight edge to the LCMV design using a diffuse noise correlation matrix. Both produce better beampatterns and directivity index than the reference target-equalize and sum 1+1 configuration with pre-processing or the target-equalize and sum 2+1 configuration.

The last fixed beamformer design considered in this section is the 2+1 configuration MVDR design tuned for diffuse noise. The main goal of using the MVDR is to take the advantage

of the high directivity index in this algorithm. Consequently, for this design it is not expected that the beam in the target direction will necessarily be wider than the reference target-equalize and sum 1+1 configuration or the target-equalize and sum 2+1 configuration, however the directivity index should be better.

The polar beampatterns for the MVDR fixed beamformer design tuned for diffuse noise in 2+1 configuration binaural beamformer with free field (left side), with free field (right side), with HRTFs (left side) and with HRTFs (right side) are shown respectively in Figure 5.33, Figure 5.34, Figure 5.35 and Figure 5.36.

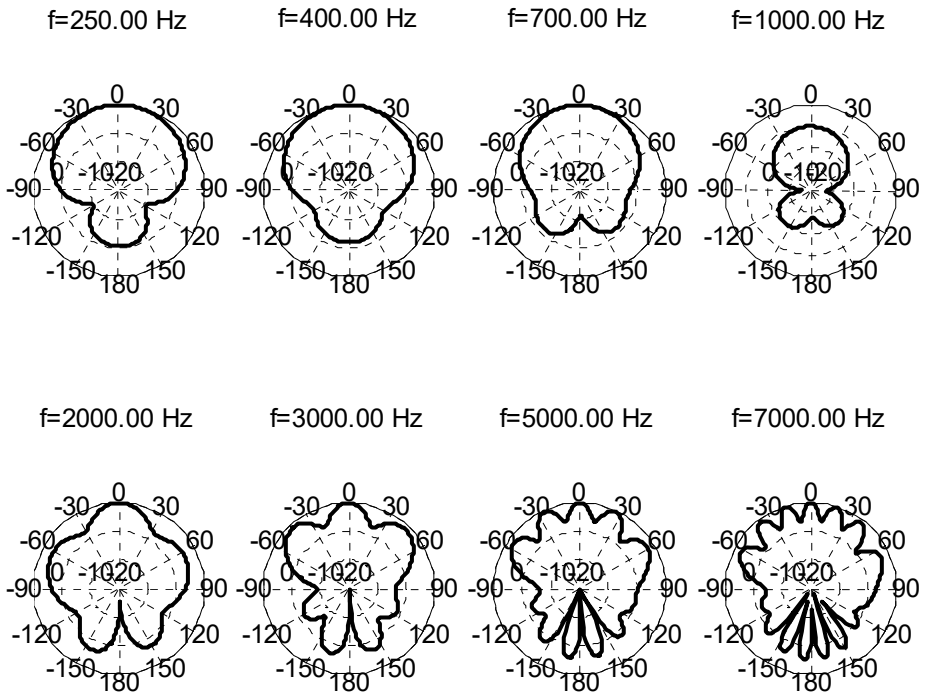


Figure 5.33 Beampatterns for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (free field, left-side reference)

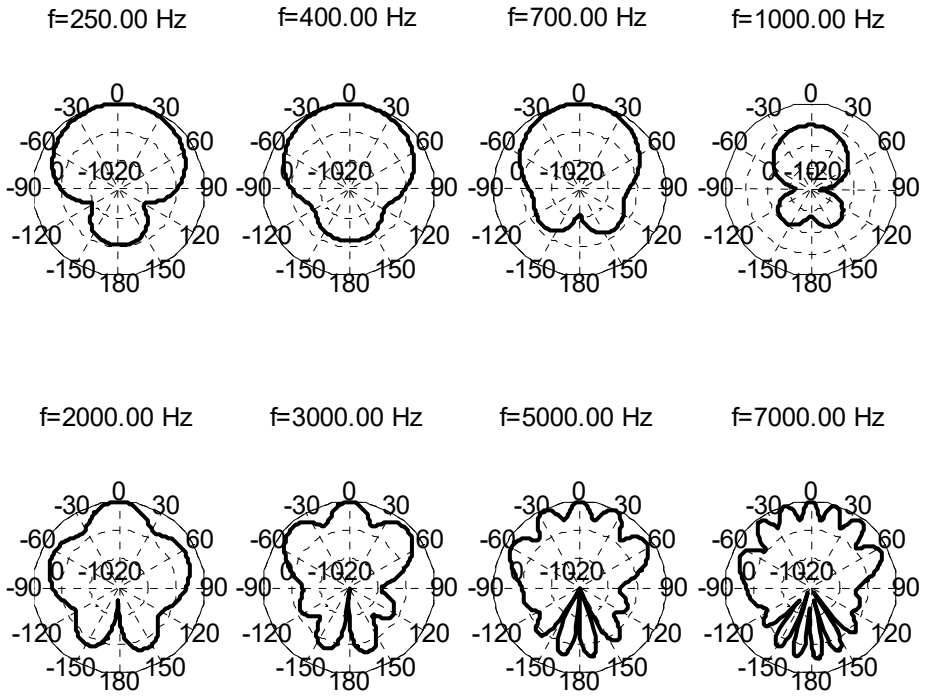


Figure 5.34 Beampatterns for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (free field, right-side reference)

In Figure 5.33 and Figure 5.34 for free field, it can be easily seen that compared to the previous designs for wide beam in the target direction (LCMV, constraint-based), this design for diffuse noise has a beam becoming more narrow starting at 2-3 kHz (similar to the target equalize and sum 2+1 design of Figure 5.17 and Figure 5.18, but not as bad as for the 1+1 with pre-processing target equalize and sum of Figure 5.13).

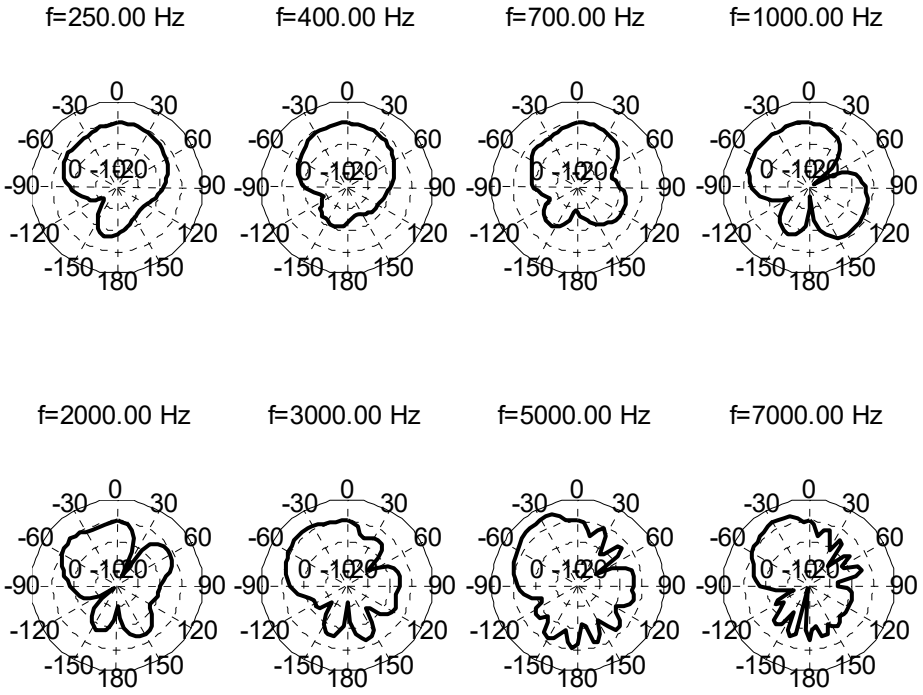


Figure 5.35 Beampatterns for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (HRTFs, left-side reference)

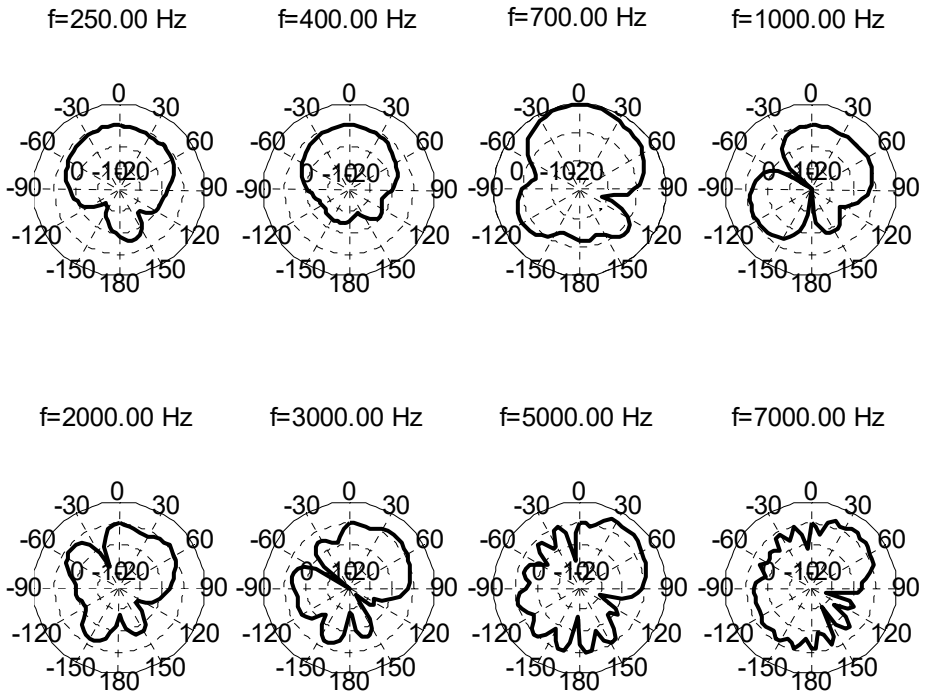


Figure 5.36 Beampatterns for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (HRTFs, right-side reference)

With HRTFs, compared to the previous designs with a wide beam in the target direction (LCMV, constraint-based), for the 2+1 configuration MVDR design tuned for diffuse noise the beam in the target direction becomes perhaps slightly more narrow starting at 2-3 kHz. But in any case it is better than the results using the target equalize and sum 2+1 design (Figure 5.19 and Figure 5.20) and the target equalize and sum 1+1 with pre-processing design (Figure 5.14).

The directivity index for the MVDR fixed beamformer tuned for diffuse noise 2+1 configuration binaural beamformer in free field (left side) and with HRTFs (left side) are shown respectively in Figure 5.37 and Figure 5.38.

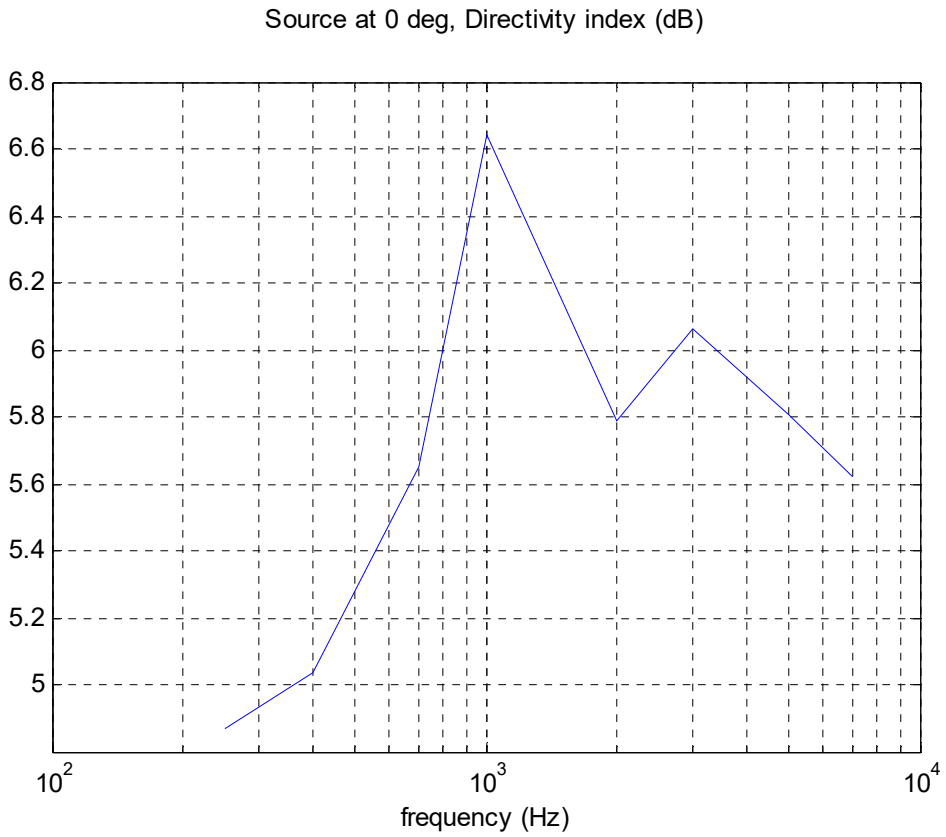


Figure 5.37 DI for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (free field, left-side reference)

In free field conditions, as expected the DI of the MVDR tuned for diffuse noise is better than for the wide beam designs (LCMV and constraint-based) or the target-equalize and sum designs: it varies from 5.5 dB to 6.6 dB, while all the previous designs had a DI below 5 dB in free field.

