

Design of Minimax Robust Broadband Beamformers with Optimized Microphone Positions

R. C. Nongpiur

Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada V8W 3P6

Abstract

A new method for the design of robust minimax far-field broadband beamformers with optimized microphone positions is proposed. The method is formulated as an iterative optimization problem where the maximum passband¹ magnitude response error is minimized and the microphone positions are optimized while ensuring that the minimum stopband attenuation is above a prescribed level. To maintain robustness, we constrain a sensitivity parameter, namely, the white noise gain, to be above prescribed levels across the frequency band. An additional feature of the method, which is quite useful in certain applications, is that it provides the capability of constraining the gain in the transition band to always lie below the maximum gain in the passband. Performance comparisons with existing methods show that the optimization of the microphone positions results in beamformers with superior performance.

Keywords: optimized sensor positions, acoustic beamforming, broadband beamformer, constrained optimization

1. Introduction

Microphone arrays are widely used in speech communication applications such as hands-free telephony, hearing aids, speech recognition, and teleconferencing systems. A technique that is widely used with microphone arrays to enhance a speech signal from a preferred spatial direction is beamforming [1]. In general, the beamforming approach can be fixed or adaptive, depending upon whether the spatial directivity pattern is fixed or varies adaptively on the basis of incoming data. Though adaptive beamforming performs better when the acoustic environment is time-varying, fixed beamforming is preferred in applications where the direction of the sound source is fixed, such as in in-car communication systems [2] or in hearing aids. In addition, fixed beamformers also have lower computational complexity and are easier to implement.

In many beamformer applications, such as in-car communication systems, voice recognition systems, video conferencing systems, etc., there is often a need to ensure that the gain across the passband has little variation from unity while that in the stopband is below a prescribed level. Consequently, for the design of such beamformers a straightforward approach is to formulate the problem as a minimax optimization problem [3]. In applications where high quality speech or audio is desired, a passband with good linear-phase characteristics is usually preferred to minimize signal distortion.

In broadband beamformer design, better frequency invariance can be achieved by arranging the microphone elements non-uniformly in an optimal manner. This is because the more widely separated sensors facilitate better performance at lower frequencies, while the closely spaced ones prevent spatial aliasing at higher frequencies. In [4]-[8], the microphone elements are arranged in the form of a nested array by appropriately combining several uniformly-spaced sub-arrays. Alternatively, in [9]-[11], the microphone positions are obtained by approximating a continuously distributed sensor as a discrete set of filtered broadband omnidirectional array elements. In both the methods, the signal from each microphone is appropriately filtered to adjust the time delay and to prevent spatial aliasing, which

¹In this paper, unless explicitly stated, the terms *passband*, *stopband*, and *transition band* refer to the *angular passband*, *angular stopband*, and *angular transition band* of the beamformer, respectively.

Email addresses: rnongpiu@ece.uvic.ca (R. C. Nongpiur), Fax: (+1) 250-721-6052 (R. C. Nongpiur)

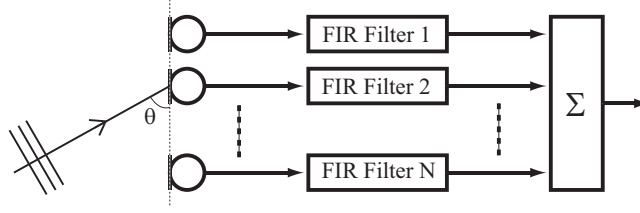


Figure 1: Filter and sum broadband beamformer.

happens when the wavelength of the signal exceeds twice the distance between the adjacent microphones. However, the microphone positions computed by using the above methods have two serious drawbacks. The first is that the positions are computed on the assumption that the array filters can have any length thereby making their performance sub-optimal when the prescribed filter length is not sufficiently long. The second drawback is the assumption that the array can have any length; such an assumption may not be applicable in certain applications such as in hearing aids and in-car communication systems where there are physical constraints on the array aperture size. Furthermore, as evident from earlier designs for superdirective narrowband arrays [12]–[16], broadband beamformers designed for physically-compact applications can likewise become very sensitive to errors in array imperfections and therefore robustness constraints need to be incorporated in the design [3, 17]–[23]. In [17]–[21], the statistics of microphone characteristics are taken into account to derive broadband beamformers that are robust to microphone mismatches, while in [3, 22, 23] the white noise gain (WNG) is incorporated in the design to ensure that the beamformer is robust to spatial white noise and array imperfections. The use of the WNG constraint is not new and has been used in earlier beamformer designs to ensure robustness in superdirective beamformers [12]–[14]. More recently, a least-squares approach [24] that uses mixed stochastic and analytic optimization to synthesize both the sensor arrays and filter coefficients for robust wideband beamformers has been proposed. In the approach, a trade-off parameter is provided so that the user can tune the mainlobe width and sidelobe energy of the beamformer.

In this paper, we propose a design method where the maximum passband magnitude response error is minimized and the microphone positions are optimized while ensuring that the minimum stopband attenuation is above a prescribed level. Although the transition region is usually treated as a “don’t care” region, in many practical applications excessive gain in this region could be undesirable. Our method provides the capability of controlling the gain in transition bands so that it does not exceed the maximum gain in the passband. To maintain robustness, we constrain a sensitivity parameter, namely, the white noise gain, to be above prescribed levels across the frequency band. The method is formulated as an iterative second-order cone programming problem (SOCP) as was done for the design of nearly linear-phase beamformers [3] and IIR filters [25, 26]. Numerical results on various types of beamformers show that the proposed method, with optimized microphone positions, results in beamformers with much lower maximum passband ripple for the same stopband attenuation when compared with beamformers where the microphone positions are fixed and not optimized.

The paper is organized as follows. In Section 2, we describe the filter-and-sum beamformer and the associated error formulations of the beamformer response and WNG for a linear array in far-field. Then in Section 3, we develop formulations for solving the optimization problem. In Section 4, performance comparisons between the proposed method and two existing methods for computing the microphone positions are carried out. Conclusions are drawn in Section 5.

2. Far-field Broadband Beamforming

In this paper, we assume a far-field signal impinging on a linear microphone array that is realized as a filter-and-sum beamformer, as shown in Fig. 1. The microphones are assumed to be omnidirectional and the filters are FIR. If N is the number of microphones and L is the length of each filter, the response of the filter-and-sum beamformer is given by [3]

$$B(\mathbf{x}, \mathbf{d}, \omega, \theta) = \sum_{n=0}^{N-1} \hat{\mathbf{g}}(d_n, \omega, \theta)^T \mathbf{x}_n = \mathbf{g}(\mathbf{d}, \omega, \theta)^T \mathbf{x} \quad (1)$$

where

$$\mathbf{x}^T = [\mathbf{x}_0^T \mathbf{x}_1^T \cdots \mathbf{x}_{N-1}^T] \quad (2)$$

$$\mathbf{d}^T = [d_0 d_1 \cdots d_{N-1}] \quad (3)$$

$$\mathbf{g}(\mathbf{d}, \omega, \theta)^T = [\hat{\mathbf{g}}(d_0, \omega, \theta)^T \hat{\mathbf{g}}(d_1, \omega, \theta)^T \cdots \hat{\mathbf{g}}(d_{N-1}, \omega, \theta)^T] \quad (4)$$

$$\mathbf{x}_n = [x_{n,0} x_{n,1} \cdots x_{n,L-1}]^T \quad (5)$$

$$\hat{\mathbf{g}}(d, \omega, \theta) = [g_0(d, \omega, \theta) g_1(d, \omega, \theta) \cdots g_{L-1}(d, \omega, \theta)]^T \quad (6)$$

$$g_l(d, \omega, \theta) = \exp \left[-j\omega \left(\frac{f_s d \cos \theta}{c} + l \right) \right] \quad (7)$$

and ω is the frequency in radians, θ is the direction of arrival, c is the speed of sound in air, f_s is the sampling frequency, d_n is the distance of the n th microphone from the origin, and $x_{n,l}$ is the l th coefficient of the n th FIR filter. If θ_d is the desired steering angle of the beamformer, the WNG of the beamformer is given by [3]

$$G_w(\mathbf{x}, \mathbf{d}, \omega) = \frac{|\mathbf{g}(\mathbf{d}, \omega, \theta_d)^T \mathbf{x}|^2}{\|\mathbf{A}(\omega) \mathbf{x}\|_2^2} \quad (8)$$

where

$$\mathbf{A}(\omega) = \mathbf{I}_N \otimes \mathbf{a}(\omega)^T \quad (9)$$

$$\mathbf{a}(\omega)^T = [1 e^{-j\omega} \cdots e^{-j(L-1)\omega}]^T$$

and \mathbf{I}_N is an $N \times N$ identity matrix, \otimes is the Kronecker product, and $\|\mathbf{v}\|_2$ is the L_2 norm of vector \mathbf{v} .