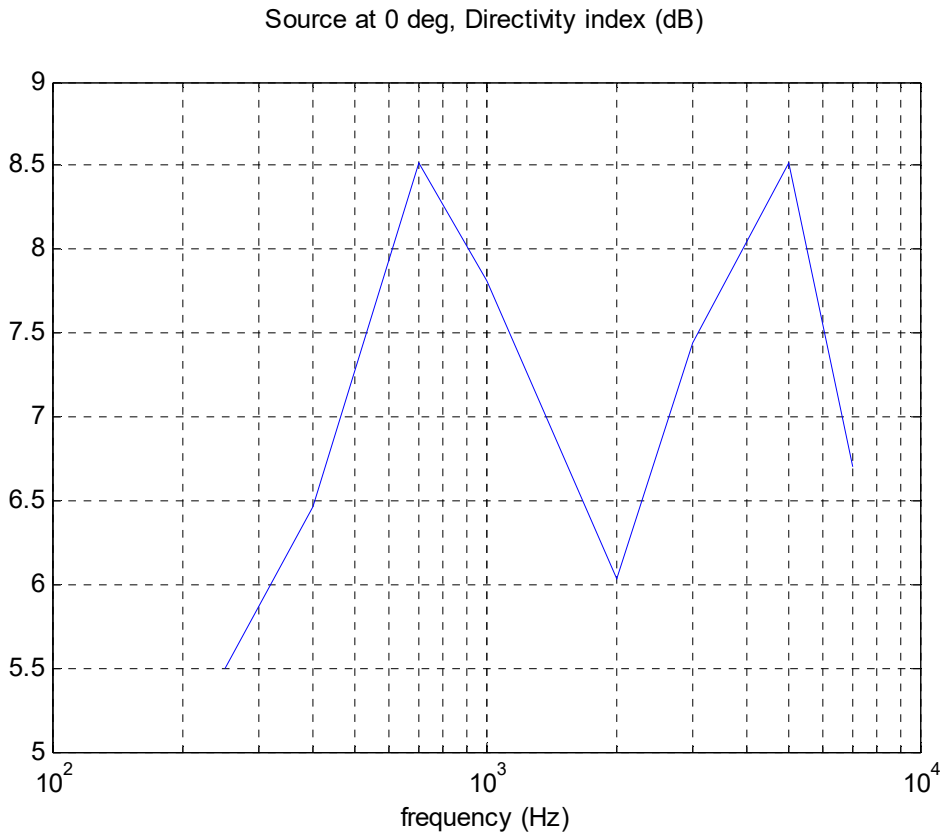


Figure 5.38 DI for fixed beamformer design, MVDR tuned for diffuse noise, 2+1 binaural configuration (HRTFs, left-side reference)

With HRTFs, again as expected the DI of the MVDR tuned for diffuse noise is better than for the wide beam designs (LCMV and constraint-based) or the target-equalize and sum designs: it varies from 6.0 dB to 8.5 dB, while the best previous design in terms of DI (with the LCMV design) had a DI varying between 4 dB and 8 dB.

5.2 Binaural Beamforming Simulation Results

This section discusses the evaluation of the algorithms described in the previous sections, utilizing realistic reverberant recordings provided by a hearing aid company. The 4-channel binaural recordings were provided for different speakers and for different source directions. 4-channel anechoic HRTFs were also provided, as well as 4-channel background noise recordings. The experimental setup for the simulations was described in detail in chapter 4. All the simulations were performed in the filter bank domain, with 10 seconds of recorded signals, and fixed source locations. For the GSC-ANC unit, all simulations were achieved with a multichannel least-squares algorithm and a solution was computed from the whole 10 seconds of signal. For the algorithms using equalization coefficients (e.g. target-equalized delay and sum fixed beamformer, target-equalized Griffiths-Jim blocking matrix), a direct estimation of the equalization coefficients from the target signals was performed.

5.2.1 Evaluation of the Target Equalized Griffiths-Jim Blocking Matrix for the 2+1 with No Pre-processing Configuration

As discussed in section 5.1, the 1+1 configuration Griffiths-Jim blocking matrix and the 2+1 configuration Griffiths-Jim blocking matrix will have similar limitations in terms of a narrow notch in the target direction, since they share one beamformer in common. Therefore, only the results of 2+1 configuration Griffiths-Jim blocking matrix are presented in this section. In order to focus on the impact of the blocking matrix DOA mismatch (and the resulting target leakage), and to enable an easy comparison with the wider-notch blocking matrix designs, in this section and the following sections comparing the performance of wide-notch blocking matrix designs the reference microphone input signal is used as the fixed beamformer output. In other words, there is no fixed beamformer, so that the target signal remains undistorted when the real target direction is not exactly the same as the target direction used in the designs.

The blocking matrix design was based on a frontal target direction (0 degree) and the coefficients were saved to be used later ('loaded' in Matlab) in cases of non-frontal target

directions. These 'save' and 'load' actions were first utilized to investigate the robustness to DOA mismatch of the target-equalized Griffiths-Jim blocking matrix for the 2+1 configuration. The noise signals in this experimental configuration include two interferers at 90 and 225 degrees and diffuse-like background noise. Throughout the thesis, the rms level of all the directional sources was adjusted such that for each source the level is normalized to 1.0 in the dominant front microphone (left or right front microphone). For the diffuse-like background noise, the rms level was adjusted in the same way, however the rms level of the diffuse-like background noise was varied depending on the experiment. Here the rms level of the diffuse-like background noise was set to 0.5 (6 dB below the level of each directional sources). The structure was tested for target directions 0, 5, 10, 20, 30, and 45 degrees. The "easier" case of signals generated by HRTFs filtering was first considered here (instead of recordings). The results for the target-equalized Griffiths-Jim blocking matrix 2+1 configuration are shown in Table 5.1:

Table 5.1 Results for GSC with target-equalized Griffiths-Jim blocking matrix, 2+1 configuration with HRTF-generated signals

	Target direction:	0	5	10	20	30	45
Left side	SNR_in_ref_sensor_dB	-1.47	-1.59	-2.74	-4.80	-6.06	-7.21
	SNR_out_dB	8.90	5.94	3.94	0.37	-2.78	-6.03
	SNR_gain_dB	10.37	7.53	6.68	5.18	3.28	1.18
	SDR dB	29.45	7.72	6.51	6.73	4.41	2.82
Right side	SNR_in_ref_sensor_dB	-1.96	-1.33	-1.33	-1.33	-1.33	-1.33
	SNR_out_dB	7.88	5.88	3.43	0.22	-2.43	-4.15
	SNR_gain_dB	9.84	7.20	4.76	1.55	-1.11	-2.83
	SDR dB	28.86	7.28	4.91	3.96	2.73	2.03

From Table 5.1, the SNR_gain and SDR for a frontal target direction without DOA mismatch (target direction 0 degree) were very good. However, even for just 5 degrees of DOA deviation in the target direction there was a significant degradation in SNR and SDR. The performance rapidly deteriorated for directions greater than 5 degrees, due to a narrow-notch blocking matrix causing target signal leakage to the output of the blocking matrix unit. When listening to the wave files for the output of the blocking matrix unit, the target leakage for target

directions starting at 10 degrees could be clearly heard. Informally, we found that when the SDR distortion criterion is higher than 7.0 dB then the target leakage component is not significant in the blocking matrix outputs.

It should be noted that that the input SNR remains the same when the target direction is changing because the same specified rms value (based on the side with the strongest level) were set for the target, the interferers and the background noise, regardless of their direction.

5.2.2 Evaluation of LCMV with Null-Space Design Blocking Matrix for 2+1 Configuration

The proposed wide-notch blocking matrix LCMV with null-space design was designed for a frontal target direction using anechoic HRTFs. One difference between the LCMV blocking matrix and the previous Griffiths-Jim blocking matrix is the number of outputs. Here the proposed LCMV blocking matrix has only one output for all scenarios, whereas the number of Griffiths-Jim blocking matrix outputs is equal to the number of microphones minus one. Therefore, there are two blocking matrix outputs for the Griffiths-Jim blocking matrix 2+1 configuration.

Simulations first were performed with HRTF-generated signals (because the designs done offline were based on HRTFs). Afterwards the designed blocking matrix was tested with reverberant signals to test for HTRF mismatch in addition to DOA mismatch. The noise signals in this experimental configuration were two interferers at 90 and 225 degrees and diffuse-like background noise (at rms level 0.5). The structure was tested for target directions 0, 5, 10, 15, 20, 25, 30, 45, 50 and 65 degrees. The results for the GSC using LCMV with null-space design as the blocking matrix, for 2+1 configuration with HRTF-generated signals and with reverberant recording signals are shown respectively in Table 5.2 and Table 5.3.

Table 5.2 Results for GSC with LCMV null-space design blocking matrix, 2+1 configuration with HRTF-generated signals

	Target direction:	0	5	10	15	20	25	30	45	50	65
Left side	SNR_in_ref_sensor	-1.47	-1.59	-2.74	-3.91	-4.80	-5.47	-6.06	-7.21	-7.46	-8.00
	SNR_out_dB	7.35	7.40	6.90	5.53	4.43	3.60	2.45	-0.46	-1.80	-4.50
	SNR_gain_dB	8.81	8.99	9.64	9.43	9.24	9.07	8.51	6.75	5.66	3.50
	SDR dB	16.18	16.35	17.56	15.70	14.52	13.12	11.33	7.22	5.94	3.85
Right side	SNR_in_ref_sensor	-1.96	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33
	SNR_out_dB	4.14	4.80	5.04	4.88	4.69	4.64	4.36	3.49	2.75	1.12
	SNR_gain_dB	6.10	6.13	6.37	6.20	6.01	5.97	5.68	4.82	4.07	2.44
	SDR dB	17.45	18.52	20.36	19.33	18.23	16.70	15.49	11.27	9.25	6.03

By comparing Table 5.1 and Table 5.2, the GSC system using the proposed wide-notch LCMV null-space blocking matrix performs slightly worse than the reference target-equalized Griffiths-Jim blocking matrix for a frontal target direction (no DOA mismatch) and for a target at 5 degrees (5 degrees DOA mismatch). However, for larger DOA mismatch values the LCMV null-space blocking matrix shows much more robustness to DOA mismatch (up to 30-45 degrees). Therefore, the proposed LCMV null-space blocking matrix exhibits superior performance for HRTF-generated signals.

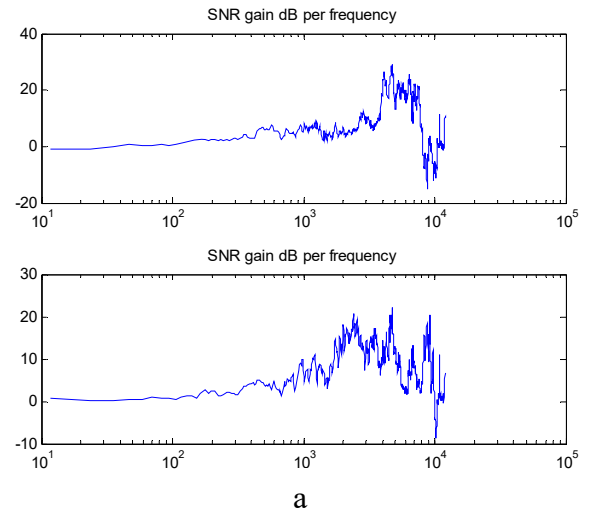
Table 5.3 Results for GSC with LCMV null-space design blocking matrix, 2+1 configuration, with reverberant recordings

	Target direction:	0	5	15	25	35	45	65
Left side	SNR_in_ref_sensor_dB	-1.94	-2.37	-3.55	-4.54	-5.26	-5.73	-6.00
	SNR_out_dB	1.29	1.26	-0.22	-1.45	-2.36	-2.97	-4.20
	SNR_gain_dB	3.23	3.63	3.33	3.09	2.90	2.76	1.79
	SDR dB	4.76	5.34	4.27	3.86	3.63	3.35	2.89
Right side	SNR_in_ref_sensor_dB	-1.88	-1.88	-1.88	-1.88	-1.88	-1.88	-1.88
	SNR_out_dB	2.67	2.97	3.11	3.02	2.78	2.30	0.33
	SNR_gain_dB	4.54	4.84	4.99	4.90	4.66	4.17	2.21
	SDR dB	7.69	7.90	8.48	8.76	8.87	8.49	6.23

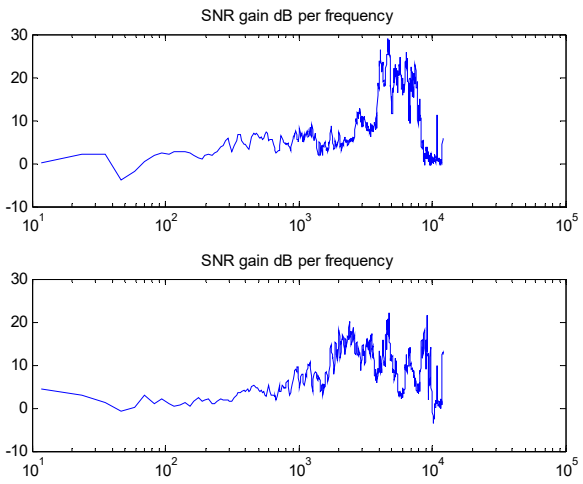
The two main conclusions from the table 5.3 are:

- Because of the mismatch between the HRTFs used for the design of the LCMV null-space blocking matrix and the recordings used for beamforming, and because the case of reverberant signals is always more challenging than the case using HRTF-filtered signals (presence of echoes), even when there is no target DOA mismatch the performance is significantly reduced compared to Table 5.2, especially for the left side beamformer. Nevertheless, a 3-5 dB SNR gain can be achieved. But the left side output is distorted from target leakage, as indicated by the lower SDR.
- The performance of the LCMV blocking matrix remains nearly constant for DOA mismatches up to 35-45 degrees, which indicates that the LCMV null-space blocking matrix performs as expected.