2.1. Passband Error

If $B_d(\omega, \theta)$ is the desired beampattern at a certain frequency and direction, the squared magnitude error between the beamformer response and the desired beampattern is given by

$$e_b(\mathbf{z}, \omega, \theta) = |B(\mathbf{z}, \mathbf{d}, \omega, \theta)|^2 - |B_d(\omega, \theta)|^2 \quad (10)$$

where

$$\mathbf{z}^T = [\mathbf{x}^T \mathbf{d}^T] \quad (11)$$

If \mathbf{z}_k is the value of \mathbf{z} at the start of the k th iteration and δ is the update to \mathbf{z}_k , the updated value of the squared magnitude error can be estimated by a linear approximation

$$e_b(\mathbf{z}_k + \delta, \omega, \theta) \approx e_b(\mathbf{z}_k, \omega, \theta) + \nabla e_b(\mathbf{z}_k, \omega, \theta)^T \delta \quad (12)$$

which becomes more accurate as $\|\delta\|_2$ gets smaller.

The L_p -norm of the passband squared magnitude error for the k th iteration is given by [3]

$$\begin{aligned} \mathcal{E}_p^{(pb)}(\mathbf{z}_k) &= \left[\int_{\Omega} \int_{\Theta_{pb}} |e_b(\mathbf{z}_{k+1}, \omega, \theta)|^p d\theta d\omega \right]^{1/p} \\ &\approx \kappa_{pb} \left[\sum_{r=1}^R \sum_{s=1}^{S_{pb}} |e_b(\mathbf{z}_{k+1}, \omega_r, \theta_s)|^p \right]^{1/p} \\ &\approx \left[\sum_{r=1}^R \sum_{s=1}^{S_{pb}} |\kappa_{pb} e_b(\mathbf{z}_k, \omega_r, \theta_s) + \right. \\ &\quad \left. \kappa_{pb} \nabla e_b(\mathbf{z}_k, \omega_r, \theta_s)^T \delta|^p \right]^{1/p}, \quad \omega_r \in \Omega, \theta_s \in \Theta_{pb} \end{aligned} \quad (13)$$

where Ω is the frequency band of interest, $\Theta_{pb} = [\theta_{pl}, \theta_{ph}]$ is the angular passband, and κ_{pb} is a constant. Expressing (13) in matrix form we get

$$\mathcal{E}_p^{(pb)}(\mathbf{z}_k) \approx \|\mathbf{C}_k \boldsymbol{\delta} + \mathbf{u}_k\|_p \quad (14)$$

where

$$\mathbf{C}_k = \begin{bmatrix} \kappa_{pb} \nabla e_b(\mathbf{z}_k, \omega_1, \theta_1)^T \\ \vdots \\ \kappa_{pb} \nabla e_b(\mathbf{z}_k, \omega_R, \theta_{S_{pb}})^T \end{bmatrix} \quad (15)$$

$$\mathbf{u}_k = [u_{11} \ u_{12} \ \cdots \ u_{RS_{pb}}]^T, \quad (16)$$

$$u_{rs} = \kappa_{pb} e_b(\mathbf{z}_k, \omega_r, \theta_s), \quad \omega_r \in \Omega, \theta_s \in \Theta_{pb} \quad (17)$$

where κ_{pb} is a constant and $\boldsymbol{\delta}$ is the update to \mathbf{z}_k that is constrained to be small so that the linear approximation in (14) is accurate. The right-hand side of (14) is the L_p -norm of an affine function of $\boldsymbol{\delta}$ and, therefore, it is convex with respect to $\boldsymbol{\delta}$ [27].

2.2. Stopband Error

Setting $B_d(\omega, \theta)$ to zero in (10), the resulting beamformer response error can be defined as

$$e_{b0}(\mathbf{z}, \omega, \theta) = B(\mathbf{x}, \mathbf{d}, \omega, \theta) \quad (18)$$

and as in the previous subsection, the corresponding L_p -norm of the error in the stopband for the k th iteration can be expressed as

$$\mathcal{E}_p^{(sb)}(\mathbf{z}_k) \approx \|\mathbf{D}_k \boldsymbol{\delta} + \mathbf{v}_k^{(sb)}\|_p \quad (19)$$

where

$$\mathbf{D}_k^{(sb)} = \begin{bmatrix} \kappa_{sb} \nabla e_{b0}(\mathbf{z}_k, \omega_1, \theta_1)^T \\ \vdots \\ \kappa_{sb} \nabla e_{b0}(\mathbf{z}_k, \omega_R, \theta_{S_{sb}})^T \end{bmatrix} \quad (20)$$

$$\mathbf{v}_k^{(sb)} = [v_{11}^{(sb)} \ v_{12}^{(sb)} \ \cdots \ v_{RS_{sb}}^{(sb)}]^T, \quad (21)$$

$$v_{rs}^{(sb)} = \kappa_{sb} e_{b0}(\mathbf{z}_k, \omega_r, \theta_s), \quad \omega_r \in \Omega, \theta_s \in \Theta_{sb} \quad (22)$$

where κ_{sb} is a constant and Θ_{sb} is the angular stopband. Note that if the sensor positions are not optimized, \mathbf{d} is a constant and, as a consequence, the L_p -norm of (18) in the stopband can be expressed as a convex function given by [3]

$$E_p^{(sb)}(\mathbf{x}) = \|\mathbf{U}_{sb} \mathbf{x}\|_p \quad (23)$$

where

$$\mathbf{U}_{sb} = [\kappa_{sb} \mathbf{g}(\mathbf{d}, \omega_1, \theta_1) \cdots \kappa_{sb} \mathbf{g}(\mathbf{d}, \omega_1, \theta_{S_{sb}}) \cdots \kappa_{sb} \mathbf{g}(\mathbf{d}, \omega_R, \theta_1) \cdots \kappa_{sb} \mathbf{g}(\mathbf{d}, \omega_R, \theta_{S_{sb}})]^T, \quad \theta_s \in \Theta_{sb} \quad (24)$$

2.3. Transition band Error

In some applications it is desirable that the gain in the transition band does not exceed the maximum gain in the passband. To this end, we compute the L_p norm of (18) in the transition band, given by

$$\mathcal{E}_p^{(tb)}(\mathbf{z}_k) \approx \|\mathbf{D}_k^{(tb)} \boldsymbol{\delta} + \mathbf{v}_k^{(tb)}\|_p \quad (25)$$

where

$$\mathbf{D}_k^{(tb)} = \begin{bmatrix} \kappa_{tb} \nabla e_{b0}(\mathbf{z}_k, \omega_1, \theta_1)^T \\ \vdots \\ \kappa_{tb} \nabla e_{b0}(\mathbf{z}_k, \omega_R, \theta_{S_{tb}})^T \end{bmatrix} \quad (26)$$

$$\mathbf{v}_k^{(tb)} = [v_{11}^{(tb)} \ v_{12}^{(tb)} \ \cdots \ v_{RS_{tb}}^{(tb)}]^T, \quad (27)$$

$$v_{mn}^{(tb)} = \kappa_{tb} e_{b0}(\mathbf{z}_k, \omega_r, \theta_s), \quad \omega_r \in \Omega, \theta_s \in \Theta_{tb} \quad (28)$$

and Θ_{tb} is the angular transition-band. Like the L_p -norm of the stopband error in (23), the L_p -norm of (18) in the transition band can be expressed as a convex function if the sensor positions are not optimized, given by

$$E_p^{(tb)}(\mathbf{x}) = \|\mathbf{U}_{tb}\mathbf{x}\|_p \quad (29)$$

where

$$\begin{aligned} \mathbf{U}_{tb} &= [\kappa_{tb} \mathbf{g}(\mathbf{d}, \omega_1, \theta_1) \cdots \kappa_{tb} \mathbf{g}(\mathbf{d}, \omega_1, \theta_{S_{tb}}) \cdots \\ &\quad \kappa_{tb} \mathbf{g}(\mathbf{d}, \omega_R, \theta_1) \cdots \kappa_{tb} \mathbf{g}(\mathbf{d}, \omega_R, \theta_{S_{tb}})]^T, \quad \theta_s \in \Theta_{tb} \end{aligned} \quad (30)$$

2.4. White Noise Gain

If $\Gamma_{wng}(\omega)$ is the prescribed lower bound of the WNG at frequency ω , the difference between the WNG of the beamformer and the prescribed lower bound is given by

$$e_w(\mathbf{z}, \omega) = G_w(\mathbf{x}, \mathbf{d}, \omega) - \Gamma_{wng}(\omega) \quad (31)$$