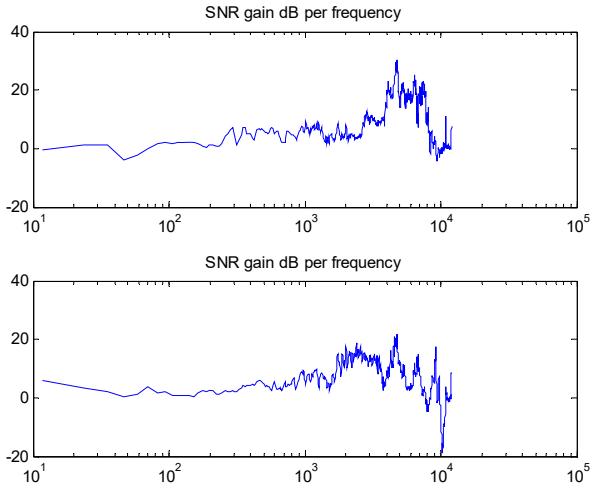
The frequency dependent SNR_gain and SDR curves for the GSC structure using the LCMV with null-space design as the blocking matrix for 2+1 configuration with HRTF-generated signals are shown in Figure 5.39 and Figure 5.40 respectively.



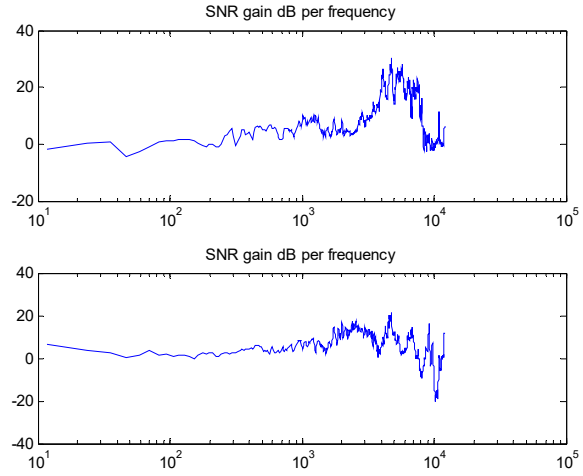
a



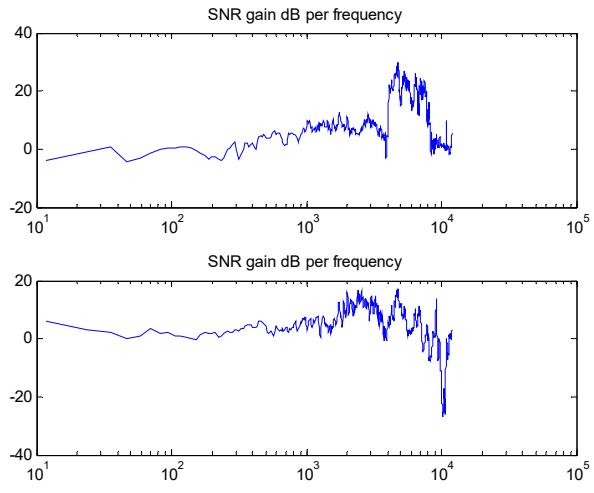
b



c

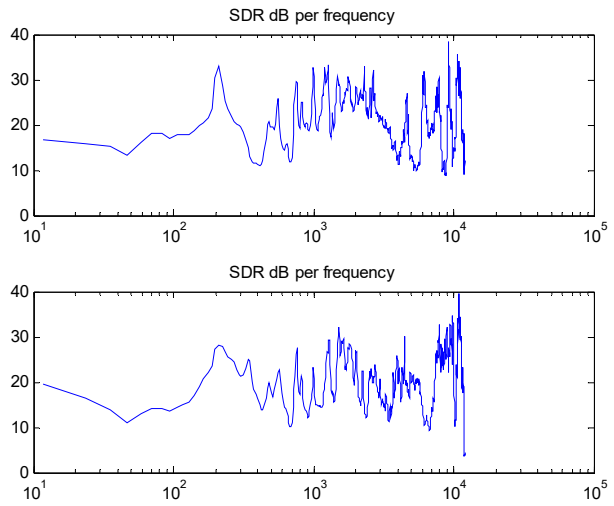


d

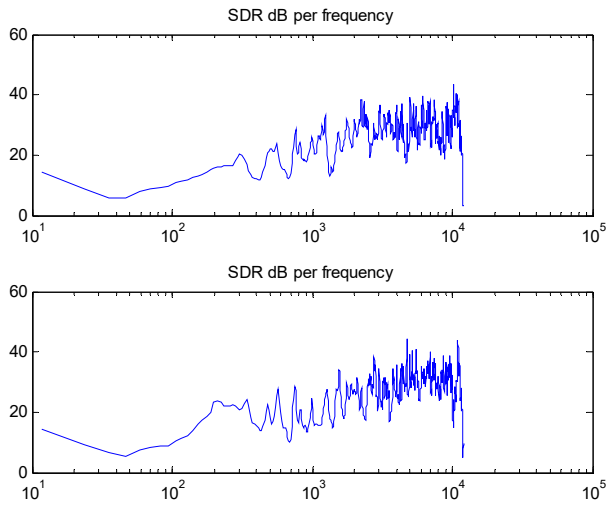


e

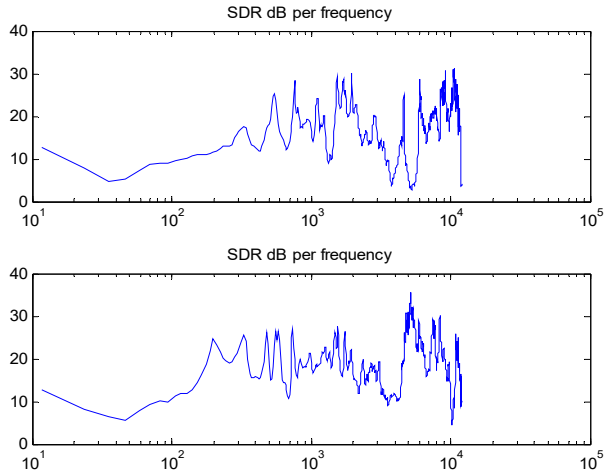
Figure 5.39 SNR_gain curves for the GSC system with LCMV null-space blocking matrix, 2+1 scenario, with HRTF generated signals, for target direction at a) 0, b) 10, c) 20, d) 30 and e) 45 degrees



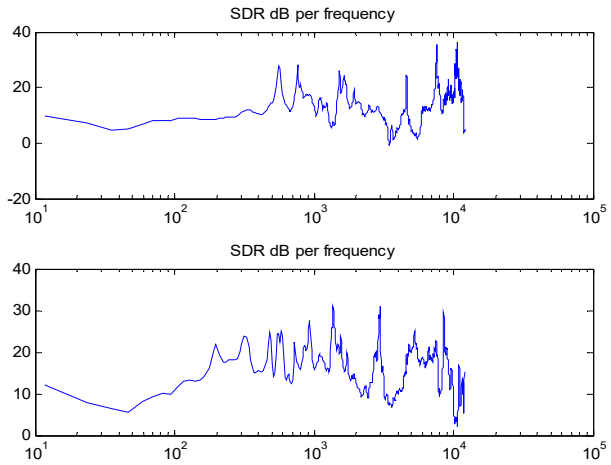
a



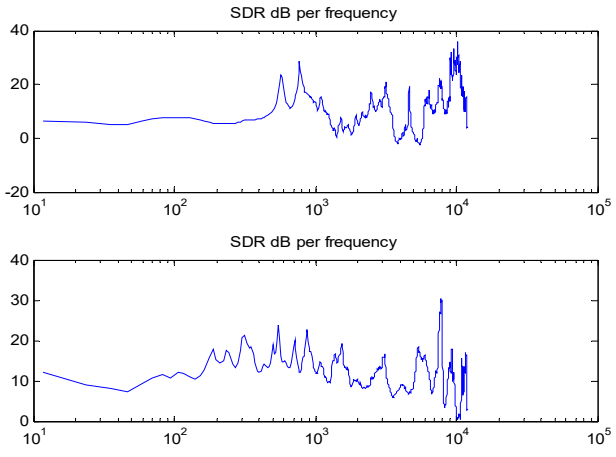
b



c



d



e

Figure 5.40 SDR curves for the GSC system with LCMV null-space blocking matrix, 2+1 scenario, with HRTF generated signals, for target direction at a) 0, b) 10, c) 20, d) 30 and e) 45 degrees

5.2.3 Evaluation of ADMA Blocking Matrix for 2+1 Configuration

Adaptive back-to-back first order differential microphone beamformers in target canceling mode were also used for the blocking matrix. For the 2+1 configuration (3 microphone signals available), as previously mentioned this could lead to three pairs of 1st order ADMAs:

- Using local front and local rear microphones (monaural ADMA)
- Using local front and remote front microphones (binaural ADMA)
- Using local rear and remote front microphones (binaural ADMA).

However, the 2nd and 3rd cases are likely to be highly redundant, so as in the classic "Griffiths-Jim" blocking matrix we only use *N*-1 blocking matrix outputs for *N* microphones, and only the 1st and 2nd cases were considered.

For the noise field configuration, two interferers were used. First they were positioned in directions such that they will not significantly affect the convergence of the ADMA adaptive nulls. For example, if the target DOA is to be varied from 0 to 35 degrees in the right front quadrant, two interferers can be positioned at 270 and 225 degrees in the back left quadrant. In this case, the interferers will not affect the monaural ADMA null positioning in the frontal half-hemisphere near the 0 to 35 degrees range and its symmetric range 0 to -35 degrees, nor will it affect the binaural ADMA null positioning in the right half-hemisphere near the 0 to 35 degrees range and its symmetric range 180 to 145 degrees. Using appropriate initial β values for the near frontal target (β=0 for the monaural ADMA for endfire direction, and β=1 for the binaural ADMA for broadside direction), this first case does not require any additional constraint in the range of allowed β values in the ADMAs. The noise signals in this experimental configuration were thus two interferers and diffuse-like background noise at rms level 0.5.

To evaluate only the performance of the ADMA blocking matrix, as in the previous sections the reference microphone input signal was used as the fixed beamformer output such that the target signal remains undistorted regardless of the conditions tested. Simulations were first done with reverberant signals. The target direction was set at 0, 15, 25 and 35 degrees. The results

for the GSC using the ADMA blocking matrix with the ranges of allowed values 'β_{min}=0, β_{max}= 1' and 'β_{min}=-0.4, β_{max}= 1.4' are shown respectively in Table 5.4 and Table 5.5.

Table 5.4 Results for the GSC using ADMA blocking matrix, for 2+1 configuration with reverberant recordings with the range 'β_{min}=0, β_{max}= 1'

	Target direction:	0	5	15	25	35
Left side	SNR_in_ref_sensor_dB	-3.56	-3.99	-5.17	-6.15	-6.88
	SNR_out_dB	2.69	1.52	1.25	0.71	0.10
	SNR_gain_dB	6.24	5.51	6.42	6.86	6.98
	SDR dB	7.90	4.77	4.89	4.65	4.14
Right side	SNR_in_ref_sensor_dB	0.50	0.50	0.50	0.50	0.50
	SNR_out_dB	2.24	1.02	1.02	0.85	0.55
	SNR_gain_dB	1.74	0.52	0.52	0.35	0.05
	SDR dB	10.85	5.91	5.64	5.16	4.99

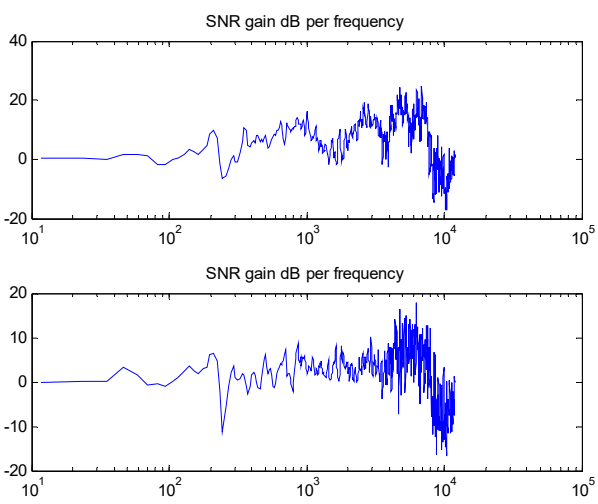
Table 5.5 Results for the GSC using ADMA blocking matrix, for 2+1 configuration with reverberant recordings with the range 'β_{min}=-0.4, β_{max}= 1.4'

	Target direction:	0	5	15	25	35
Left side	SNR_in_ref_sensor_dB	-3.56	-3.99	-5.17	-6.15	-6.88
	SNR_out_dB	3.45	1.79	1.35	0.76	0.13
	SNR_gain_dB	7.00	5.77	6.52	6.91	7.01
	SDR dB	8.88	5.01	5.04	4.79	4.28
Right side	SNR_in_ref_sensor_dB	0.50	0.50	0.50	0.5	0.50
	SNR_out_dB	3.42	1.82	1.64	1.33	1.00
	SNR_gain_dB	2.92	1.32	1.14	0.83	0.50
	SDR dB	11.85	6.08	5.78	5.34	5.20

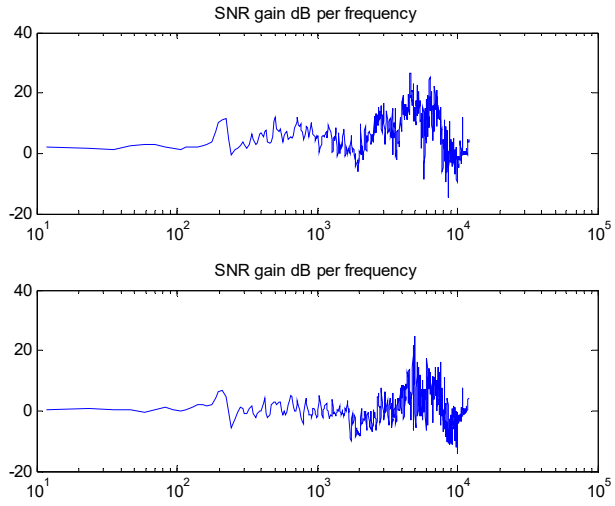
Considering the initial values of β (initially set for a frontal direction) and the fact that there is always an active target, from the results we see good robustness in ADMA convergence to the target direction (i.e., good SNR gain on the side with low input SNR, for different target DOA).