As in [3], the update of $e_w(\mathbf{z}, \omega)$ for the k th iteration can be approximated as

$$e_w(\mathbf{z}_k + \boldsymbol{\delta}, \omega) \approx e_w(\mathbf{z}_k, \omega) + \nabla e_w(\mathbf{z}_k, \omega)^T \boldsymbol{\delta}, \quad \omega \in \Omega \quad (32)$$

and the RHS of (32) can be expressed in matrix form across Ω as

$$\mathbf{w}(\mathbf{z}_k) = \mathbf{Q}_k \boldsymbol{\delta} + \mathbf{h}_k \quad (33)$$

where

$$\mathbf{Q}_k = \begin{bmatrix} \nabla e_w(\mathbf{z}_k, \omega_1)^T \\ \vdots \\ \nabla e_w(\mathbf{z}_k, \omega_R)^T \end{bmatrix} \quad (34)$$

$$\mathbf{h}_k = [e_w(\mathbf{z}_k, \omega_1) \cdots e_w(\mathbf{z}_k, \omega_R)]^T, \quad (35)$$

and $\omega_r \in \Omega$. The right-hand side of (33) is an affine function of $\boldsymbol{\delta}$ and, therefore, it is convex with respect to $\boldsymbol{\delta}$.

3. The Optimization Problem

The optimization problem is solved by minimizing the passband magnitude response error, $\mathcal{E}_p^{(pb)}$, while constraining the stopband error to be below a prescribed threshold, the WNG to be above prescribed levels across the frequency band, and the L_2 norm of the update to be small; in addition, we also include an optional constraint to ensure that the gain in the transition-band is always below the maximum gain in the passband. Consequently, we have the optimization problem

$$\begin{aligned} &\text{minimize} && \mathcal{E}_p^{(pb)}(\mathbf{z}) \\ &\text{subject to:} && \mathcal{E}_p^{(sb)}(\mathbf{z}) \leq \Gamma_{sb} \\ &&& \mathcal{E}_p^{(tb)}(\mathbf{z}) \leq \Gamma_{pb} \quad (\text{optional}) \\ &&& G_w(\mathbf{x}, \mathbf{d}, \omega_m) \geq \Gamma_{wng}(\omega_m) \quad \forall \omega_m \in \Omega \\ &&& \|\boldsymbol{\delta}\|_2 \text{ is small} \end{aligned} \quad (36)$$

where Γ_{sb} is the minimum stopband attenuation, $\Gamma_{wng}(\omega)$ is the minimum WNG at frequency ω , and Γ_{pb} is the maximum gain in the passband.

The value of p in the above optimization problem can be any positive integer. The most significant values for p are 2 and ∞ . In the first case, the L_2 -norm would be minimized, which would result in a least-squares solution, whereas, in the second case, the L_∞ -norm would be minimized, which would result in a minimax solution. In this paper, we explore the minimax design and therefore set $p = \infty$. Consequently, substituting $\mathcal{E}_p^{(pb)}$, $\mathcal{E}_p^{(sb)}$, $\mathcal{E}_p^{(tb)}$, and G_w by the expressions in (14), (19), (25), and (33), respectively, the optimization problem for the k th iteration can be expressed as

$$\begin{aligned} & \text{minimize} && \|\mathbf{C}_k \boldsymbol{\delta} + \mathbf{u}_k\|_\infty + W \delta_{rlx} \\ & \text{subject to:} && \mathbf{Q}_k \boldsymbol{\delta} + \mathbf{h}_k \geq \mathbf{0} - \delta_{rlx} \mathbf{1} \\ & && \|\mathbf{D}_k^{(sb)} \boldsymbol{\delta} + \mathbf{v}_k^{(sb)}\|_\infty \leq \Gamma_{sb} + \delta_{rlx} \\ & && \|\mathbf{D}_k^{(tb)} \boldsymbol{\delta} + \mathbf{v}_k^{(tb)}\|_\infty \leq \Gamma_{pb}(k) + \delta_{rlx} \text{ (optional)} \\ & && \|\boldsymbol{\delta}\|_2 \leq \Gamma_\delta(k) + \delta_{rlx} \\ & && \delta_{rlx} \geq 0 \end{aligned} \tag{37}$$

where $\boldsymbol{\delta} \in \mathbf{R}^{LN+N}$ and $\delta_{rlx} \in \mathbf{R}^1$ are the optimization variables, $W > 0$,