With the range 'β_{min}=-0.4, β_{max}= 1.4' the system works slightly better than when the range 'β_{min}=-0.0, β_{max}= 1.0' is used. By listening to the wave files, we find that there is target signal leakage in the blocking matrix output when the SDR is below 7 dB.

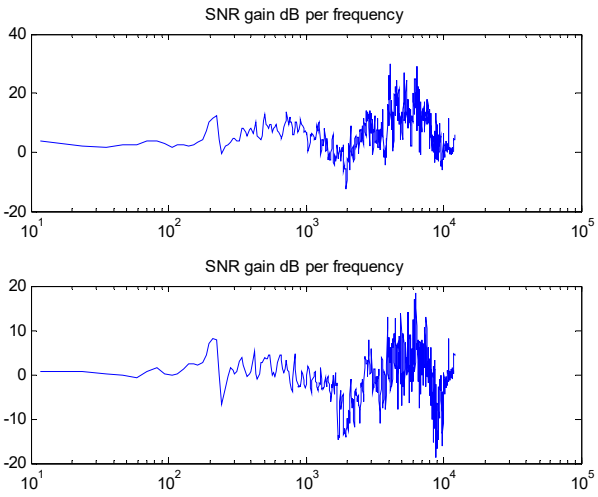
The frequency dependent SNR_gain and SDR curves for the GSC structure using ADMA blocking matrix ('β_{min}=-0.4, β_{max}= 1.4'), for the 2+1 configuration with reverberant recordings are shown in Figure 5.41 and Figure 5.42 respectively.



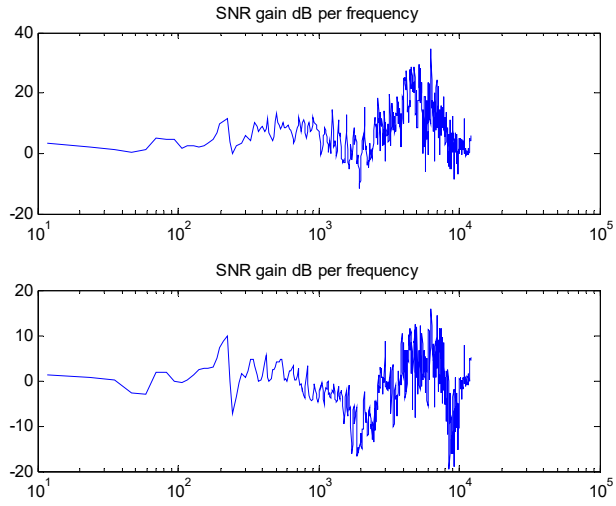
a



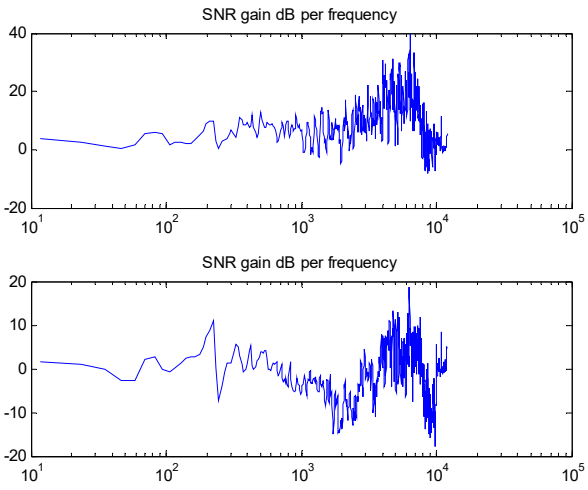
b



c



d



e

Figure 5.41 SNR_gain for the GSC system with ADMA blocking matrix, reverberant recordings, target direction at a) 0, b) 5, c) 15, d) 25 and e) 35

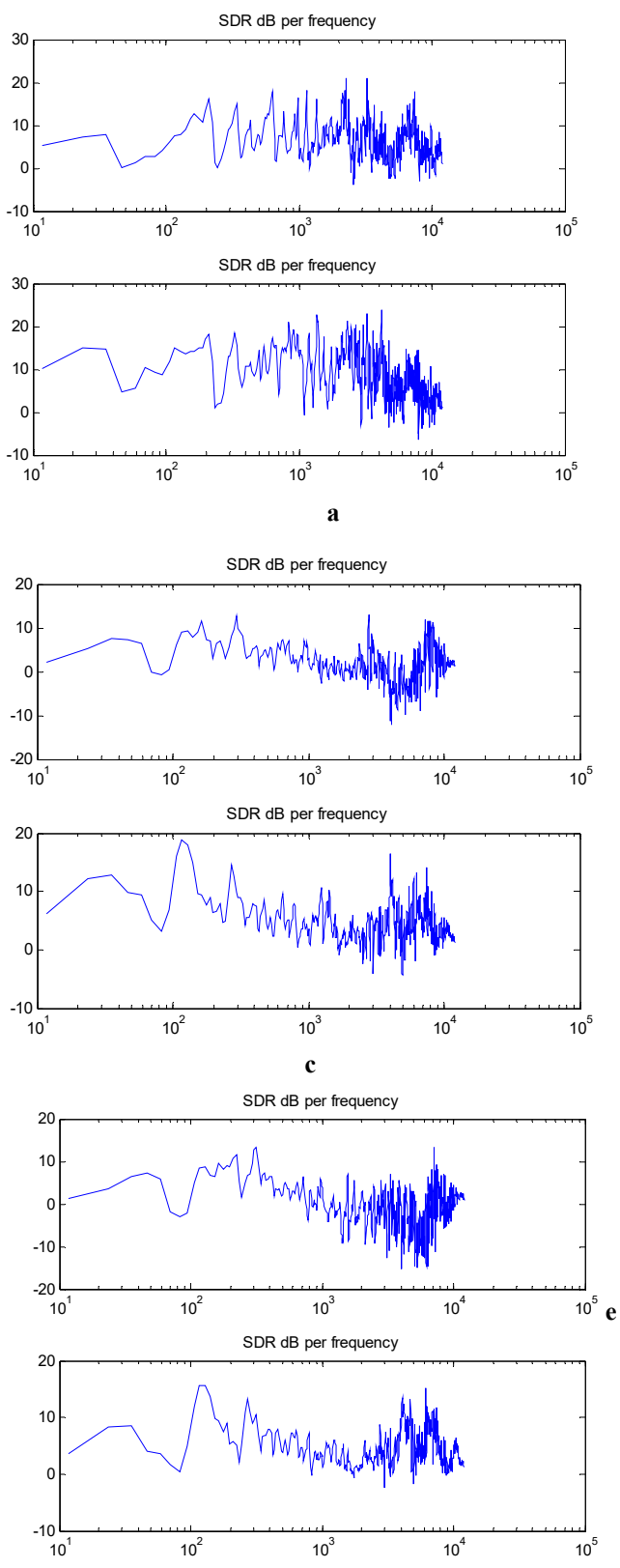


Figure 5.42 SDR for the GSC system with ADMA blocking matrix, reverberant recordings, target direction at a) 0, b) 5, c) 15, d) 25 and e) 35

For cases where the interferers may be closer to directions where they can create a conflict for the monaural and binaural ADMA blocking matrix to position their null in the target direction, it is possible to constrain or restrict the range of β values near the expected β (which is computed from the expected DOA) and consequently it may be possible to prevent the ADMA from converging to the interferer location. In order to obtain an assessment for such interferer directions out of the "safe region", the ADMA blocking matrix was tested for different pairs of interferer directions. The experimental configuration consists of a frontal target source, two interferers and diffuse-like background noise at rms level 0.5. First, the results of the GSC system using ADMA blocking matrix with the range 'β_{min}=-0.4, β_{max}= 1.4' (no constrained or restricted range) with reverberant recordings are shown in Table 5.6. The bold written angles indicate cases with a potential for conflict in the null convergence.

Table 5.6 Results of the GSC system using ADMA blocking matrix with the range 'β_{min}=-0.4, β_{max}= 1.4' with reverberant recordings

	Interferer direction	225 **270**	225 **135**	225 **90**	**135** **90**	**135** **45**
Left side	SNR_in_ref_sensor_dB	-3.56	-1.86	-1.94	0.76	0.16
	SNR_out_dB	3.45	3.14	2.84	3.91	2.37
	SNR_gain_dB	7.00	5.00	4.78	3.15	2.21
	SDR dB	8.88	9.99	10.14	11.00	10.93
Right side	SNR_in_ref_sensor_dB	0.50	-1.92	-1.88	-3.37	-3.45
	SNR_out_dB	3.42	2.60	1.34	2.82	1.48
	SNR_gain_dB	2.92	4.52	3.22	6.20	4.93
	SDR dB	11.85	10.79	10.75	8.29	8.60

Considering the the SNR gains for the sides with lowest input SNR in Table 5.6, the ADMA blocking matrix overall still works fairly well because of the initial conditions on β (set for the target direction) and because there is always an active target at 0 degrees (so the null does not move away from the target direction). But in practice this setup would likely not be stable.

Next, the β range was constrained from 0.0 to 0.2 for the monaural ADMA and from 0.8 to 1.0 for the binaural ADMA, restricting the null positining near 0 degrees. A slightly wider range

of β from -0.1 to 0.3 for the monaural ADMA and 0.7 to 1.1 for the binaural ADMA was also considered. The results of the ADMA blocking matrix while using the restricted β values with reverberant recorded signals are listed in Table 5.7.

Table 5.7 Comparison of ADMA blocking matrix while using restricted β values for monaural and binaural configurations with reverberant recordings

	Interferer Direction	135 90			135 45		
	β _min and β _max	-0.4 1.4	Mon. β: 0 to 0.2 Bin. β: 0.8 to 1	Mon. β: -0.1 to 0.3 Bin. β: 0.7 to 1.1	-0.4 1.4	Mon. β: 0 to 0.2 Bin. β: 0.8 to 1	Mon. β: -0.1 to 0.3 Bin. β: 0.7 to 1.1
Left side output	SNR_in,ref. sen.	0.76			0.16		
	SNR out	3.91	3.19	3.65	2.37	2.28	2.50
	SNR_gain	3.15	2.43	2.90	2.21	2.11	2.33
	SDR (dB)	11.00	12.51	12.93	10.93	11.37	12.03
Right side output	SNR_in,ref. sen.	-3.37			-3.45		
	SNR out	2.82	3.73	3.83	1.48	2.63	2.51
	SNR_gain	6.20	7.10	7.20	4.93	6.08	5.96
	SDR (dB)	8.29	9.57	9.50	8.60	9.70	9.49

As seen from Table 5.7, constraining the β range near the expected target direction does help performance (i.e., the SNR gains for the sides with lowest input SNR), making the ADMA blocking matrix more robust in the presence of interferers that can potentially affect ADMA convergence.

Next, the same scenarios were also evaluated for the case of HRTF-generated signals, which is a simpler, less-realistic case because there are then no echoes to deal with. The results of the GSC system using the ADMA blocking matrix for 2+1 configuration with the non-restricted range 'β_{min}=-0.4, β_{max}= 1.4' are shown in Table 5.8. The results of the ADMA blocking matrix while using restricted β values for the monaural and binaural ADMA blocking matrix are listed in Table 5.9.

Table 5.8 Results of the GSC system using ADMA blocking matrix with the range 'β_{min}=-0.4, β_{max}= 1.4' with HRTF-generated signals

	Interferer direction	225 **270**	225 **135**	225 **90**	**135** **90**	**135** **45**
Left side	SNR_in_ref_sensor_dB	-3.50	-1.55	-1.47	3.09	2.75
	SNR_out_dB	3.16	3.91	3.24	3.02	2.15
	SNR_gain_dB	6.66	5.46	4.73	-0.07	-0.59
	SDR dB	7.52	9.53	9.59	7.04	7.13
Right side	SNR_in_ref_sensor_dB	2.23	-1.96	-1.96	-4.06	-4.07
	SNR_out_dB	3.95	2.93	1.09	1.78	1.22
	SNR_gain_dB	1.72	4.89	3.05	5.84	5.29
	SDR dB	7.69	9.08	7.06	7.41	8.33

As in the corresponding case with reverberant recorded signals (Table 5.6), considering the SNR gains for the sides with lowest input SNR in Table 5.8, the ADMA blocking matrix overall still works fairly well because of the initial conditions on β (set for the target direction) and because there is always an active target at 0 degrees (so the null does not move away from the target direction). But as mentioned earlier, in practice this setup would likely not be stable.

Table 5.9 Comparison of ADMA blocking matrix while using restricted β values with HRTF-generated signals

	Interferer Direction	135 90			135 45		
	β _min and β _max	-0.4 1.4	Mon. β: 0 to 0.2 Bin. β: 0.8 to 1	Mon. β: -0.1 to 0.3 Bin. β.: 0.7 to 1.1	-0.4 1.4	Mon. β: 0 to 0.2 Bin. β: 0.8 to 1	Mon. β: -0.1 to 0.3 Bin. β: 0.7 to 1.1
Left side output	SNR_in,ref. sen.	3.09			2.75		
	SNR out	3.02	3.46	3.97	2.15	2.55	2.82
	SNR_gain	-0.07	0.37	0.88	-0.59	-0.02	0.08
	SDR (dB)	7.04	6.43	7.09	7.13	6.31	7.02
Right side output	SNR_in,ref. sen.	-4.06			-4.07		
	SNR out	1.78	3.10	3.28	1.22	2.25	2.26
	SNR_gain	5.84	7.16	7.34	5.29	6.32	6.33
	SDR (dB)	7.41	9.26	9.28	8.33	9.34	9.58

As seen from Table 5.9, constraining the β range near the expected target direction does overall help performance for the case of reverberant recordings, making the ADMA blocking matrix more robust in the presence of interferers that can potentially affect ADMA convergence.

5.2.4 Evaluation of MVDR Fixed Beamformer Tuned for Diffuse Noise for 2+1 Configuration

This section evaluates the MVDR fixed beamformer tuned for diffuse noise for the 2+1 configuration with no pre-processing. The simulations were performed with reverberant recordings, and the correlation matrix of the MVDR beamformer was computed from the reverberant diffuse noise signal. Target DOA mismatch was not evaluated for this section, as improving robustness to target DOA mismatch is not the objective for the MVDR fixed beamformer tuned for diffuse noise. The target direction was constant at 0 degrees (frontal direction) and the background diffuse noise was at rms level 1.0. The GSC system with MVDR fixed beamformer was combined with the reference target-equalized Griffiths-Jim blocking matrix, described earlier (with equalization coefficients obtained from the target signals).

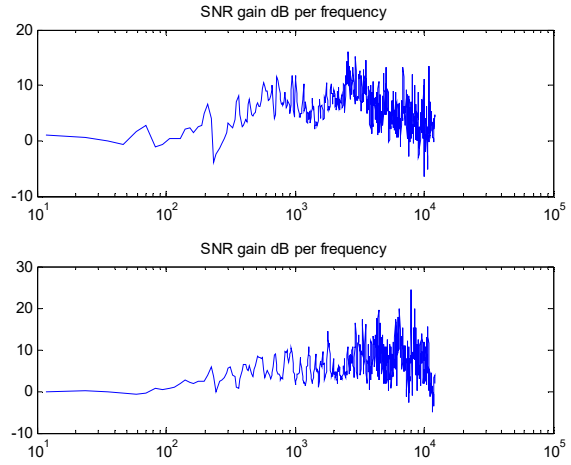
In addition to the SNR-gain and SDR, it was necessary to assess the diffuse noise SNR-gain and the interference SNR-gain compared with the reference setup where both the fixed beamformer and the blocking matrix are done with target-equalized signals (target-equalize and sum fixed beamformer, and target-equalized Griffiths-Jim blocking matrix). The results are listed in Table 5.10.

Table 5.10 Comparison of the GSC system using MVDR fixed beamformer tuned for diffuse noise and Griffiths-Jim blocking matrix with reference GSC structure, 2+1 configuration with reverberant recordings

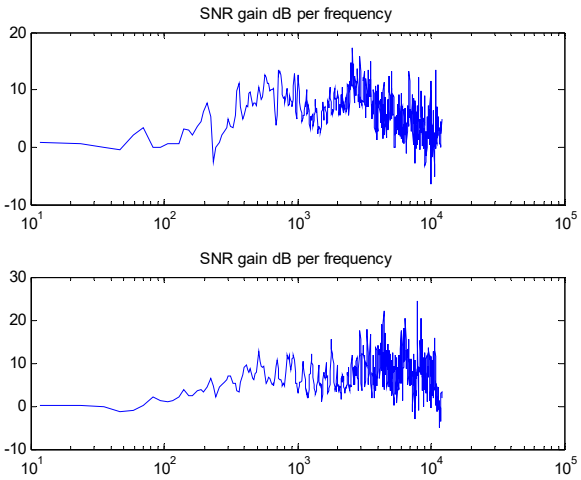
		Interferer at 225		Interferers at 90 and -90	
	Fixed beamformer type:	Target-equalized	MVDR	Target-equalized	MVDR
Left side	SNR_in_ref_sensor_dB	-3.04		-3.68	
	SNR_out_dB	3.29	4.34	1.21	2.72
	SNR_gain_dB	6.33	7.38	4.89	6.39
	backnoiseSNR_in_ref_sensor_dB	-0.04		-0.4	
	backnoiseSNR_out_dB	4.37	5.08	4.30	5.07
	backnoiseSNR_gain_dB	4.41	5.12	4.34	5.11
	SIR_in_ref_sensor_dB	-0.04		-1.13	
	SIR_out_dB	9.73	12.23	4.15	6.49
	SIR_gain_dB	9.78	12.27	5.28	7.62
	SDR dB	13.33	11.85	14.37	11.69
Right side	SNR_in_ref_sensor_dB	-0.66		-3.33	
	SNR_out_dB	3.03	4.24	1.07	2.66
	SNR_gain_dB	3.69	4.90	4.40	5.99
	backnoiseSNR_in_ref_sensor_dB	0.85		0.85	
	backnoiseSNR_out_dB	4.15	5.17	4.05	5.17
	backnoiseSNR_gain_dB	3.31	4.33	3.21	4.32
	SIR_in_ref_sensor_dB	4.59		-1.34	
	SIR_out_dB	9.29	11.20	4.11	6.24
	SIR_gain_dB	4.69	6.61	5.45	7.58
	SDR dB	14.64	11.98	14.13	11.27

The simulation was performed for two scenarios. The first scenario included the presence of only one interferer while the second scenario included two interferers. In both scenarios, the MVDR fixed beamformer performed better than the reference fixed beamformer (in terms of SNR, background noise SNR and SIR).

The frequency dependent SNR_gain curves for the GSC using the target equalized and sum fixed beamformer and the MVDR fixed beamformer are shown in Figure 5.43.a and Figure 5.43.b respectively.



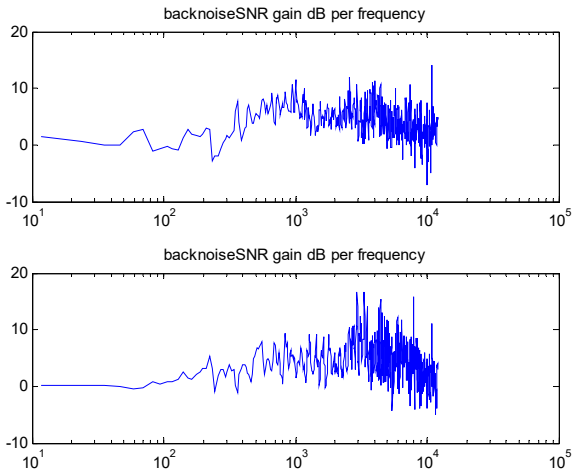
a



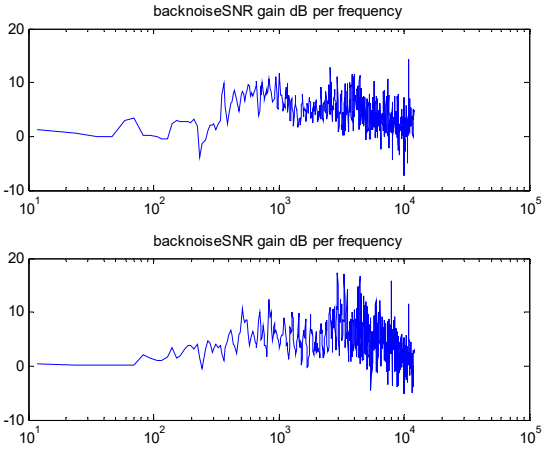
b

Figure 5.43 SNR_gain curves for GSC using a) target equalized and sum fixed beamformer and b) MVDR fixed beamformer

The frequency dependent background noise SNR-gain curves for the GSC using the target equalized and sum fixed beamformer and the MVDR fixed beamformer are shown in Figure 5.44.a and Figure 5.44.b respectively.



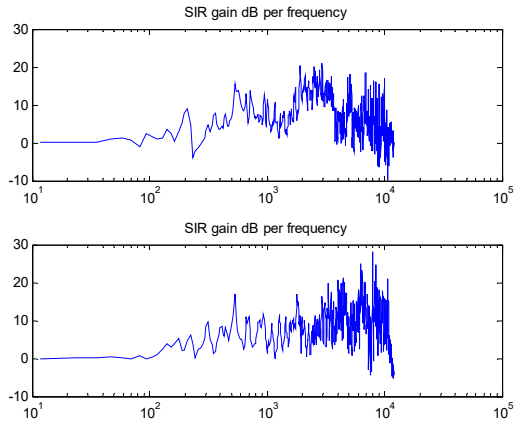
a



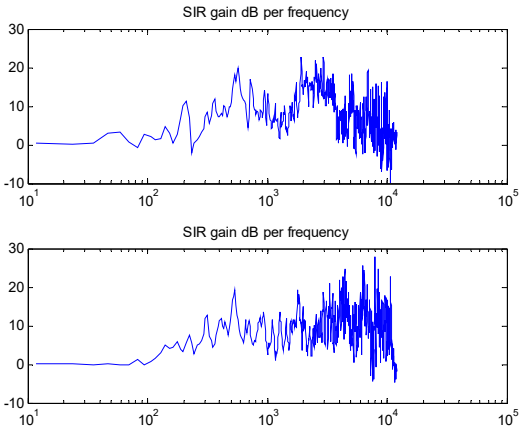
b

Figure 5.44 Background noise SNR_gain curves for GSC using a) target equalized and sum fixed beamformer and b) MVDR fixed beamformer

The frequency dependent SIR_gain curves for the GSC using the target equalized and sum fixed beamformer and the MVDR fixed beamformer are shown in Figure 5.45.a and Figure 5.45.b respectively.



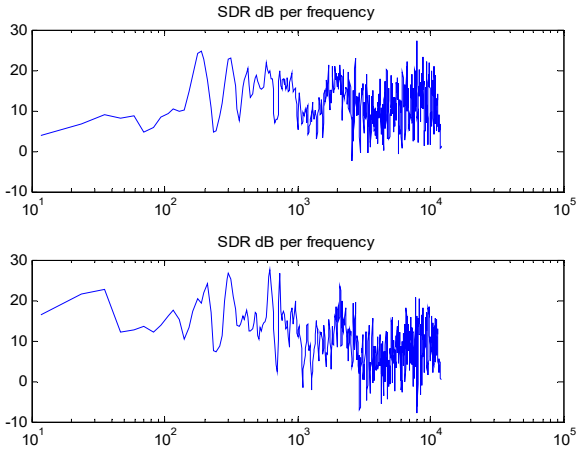
a



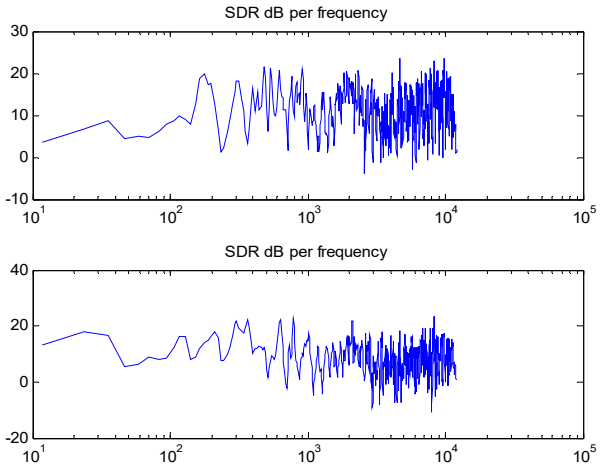
b

Figure 5.45 SIR_gain curves for GSC using a) target equalized and sum fixed beamformer and b) MVDR fixed beamformer

Finally, the frequency dependent SDR curves for GSC using the target equalized and sum fixed beamformer and the MVDR fixed beamformer are shown in Figure 5.46.a and Figure 5.46.b respectively.



a



b

Figure 5.46 SDR curves for GSC using a) target equalized and sum fixed beamformer and b) MVDR fixed beamformer

5.2.5 Evaluation of LCMV Fixed Beamformer for 2+1 Configuration

The proposed wide-beam fixed beamformer using an LCMV design was designed for a frontal target direction using anechoic HRTFs. To evaluate the wide-beam fixed beamformer performance in terms of robustness to target DOA mismatch, it was combined with a LCMV null-space wide-notch blocking matrix design, also using anechoic HRTFs. The comparison was done with a setup where the fixed beamformer is the reference target-equalize and sum (with equalization coefficients obtained from the target signals), while the blocking matrix is the same wide-notch LCMV null-space beamformer.

Simulations were first performed with HRTF-generated signals due to the offline design of the LCMV fixed beamformer and LCMV blocking matrix based on HRTFs. Subsequently simulations were also performed with reverberant signals, to evaluate the effect of HRTF mismatch. The noise signals in this experiment included two interferers at 90 and 225 degrees and diffuse-like background noise with rms level 0.5. The structure was tested for target directions at 0, 5, 10, 15, 20, 25, 30, 40 and 50 degrees. The results for the GSC using a wide-beam fixed beamformer and a wide-notch blocking matrix for the 2+1 configuration are listed in Table 5.11, for HRTF-generated signals. The corresponding results for the GSC with the target-equalize and sum fixed beamformer and wide-notch LCMV null-space blocking matrix are listed in Table 5.12.

Table 5.11 Results for GSC system with wide-beam fixed beamformer and wide-notch blocking matrix, 2+1 configuration, with HRTF-generated signals

	Target direction:	0	5	10	15	20	25	30	40	50
Left side	SNR_in_ref_sensor_dB	-1.47	-1.59	-2.74	-3.91	-4.80	-5.47	-6.06	-6.91	-7.46
	SNR_out_dB	7.23	7.32	6.83	5.45	4.41	3.63	2.50	0.38	-1.66
	SNR_gain_dB	8.70	8.91	9.58	9.35	9.21	9.10	8.55	7.30	5.80
	SDR dB	16.40	16.72	18.20	16.24	15.13	13.75	11.92	9.03	6.30
Right side	SNR_in_ref_sensor_dB	-1.96	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33
	SNR_out_dB	4.26	4.92	5.08	4.87	4.67	4.61	4.32	3.81	2.73
	SNR_gain_dB	6.22	6.24	6.40	6.19	5.99	5.93	5.64	5.14	4.06
	SDR dB	17.59	18.73	20.51	19.22	18.08	16.61	15.42	12.75	9.14

Table 5.11 shows that the performance of the GSC with wide-beam LCMV fixed beamformer and wide-notch LCMV blocking matrix is quite robust to DOA mismatch, up to at least 40 degrees of DOA mismatch.