$$\Gamma_{pb}(k) = \sup_{\omega \in \Omega, \theta \in \Theta_{pb}} |B(\mathbf{x}_k, \mathbf{d}_k, \omega, \theta)| \tag{38}$$

and [3]

$$\Gamma_\delta(k) = \begin{cases} \gamma_k & k < T \\ \gamma_{small} & \text{otherwise} \end{cases} \tag{39}$$

such that $\gamma_i > \gamma_{i+1}$. Note that W is a weighting factor that is made sufficiently large so that δ_{rlx} converges to 0. To speed up the convergence, $\Gamma_\delta(k)$ is made relatively large during the starting iteration and gradually reduced to a small fixed value after a certain number of iterations. As in [3], parameter δ_{rlx} is a slack variable that is introduced in the optimization to ensure that the problem does not become infeasible if the stopband or WNG constraint is not satisfied during the starting phase of the optimization iterations. The optimal value, $\boldsymbol{\delta}_{opt}$, obtained by solving the convex optimization problem in (37) is then used to update \mathbf{z}_k , i.e.,

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \boldsymbol{\delta}_{opt} \tag{40}$$

3.1. Deriving the Initializing Beamformer

To derive the initializing beamformer for the optimization problem above, we consider the optimization problem where the beamformer passband response error is minimized under the constraint that the stopband attenuation and WNG are above prescribed levels. In [3], it is shown that such a problem is nonconvex since the WNG is nonconvex, but can be made convex by approximating the magnitude of the numerator of the WNG to unity. In addition, it is observed in [3] that a regularized version of the convex formulation facilitates faster convergence. Therefore, to derive the initialization beamformer we solve the regularized convex formulation as in [3, c.f. eqn. (55)]. The microphone positions for the initializing beamformer are nonuniformly spaced and either symmetric or nonsymmetric as per the design requirement.

3.2. Perfect linear-phase and maximum array length constraints

Depending on whether or not a beamformer with perfect linear-phase across the frequency band is desired, the location of the microphones along the linear array can be either symmetric or non-symmetric. In [3], it is shown that for perfect linear-phase beamformers, the sensor array should be symmetric about the array center such that

$$d_{N-n-1} = -d_n \tag{41}$$

and the filter coefficients should satisfy the condition

$$x_{n,l} = x_{N-n-1,L-l-1} \quad (42)$$

with a resulting group-delay of $(L-1)f_s^{-1}/2$. The two conditions above are satisfied by incorporating the following constraints in the iterative optimization problem:

$$d_n^{(k)} + \delta_{d_n} = -d_{N-n-1}^{(k)} - \delta_{d_{N-n-1}} \quad (43)$$

$$x_{n,l}^{(k)} + \delta_{n,l} = x_{N-n-1,L-l-1}^{(k)} + \delta_{N-n-1,L-l-1} \quad (44)$$

where $n \in [0, N-1]$, $l \in [0, L-1]$, and δ_{d_i} and $\delta_{n,l}$ are the updates for d_i and $x_{n,l}$, respectively, during the k th iteration. Note that if the initialization filter coefficients do not satisfy the constraint in (44) the optimization problem may become infeasible. To avoid the possibility, we can frame the constraint by including the slack parameter δ_{rlx} as

$$|x_{n,l}^{(k)} + \delta_{n,l} - x_{N-n-1,L-l-1}^{(k)} - \delta_{N-n-1,L-l-1}| \leq \delta_{rlx} \quad (45)$$

which becomes equal to (44) when δ_{rlx} is minimized to 0.

If a perfectly linear-phase beamformer is not required, ignoring the constraints in (41) and (42) will result in greater degrees of freedom in the optimization thereby improving the passband and stopband performances further.

In applications where the maximum length of the array should be less than a prescribed length, d_{max} , we can incorporate a convex inequality constraint in the optimization problem in (37) to ensure that the array length of the optimized microphone positions is always below d_{max} . Such a constraint is given by

$$\sup_i \{d_i^{(k)} + \delta_{d_i}\} - (d_n^{(k)} + \delta_{d_n}) \leq d_{max} \quad \forall n \in [0, N-1] \quad (46)$$

3.3. Practical Considerations

To significantly reduce the computational complexity in the optimization, the 2-D nonuniform variable sampling technique in [3] is used. The weights W for the slack parameter, δ_{rlx} , in (37) should not be too small as this can make the optimization algorithm unstable and prevent it from converging; at the same time, it should also not be too large as this can slow down the convergence process. Typical values of W that have been found to work well range between 500 to 5000.