Table 5.12 Results of the GSC with the target-equalize and sum fixed beamformer and wide-notch LCMV null-space blocking matrix, 2+1 configuration, with HRTF-generated signals

	Target direction:	0	5	10	15	20	25	30	40	50
Left side	SNR_in_ref_sensor_dB	-1.47	-1.59	-2.74	-3.91	-4.80	-5.47	-6.06	-6.91	-7.46
	SNR_out_dB	6.60	6.42	4.62	5.08	4.82	2.63	2.04	1.16	0.04
	SNR_gain_dB	8.07	8.01	7.36	8.99	9.62	8.10	8.10	8.07	7.50
	SDR dB	19.79	11.75	6.17	8.77	7.41	6.85	6.41	4.78	2.42
Right side	SNR_in_ref_sensor_dB	-1.96	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33	-1.33
	SNR_out_dB	5.95	5.98	4.99	5.18	4.85	4.23	3.29	3.46	2.42
	SNR_gain_dB	7.91	7.31	6.32	6.51	6.18	5.55	5.25	4.79	3.75
	SDR dB	18.47	12.66	9.16	10.49	9.77	8.22	8.06	7.18	5.58

Considering Table 5.11 and Table 5.12, for non-frontal target direction (DOA mismatch) the "GSC with LCMV wide-beam fixed beamformer and LCMV wide-notch blocking matrix" outperforms the "GSC with the target-equalize and sum fixed beamformer and wide-notch LCMV null-space blocking" in terms of SDR. This essentially means that the wide beam fixed beamformer reduces the distortion on the target signal for non-frontal target directions (DOA mismatch). However, in terms of SNR-gain, the wide beam fixed beamformer does not significantly improve the performance for non-frontal target directions (DOA mismatch).

The same setups are next tested with realistic reverberant recordings used as the input signals. The results for the GSC using a wide-beam fixed beamformer and a wide-notch blocking matrix for the 2+1 configuration are listed in Table 5.13, for recorded signals. The corresponding

results for the GSC with the target-equalize and sum fixed beamformer and wide-notch LCMV null-space blocking matrix are listed in Table 5.14.

Table 5.13 Results for GSC system with wide-beam fixed beamformer and wide-notch blocking matrix, 2+1 configuration, with reverberant recordings

	Target direction:	0	5	15	25	35	45
Left side	SNR_in_ref_sensor_dB	-1.94	-2.37	-3.55	-4.54	-5.26	-5.73
	SNR_out_dB	2.79	3.16	0.25	-1.00	-1.99	-2.68
	SNR_gain_dB	4.73	5.53	3.81	3.54	3.27	3.04
	SDR dB	6.11	5.93	4.68	4.26	3.97	3.61
Right side	SNR_in_ref_sensor_dB	-1.88	-1.88	-1.88	-1.88	-1.88	-1.88
	SNR_out_dB	5.37	5.37	3.25	3.10	2.82	2.30
	SNR_gain_dB	7.25	7.25	5.12	4.98	4.70	4.17
	SDR dB	8.69	8.33	8.65	8.87	8.93	8.50

Table 5.13 shows two effects:

- The SNR-gains are significantly weaker than for the case of HRTF-generated signals (Table 5.11), because of the HRTF mismatch between the LCMV designs and the actual conditions found in the recordings (and the fact that reverberant recordings are also more challenging by nature, with echoes). Nevertheless, SNR-gains between 3 and 5 dB can be achieved. But the left-side SDR is also low, indicating distortion in the output and target leakage in the output of the blocking matrix.
- The performance of the GSC with wide-beam LCMV fixed beamformer and wide-notch LCMV blocking matrix is quite robust to DOA mismatch in the case of recordings, up to at least 35 degrees of DOA mismatch.

Table 5.14 GSC with the target-equalize and sum fixed beamformer and wide-notch LCMV null-space blocking matrix, 2+1 configuration, with reverberant recordings

	Target direction:	0	5	15	25	35	45
Left side	SNR_in_ref_sensor_dB	-1.94	-2.37	-3.55	-4.54	-5.26	-5.73
	SNR_out_dB	2.53	2.93	1.30	-0.72	-2.35	-3.48
	SNR_gain_dB	4.47	5.31	4.86	3.82	2.91	2.24
	SDR dB	6.28	6.11	5.45	4.38	3.43	2.66
Right side	SNR_in_ref_sensor_dB	-1.88	-1.88	-1.88	-1.88	-1.88	-1.88
	SNR_out_dB	4.84	4.73	4.31	3.51	2.79	2.01
	SNR_gain_dB	6.72	6.61	6.18	5.39	4.67	3.89
	SDR dB	9.09	8.94	7.30	6.36	5.65	5.07

Comparing Table 5.13 and Table 5.14, we see that for reverberant recordings the SNR-gain values are marginally better for the GSC with wide-beam fixed beamformer and wide-notch blocking matrix, regardless of the DOA mismatch. However, the SDR values are significantly better for the right side GSC with wide-beam fixed beamformer and wide-notch blocking, regardless of the DOA mismatch.

5.2.6 Using Two Null-beamformers for the Wide-Notch Blocking Matrix Design

The concept of producing more than one wide-notch null-beamformer output in the blocking matrix is to provide the ANC system with more information/reference about the noise signals to be canceled from the fixed beamformer output. Using the two last columns of $V(\omega)$ in equation (3.4) for null-space design, we can obtain two different null-beamformers. In this section, we have tested this for the 2+1 configuration. As in some previous sections, for these tests the reference microphone was used as the fixed beamformer output (in other words there is no fixed beamformer used in these tests). The LCMV design with null-space was used with the constraint gain vector g=[1.0; 1.0] and the angles theta_constraints=[10,350].

The left side beampatterns (with the left frontal microphone used as the reference microphone) obtained when the last column from the eigenvector matrix $V(\omega)$ is used as the null beamformer for a 2+1 configuration in free field conditions and HRTF conditions are shown is Figure 5.47 and Figure 5.48, respectively (same as Figure 5.4 and Figure 5.5 in section 5.1.1). Also, the left side beampatterns obtained when the second last column of $V(\omega)$ is used as a null beamformer under the same conditions are shown is Figure 5.49 and Figure 5.50 respectively. Note that theoretically with 3 sensors and 2 constraints, $\hat{C}(\omega)\hat{C}^H(\omega)$ in (3.4) is a matrix of rank 2, so the first two columns of $V(\omega)$ should not correspond to null beamformers, and only the third column of $V(\omega)$ should be a null beamformer.

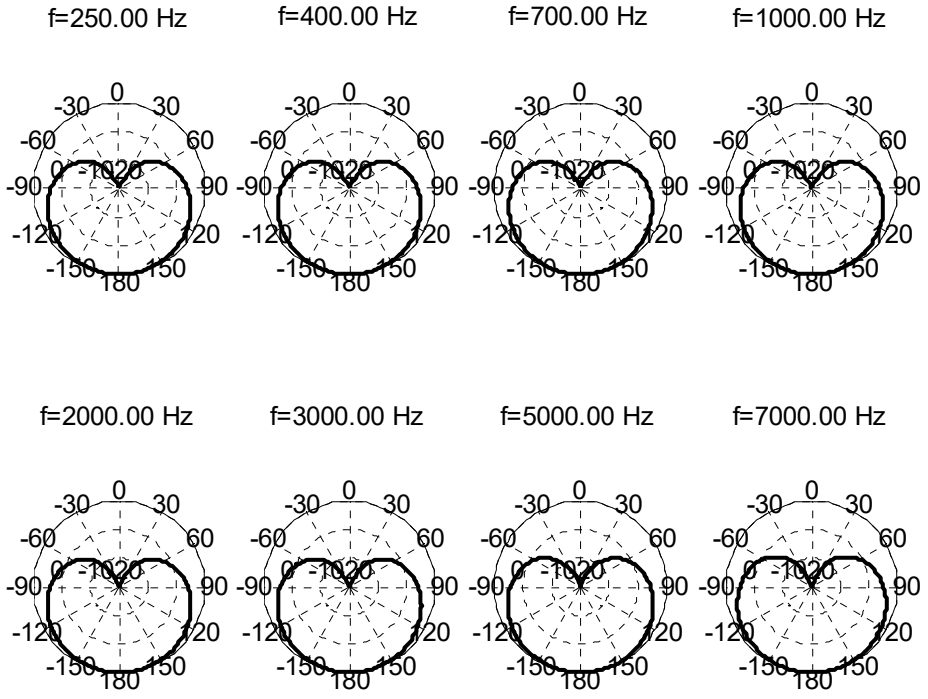


Figure 5.47 Left side beampatterns when the last column of matrix V is used as the null beamformer, 2+1 configuration, free field conditions

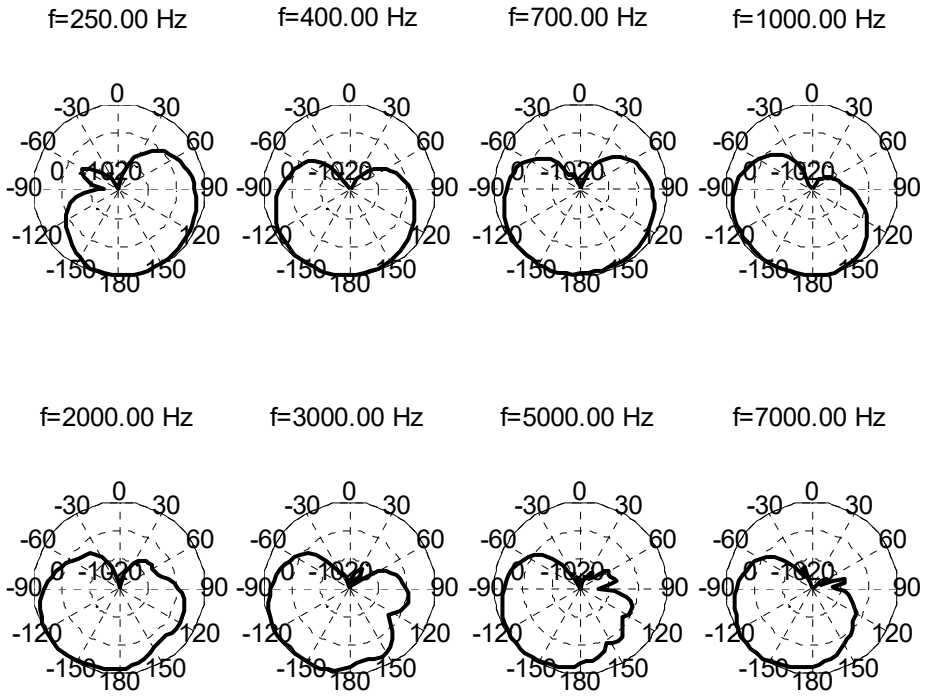


Figure 5.48 Left side beampatterns when the last column of matrix V is used as the null beamformer, 2+1 configuration, HRTFs conditions

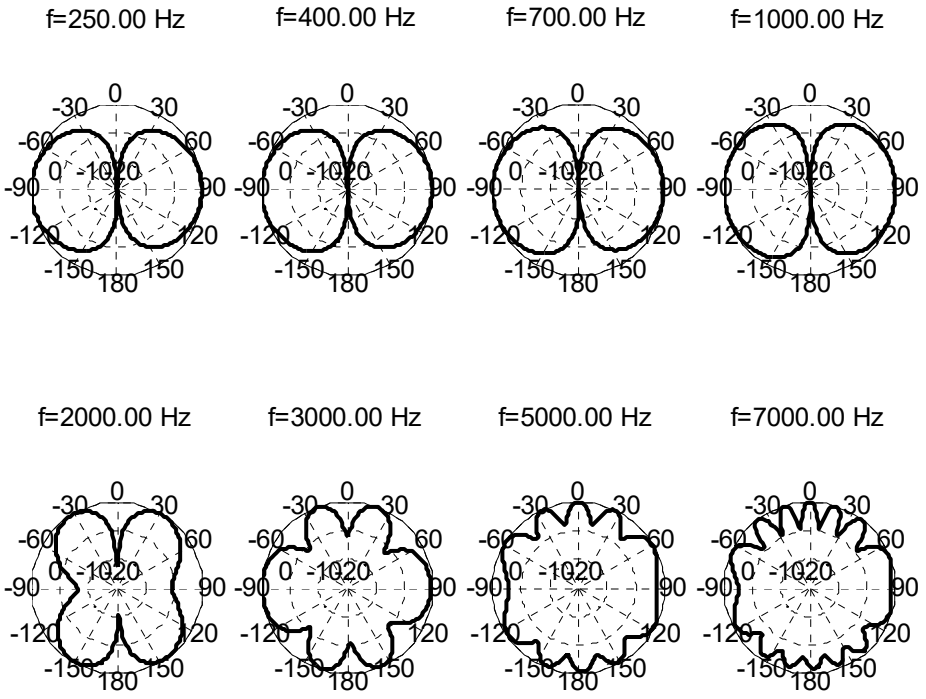


Figure 5.49 Left side beampatterns when the second last column of matrix V is used as the null beamformer, 2+1 configuration, free field conditions

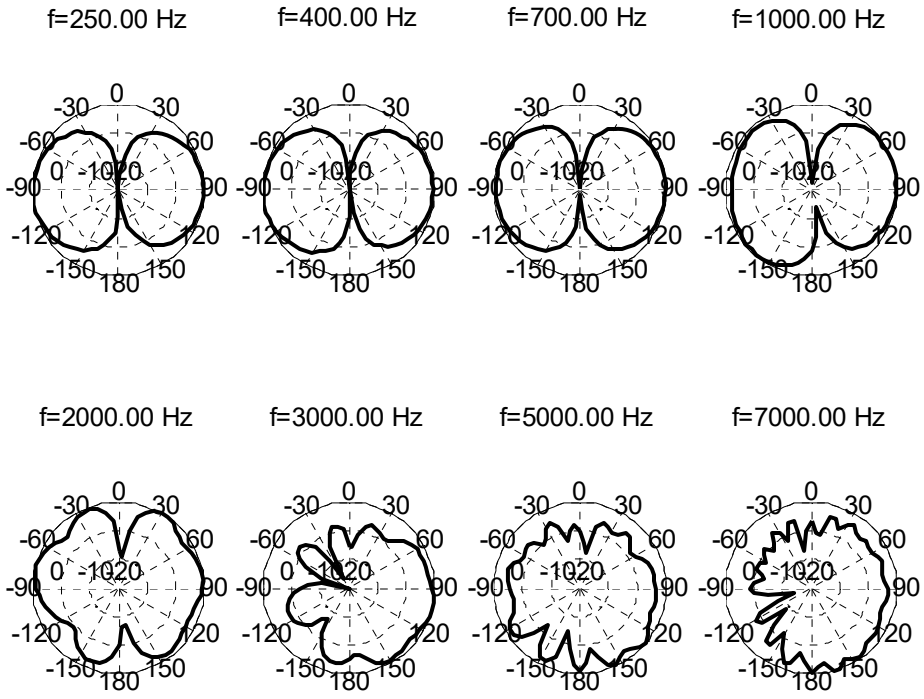


Figure 5.50 Left side beampatterns when the second last column of matrix V is used as the null beamformer, 2+1 configuration, HRTFs conditions

It can be seen from Figure 5.49 and Figure 5.50 (when using the second last column of $V(\omega)$) that the null is much more narrow compared to Figure 5.47 and Figure 5.48 (when using the last column of $V(\omega)$). Moreover, there is no null in the beampatterns above 3 kHz in Figure 5.49 and Figure 5.50.