To ensure that the optimization is not prematurely terminated, the termination condition is decided by monitoring the values of the objective function typically for the last 5 iterations as was done in [3].

4. Numerical Results

In this section, we compare the performance of beamformers with optimized microphone positions proposed in this paper with those that use the microphone positions derived by two existing methods. The experiments are divided into two subsections. In the first subsection, the beamformer is symmetric about $\theta = \pi/2$ while in the second subsection the beamformer is nonsymmetric.

To compare our proposed method, we consider the methods in [9],[11] and [5]-[8]. In [9] and [11] the microphone positions are derived for symmetric and nonsymmetric beampatterns, respectively, by approximating a continuously distributed sensor. In [5]-[8] the microphone positions are computed by using the harmonic nesting approach.

Performance Comparison Procedure

In each subsection, we consider design examples of beamformers with perfect linear phase and no linear phase requirements. As described in Section 3.2, the perfect linear phase designs are obtained by incorporating the constraints in (43) and (45) in the optimization problem. Consequently, depending upon the specifications, we have the following design choices for the proposed variants:

Design P: The initializing microphone array is non-symmetric and is obtained by using the method in [11]. The beamformer is obtained by solving the optimization problem in (37).

Design P(LP): The beamformer has perfect linear phase, which implies that conditions (41) and (42) are satisfied. The initializing microphone array is symmetric to satisfy condition (41) and is computed using the harmonic nested array method in [5]-[8] or the method in [9]². The beamformer is obtained by solving the optimization problem in (37); to satisfy conditions (41) and (42), the linear-phase constraints in (43) and (45) are incorporated in the optimization.

For a fair comparison, the competing beamformers are designed using an algorithm that is equivalent to the proposed optimization method but where the microphone positions are fixed and not optimized; such an optimization problem is given by (59) in the Appendix. As a consequence, we have the following competing methods for comparison:

Design C: The microphone array is non-symmetric and the microphone positions are computed using the method in [11]. The beamformer is obtained by solving the optimization problem in (59) where the microphone positions are fixed and not optimized.

Design C(LP): The beamformer has perfectly linear phase. The microphone array is symmetric in order to satisfy condition (41) and the microphone positions are computed using the harmonic nested array method in [5]-[8]. The beamformer is obtained by solving the optimization problem in (59) where the microphone positions are fixed and not optimized. To satisfy condition (42) for linear phase, constraint (45) is incorporated in the optimization.

For the symmetric beampattern case, the desired steering angle θ_d , which is used in (8), is set to $\pi/2$ and for the nonsymmetric case to $2\pi/3$. In all designs the WNG is constrained to be above 0 dB, and therefore $\Gamma_{wng} = 1$.

The initializing beamformer for the optimization problem in (37) is designed using the method described in Section 3.1 by setting the desired beamformer response in the passband to be

$$B_d(\omega, \theta) = e^{-j\omega\tau_d} \quad (47)$$

where $\tau_d = (L - 1)f_s^{-1}/2$.

For the iterative optimization problem in (37), W is set to 1000, γ_{small} to 0.001, and $\Gamma_\delta(k)$ is defined as

$$\Gamma_\delta(k) = \begin{cases} \gamma_k & k < 20 \\ 0.001 & k \geq 20 \end{cases} \quad (48)$$

where

$$\gamma_k = \gamma_1 - \frac{(\gamma_1 - \gamma_{19})(k - 1)}{20 - 1} \quad (49)$$

$\gamma_1 = 0.5$ and $\gamma_{19} = 0.001$. The speed of sound, c , is assumed to be 340 m/s while the sampling frequency, f_s , is assumed to be 8 kHz. The frequency and angle dependent parameters are evaluated using the 2-D nonuniform variable sampling technique as described in [3] where the number of virtual points along the angular and frequency dimensions are 500 and 200, respectively. For each passband, stopband or transition-band, we set one actual sampling point to correspond to approximately 9 virtual sampling points along each of the angular and frequency dimensions; however, near the band edges, we set one actual sampling point to correspond to one virtual sampling point, for the last three actual sampling points from an edge, as illustrated in Fig. 2. In our experiments, we have always included the optional constraints in (37) and (59) to ensure that the transition-band gain is always below the maximum passband gain.