Nevertheless, simulation results were performed using these two LCMV null-space designs for the blocking matrix. The noise signals in this experiment included two interferers at 270 and 225 degrees and diffuse-like background noise with rms level 0.5. The structure was tested for target directions at 0, 10, 20 and 30 degrees. The results are listed in Table 5.15, for HRTF-generated signals.

Table 5.15 Results for GSC using two null-beamformers from LCMV null-space design, 2+1 configuration, HRTFs conditions, interferers at 225 and 270 degrees

	Target direction	0		10		20		30	
	No. of BM outputs	1	2	1	2	1	2	1	2
Left side	SNR_in_ref_sensor	-3.51		-4.78		-6.84		-8.10	
	SNR_out_dB	4.36	5.13	3.66	4.54	1.40	-0.14	-0.38	-2.43
	SNR_gain_dB	7.87	8.64	8.44	9.32	8.24	6.70	7.72	5.67
	SDR dB	18.80	4.37	22.86	10.23	16.31	5.48	12.78	5.13
Right side	SNR_in_ref_sensor	2.59		3.23		3.23		3.23	
	SNR_out_dB	7.35	6.14	8.59	3.98	7.05	0.13	6.43	-1.47
	SNR_gain_dB	4.76	3.54	5.36	0.75	3.82	-3.10	3.20	-4.70
	SDR dB	19.90	6.96	24.15	5.59	18.90	3.80	16.26	3.02

As expected, the results show that using an additional null-beamformer from the LCMV null-space design in the 2+1 configuration does not improve the performance, because the second null-beamformer used in the experiment either has a narrow notch or no notch at all for some frequencies (leading to target leakage).

The next results listed in Table 5.16 are for the same setup but with interferers at 90 and 225 degrees, to verify if the same trends are obtained for another interferer scenario.

Table 5.16 Results for GSC using two null-beamformers from LCMV null-space design, 2+1 configuration, HRTFs conditions, interferers at 225 and 90 degrees

	Target direction	0		10		20		30	
	No. of BM outputs	1	2	1	2	1	2	1	2
Left side	SNR_in_ref_sensor	-1.47		-2.74		-4.80		-6.06	
	SNR_out_dB	7.35	4.19	6.90	3.58	4.43	-0.11	2.45	-2.00
	SNR_gain_dB	8.81	5.66	9.64	6.32	9.24	4.69	8.51	4.05
	SDR dB	16.18	5.20	17.56	9.29	14.52	5.58	11.33	4.76
Right side	SNR_in_ref_sensor	-1.96		-1.33		-1.33		-1.33	
	SNR_out_dB	4.14	4.96	5.04	2.82	4.69	-0.01	4.36	-1.94
	SNR_gain_dB	6.10	6.29	6.37	4.14	6.01	1.32	5.68	-0.62
	SDR dB	17.45	5.71	20.36	4.28	18.23	2.86	15.49	2.33

From Table 5.16, we obtain the same conclusion as previously with Table 5.15: the binaural beamformer performs worse when two null-beamformers from the LCMV null-space design are used in the 2+1 configuration, as expected.

In addition to the LCMV null-space design, in Chapter 3 two other approaches were presented for null-beamformer design, with corresponding beampattern results presented in section 5.1.1: the constraint-based design using null-space, and the direct constraint-based design. So we now consider to combine two designs obtained from different methods to get a blocking matrix with two outputs. As previously, the first blocking matrix output is obtained from an LCMV null-space design using the last column of $V(\omega)$). The second output of the blocking matrix is obtained from a direct constraint-based design with a constraint gain vector g=[0.0; 0.0; 1.0] and theta_constraints=[10,350,180]. The beampatterns for the direct constraint-based design of the null-beamformer in the 2+1 configuration for free field conditions and HRTFs conditions are shown in Figure 5.51 and Figure 5.52 (which are actually the same as Figure 5.7 and Figure 5.8 from section 5.1.1).

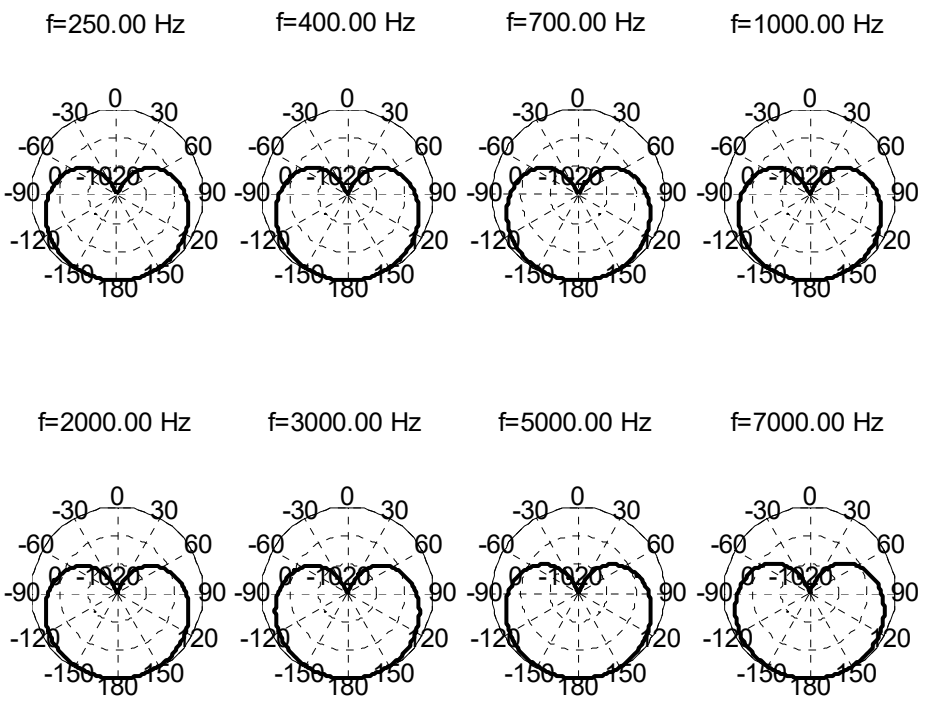


Figure 5.51 Beampatterns, direct constraint-based design of null-beamformer, 2+1 configuration, free field conditions (only shown for left side)

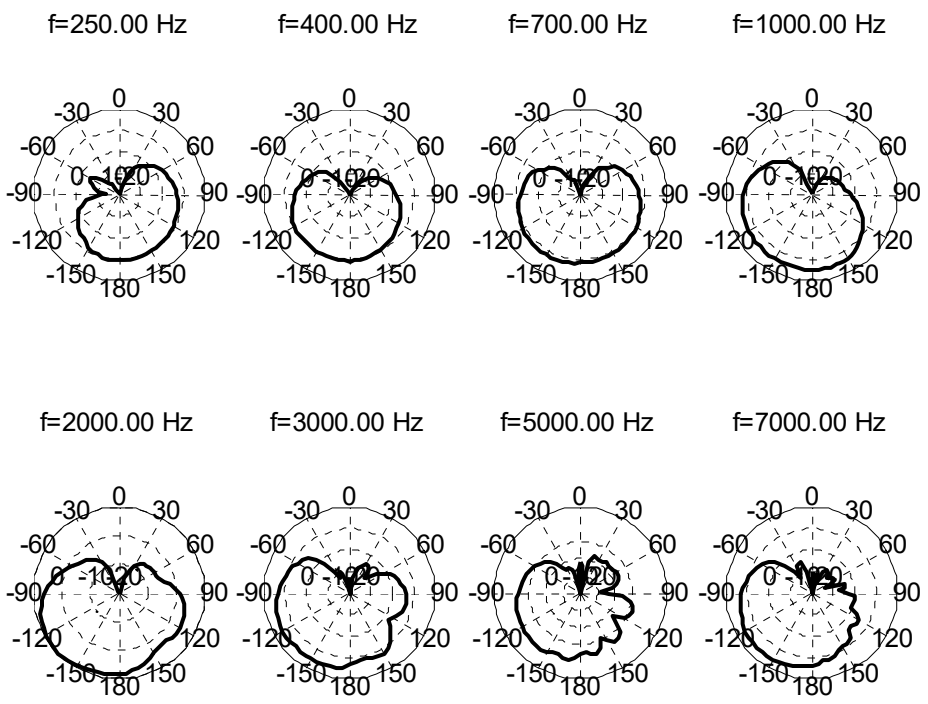


Figure 5.52 Beampatterns, direct constraint-based design of null-beamformer, 2+1 configuration, HRTFs conditions (only shown for left side)

These beampatterns show a wide-notch over a wide range of frequencies. However, the beampatterns in Figure 5.51 and Figure 5.52 for direct constraint-based design are very similar to the beampatterns for the LCMV null-space design (using last column of $V(\omega)$) in Figure 5.47 and Figure 5.48. So it is not certain that the two blocking matrix outputs will be sufficiently independent to be considered as two useful distinct outputs.

The results for the GSC using those two null-beamformers (LCMV null-space design and direct constraint-based design) in a 2+1 configuration, with interferers at 270 and 225 degrees, are listed in Table 5.17 for HRTF-generated signals.

Table 5.17 GSC results, two null-beamformers (LCMV null-space design and direct constraint-based design), 2+1 configuration, HRTFs conditions

	Target direction	0		10		20		30	
	No. of BM outputs	1	2	1	2	1	2	1	2
Left side	SNR_in_ref_sensor	-3.51		-4.78		-6.84		-8.10	
	SNR_out_dB	4.40	4.47	3.68	3.77	1.44	1.52	-0.26	-0.35
	SNR_gain_dB	7.91	7.98	8.47	8.55	8.29	8.36	7.84	7.75
	SDR dB	18.74	18.49	22.49	21.77	16.19	15.80	12.59	12.52
Right side	SNR_in_ref_sensor	2.59		3.23		3.23		3.23	
	SNR_out_dB	7.49	7.53	8.82	8.63	7.24	6.98	6.57	6.25
	SNR_gain_dB	4.76	4.93	5.59	5.40	4.01	3.76	3.34	3.03
	SDR dB	19.90	18.62	22.83	20.00	18.77	14.80	16.00	13.28

Based on Table 5.17, we see that using the two null-beamformers (LCMV null-space design and direct constraint-based design) produces essentially the same performance as using only the LCMV null-space design with the last column of $V(\omega)$, because the two designs are nearly identical.

In order to generate a 2nd null-beamformer output using a design that will not be nearly identical to the LCMV null-space design, a null-beamformer design using a direct LCMV-design is next considered. This design had not previously been considered in Section 5.1.1 for the blocking matrix, because the idea of using an LCMV design and its minimization of the trace of a signal or noise correlation matrix does not seem to be a good fit for a null-beamformer. For the direct null-beamformer LCMV design, the constraint gain vector (g) and the angles in theta_constraints are set to: g=[0.0; 1.0] and theta_constraints=[0,180]. The beampatterns for the direct null-beamformer LCMV design with 2+1 configuration in free field conditions and HRTFs conditions are shown in Figure 5.53 and Figure 5.54, respectively.

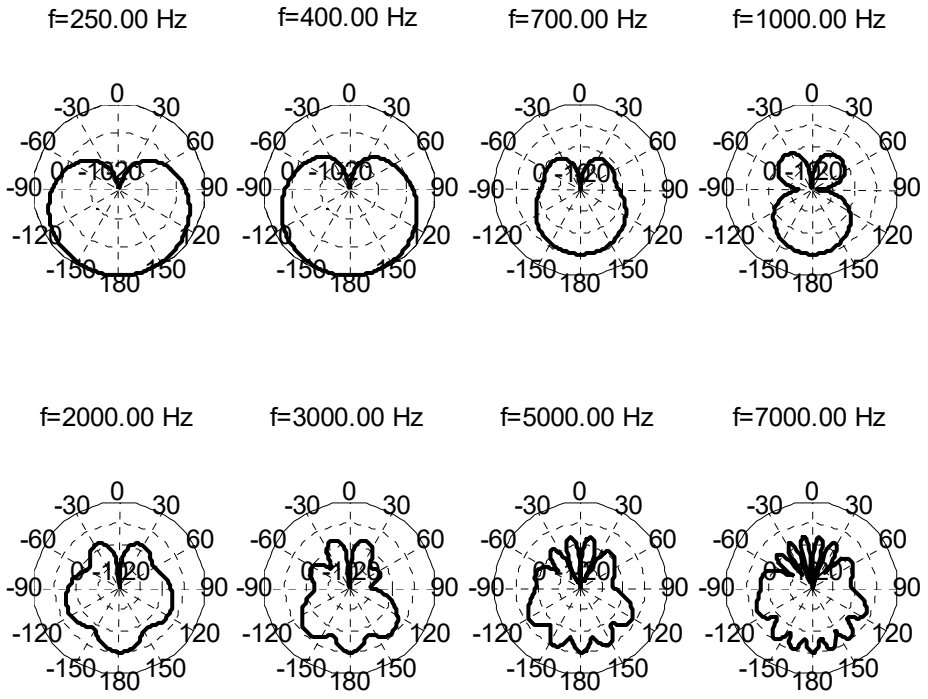


Figure 5.53 Beampatterns, direct null-beamformer LCMV design, 2+1 configuration, free field conditions (only shown for left side)

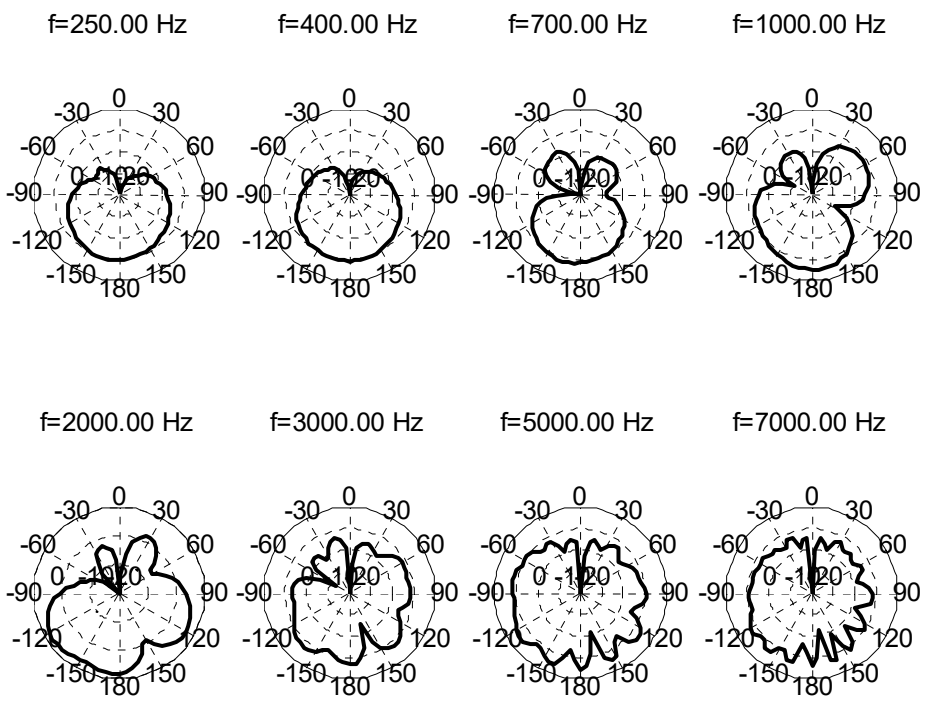


Figure 5.54 Beampatterns, direct null-beamformer LCMV design, 2+1 configuration, HRTFs conditions (only shown for left side)

Considering Figure 5.53 and Figure 5.54, the notch-width is only wide for very low frequencies and it is getting more narrow starting at 400 Hz. This is because the null constraint is only for angle zero and it is the only choice that can be considered for a 2+1 configuration while using a LCMV direct design. Indeed, for 3 sensors only 2 constraints can be used in the LCMV design, and only one constraint can be zero in order to avoid a zero gain beamformer. Therefore, the natural choice for the LCMV direct design is applying a null-constraint in the direction of the target signal and a constraint of unity in another opposite direction (e.g. 180 degrees for frontal target). Thus, based on the beampatterns observed in Figure 5.53 and Figure 5.54, the approach of using a direct null-beamformer LCMV design to generate a 2nd output in a wide-notch blocking matrix is not suitable.

Chapter 6 Conclusions and Future Work

6.1 Conclusions

The main purpose of this work package was to develop a binaural beamformer robust to DOA mismatch, while obtaining a good noise/interferers reduction. The proposed algorithms had to be able to work under real-life environments and were tested using both HRTF-generated signals and reverberant recorded signals, and processed with a realistic low-delay filter bank. The project was part of a research contract with the Siemens Audiologische Technik GmbH Group, which provided the requirements.

Different binaural beamforming configurations were proposed in Chapter 3, to be used as alternative algorithms for both the blocking matrix and the fixed beamformer units of the GSC beamformer typically used in the hearing aid industry. In Chapter 5, all the proposed algorithms were tested for a 2+1 microphone configuration. The algorithms were first assessed by their beampatterns at different frequencies, to see how they were successful to create a wide notch or a wide beam in the desired direction. The beampatterns showed that a wider notch blocking matrix and a wider beam fixed beamformer can be obtained when using the proposed LCMV and constraint-based designs compared to the reference target-equalized Griffiths-Jim blocking matrix and the target-equalized delay and sum fixed beamformer algorithms. Moreover, the directivity index of a fixed beamformer using a proposed MVDR design tuned for diffuse noise was shown to be improved compared to the reference target-equalized delay and sum fixed beamformer and other fixed beamformers.

An experimental evaluation was then performed in section 5.2 to investigate the performance of the considered algorithms in complex noisy acoustic environments including reverberation, diffuse-like noise and multiple directional interfering sources. Different metrics were introduced in Chapter 4 to evaluate the performance. The metrics include the SNR-gain, the speech to distortion ratio (SDR), the SIR-gain and the background noise SNR gain. The wide-notch blocking matrix approach outperformed the target-equalized Griffiths-Jim approach both in terms of SNR-gain and SDR for cases with DOA mismatch, i.e., when the target direction for the

design did not correspond to the actual target direction in the simulations. In the case of reverberant recordings, where there is also the effect of HRTFs mismatch between the design and the simulations, the proposed wide-notch approach still worked reasonably well and showed to be robust to DOA mismatch, for a design based on a frontal target.

To provide robustness to DOA mismatch, the use of an adaptive ADMA-based blocking matrix was tested, and it was found that initializing the β value based on the expected target direction significantly helped the binaural and monaural ADMA algorithms to converge to the correct direction, in target cancelation mode. Moreover, for cases where there were no interferer preventing the ADMAs to position a null in the target direction, allowing a wider β range was found to make the results slightly better. However, restrictions were then set on the allowed β values so that the null can only be positioned near a direction where the target is expected to be. This is to reduce the impact of interferers that can potentially create a conflict with the target when the monaural and binaural ADMAs attempt to position their null in the target direction. Using this restricted range of β values has also shown to improve the performance. Overall, the results of the adaptive ADMA-based blocking have shown to be robust to target DOA, for either HRTF-generated signals or reverberant signals.

By using the proposed MVDR fixed beamformer tuned for diffuse noise, for a frontal target direction the overall GSC performance was found to be slightly better (approximately 1 dB) in terms of SNR-gain, SIR-gain and background noise SNR-gain.

The wide-beam LCMV fixed beamformer showed a slightly better SNR-gain performance compared to the reference target-equalize and sum fixed beamformer when HRTF-generated signals were used for testing, however with the more realistic case of reverberant recordings there was no significant different in SNR-gain performance. In terms of SDR, for cases with significant DOA the wide-beam LCMV fixed beamformer outperformed the reference fixed beamformer for one of the sides (for the other side the performance was similar).

Finally, some ideas to produce more than one output in a wide-notch blocking matrix for a 2+1 microphone configuration were implemented, in order to provide more reference input signals to the ANC unit. Using the last two columns of $V(\omega)$ from equation (3.4) was first considered in a null-space design. However, it did not produce acceptable results because the second null-beamformer did not have a wide enough notch. A constraint-based direct design for the second null-beamformer was then used. However, this design was too redundant with the null-space design of the first null-beamformer (using only last column of $V(\omega)$), so it did not lead to improvements. At last, a direct LCMV design was attempted for the null-beamformer, but it did not lead to a wide notch beampattern. So the problem of producing more than one output in a wide-notch blocking matrix for the 2+1 microphone configuration remains open.

6.2 Suggestions for Future Work

Some possible items worth to be considered in future efforts related to this thesis are listed below:

- To test the proposed algorithms with dynamic (moving) sources;

- To perform the simulations with an adaptive ANC unit rather than the fixed ANC which has been used in this thesis, based on a least-squares solution computed from the whole recording;

- To develop and test multi-tap designs for the frequency domain/subband algorithms considered in this thesis: widebeam LCMV fixed beamformer, MVDR diffuse noise fixed beamformer, wide notch LCMV blocking matrix, blocking matrix ADMA (and in general for MVDR, LCMV, ADMA beamformers);

- To perform subjective tests to truly evaluate the performance of the algorithms proposed in the thesis, with people having normal hearing and people with impaired hearing.

References

[1] Puder, H. (2006). Adaptive signal processing for interference cancellation in hearing aids. *Signal Processing*, *86*(6), 1239-1253.

[2] Schaub, A. (2008). *Digital hearing aids*. New York: Thieme Medical Publisher.

[3] Gordy, J. D., Bouchard, M., & Aboulnasr, T. (2008, May). Beamformer performance limits in monaural and binaural hearing aid applications. *Canadian Conference on Electrical and Computer Engineering, 2008. CCECE 2008.* 381-386.

[4] Dillon, H. (2012). *Hearing aids* (2nd Edition ed.) Thieme Publishers.

[5] Glyde, H., Cameron, S., Dillon, H., Hickson, L., & Seeto, M. (2013). The effects of hearing impairment and aging on spatial processing. *Ear and Hearing, 34*(1), 15-28.

[6] Kawamura, A., Thanhikam, W., & Iiguni, Y. (2012). Single channel speech enhancement techniques in spectral domain. *International Scholarly Research Notices (ISRN) Mechanical Engineering, Volume 2012 (2012), Article ID 919234*

[7] Benesty, J., Chen, J., & Huang, Y. (2008). *Microphone array signal processing,* Springer Science & Business Media.

[8] Flanagan, J., Johnston, J., Zahn, R., & Elko, G. (1985). Computer-steered microphone arrays for sound transduction in large rooms. *The Journal of the Acoustical Society of America, 78*(5), 1508-1518.

[9] Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE, 57*(8), 1408-1418.

[10] McCowan, I. (2001). Microphone arrays: A tutorial. Queensland University, Australia*,* 38p. https://www.idiap.ch/~mccowan/arrays/tutorial.pdf (accessed July 2015)

[11] Doclo, S., Gannot, S., Moonen, M., & Spriet, A. (2008). Acoustic beamforming for hearing aid applications. *Handbook on Array Processing and Sensor Networks,* 269-302.

[12] Frost III, O. L. (1972). An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE, 60*(8), 926-935.

[13] Griffiths, L. J., & Jim, C. W. (1982). An alternative approach to linearly constrained adaptive beamforming. *IEEE Transactions On Antennas and Propagation, 30*(1), 27-34.

[14] Breed, B., & Strauss, J. (2002). A short proof of the equivalence of LCMV and GSC beamforming. *Signal Processing Letters, IEEE, 9*(6), 168-169.

[15] Elko, G. W. (1995). A simple adaptive first-order differential microphone. *IEEE ASSP Workshop On Applications of Signal Processing to Audio and Acoustics, 1995.,* 169-172.

[16] Elko, G. W. (2000). Steerable and variable first-order differential microphone array. *U.S. Patent No. 6,041,127*. Washington, DC: U.S. Patent and Trademark Office.

[17] Chatlani, N., Fischer, E., & Soraghan, J. (2010). Spatial noise reduction in binaural hearing aids. *European Signal Processing Conference (EUSIPCO-2010)*.

[18] Brown, C. P., & Duda, R. O. (1998). A structural model for binaural sound synthesis. *IEEE Transactions on Speech and Audio Processing, 6*(5), 476-488.

[19] S. Leukimmiatis, S., & Maragos, P. (2006). Optimum post-filter estimation for noise reduction in multichannel speech processing. *14th European Signal Processing Conference, 2006*. 1-5.

[20] Rohdenburg, T., Hohmann, V., & Kollmeier, B. (2007). Robustness analysis of binaural hearing aid beamformer algorithms by means of objective perceptual quality measures. *IEEE Workshop* on *Applications of Signal Processing to Audio and Acoustics 2007.* 315-318.

[21] Souden, M., Benesty, J., & Affes, S. (2010). A study of the LCMV and MVDR noise reduction filters. *IEEE Transactions on Signal Processing, 58*(9), 4925-4935.

[22] Aarabi, P., & Shi, G. (2002). Multi-channel time-frequency data fusion. *Proceedings of the Fifth International Conference Information Fusion, 2002.* 404-411.

[23] Brandstein, M., & Ward, D. (2001). *Microphone arrays: signal processing techniques and applications*. Springer Science & Business Media.

[24] Cox, H. (1973). Resolving power and sensitivity to mismatch of optimum array processors. *The Journal of the Acoustical Society of America, 54*(3), 771-785.

[25] Teutsch, H., & Elko, G. W. (2001). First-and second-order adaptive differential microphone arrays. *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)* (pp. 35-38).

[26] Khayeri, P., Abutalebi, H. R., & Abootalebi, V. (2011). A nested superdirective generalized sidelobe canceller for speech enhancement. *8th International Conference* on *Information, Communications and Signal Processing (ICICS) 2011*. 1-5.

[27] Khayeri, P., Abutalebi, H. R., & Abootalebi, V. (2012). A hybrid near-field superdirective GSC and post-filter for speech enhancement. *Proceedings of the 20th European Signal Processing Conference (EUSIPCO), 2012*. 2065-2069.