Performance Measures

The performance of the beamformer is evaluated using the following parameters:

Maximum passband ripple: The parameter is defined as

$$A_p = 20 \log \frac{M_{max}^{(p)}}{M_{min}^{(p)}} \quad (50)$$

where

$$M_{max}^{(p)} = \max_{\omega \in \Omega, \theta \in \Theta_{pb}} B(\omega, \theta) \quad (51)$$

²When the aperture size, Q , in the method in [9] is set to 2, the resulting array is identical to the harmonic nested array.

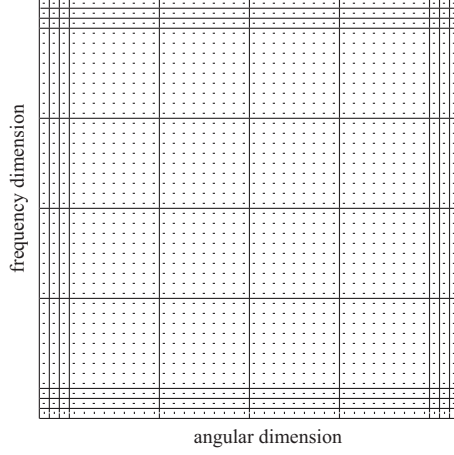


Figure 2: Diagram to illustrate nonuniform variable sampling across the angular and frequency dimensions of a desired region where each point corresponds to a virtual sampling point. From each rectangular block, a virtual point with the minimum error is selected as the actual sampling point for that block, as described in [3]. Note that close to the edges the number of virtual points per actual sampling point is reduced in order to minimize undesirable error peaks at the edges.

$$M_{min}^{(p)} = \min_{\omega \in \Omega, \theta \in \Theta_{pb}} B(\omega, \theta) \quad (52)$$

Minimum stopband attenuation: The minimum stopband attenuation is defined as the negative of the maximum stopband gain, given by

$$A_a = -20 \log M_{max}^{(a)} \quad (53)$$

where

$$M_{max}^{(a)} = \max_{\omega \in \Omega, \theta \in \Theta_{sb}} B(\omega, \theta) \quad (54)$$

Passband average group delay: The average group delay is evaluated by taking the average of the maximum and minimum group delay in the passband, given by

$$\tau_{avg} = \frac{\tau_{min} + \tau_{max}}{2} \quad (55)$$

where

$$\tau_{min} = \min_{\omega \in \Omega, \theta \in \Theta_p} \tau(\omega, \theta) \quad (56)$$

$$\tau_{max} = \max_{\omega \in \Omega, \theta \in \Theta_p} \tau(\omega, \theta) \quad (57)$$

Passband group-delay deviation: The passband group delay deviation is given by

$$\sigma_\tau = \tau_{max} - \tau_{min} \quad (58)$$

4.1. Examples 1 and 2: Symmetric Beampatterns

Here we consider the design of beamformers that are symmetric about $\theta = \pi/2$. The design specifications for both the examples are given in Table 1. The first example requires no phase linearity across the frequency band while the second example requires perfect linear phase. Consequently, for the first example we compare designs P and C, while for the second example we compare designs P(LP) and C(LP).

The comparison results for examples 1 and 2 are summarized in Tables 2 and 3 and the beamformer response, white noise gain, and sensor positions for the two examples are plotted in Figs. 3 & 4 and Figs. 5 & 6, respectively. As

Table 1: Design specifications for beamformers symmetric about $\theta = \pi/2$ for Examples 1 and 2

| Parameters | Values |
|--|--|
| No. of elements of beamformer, N | 7 |
| Sampling frequency, f_s , (Hz) | 8000 |
| FIR filter length | 20 |
| Passband region, Θ_{pb} , (deg) | $[80^\circ - 100^\circ]$ |
| Stopband region, Θ_{sb} , (deg) | $[0^\circ - 60^\circ] \cup [120^\circ - 180^\circ]$ |
| Frequency band, $[f_l - f_h]$, (Hz) | [400 - 3500] (example 1) [600 - 3500] (example 2) |
| Perfect Linear Phase | No (example 1) Yes (example 2) |
| Minimum stopband attenuation (dB) | 6.5 |
| Minimum WNG (dB) | 0 |

Table 2: Design results for example 1 of beamformers with symmetric beampatterns and no linear-phase requirement

| Parameters | P | C |
|--------------------------|--------------|-------|
| Max PB ripple A_p , dB | 0.095 | 0.696 |
| Min SB atten. A_a , dB | 6.99 | 6.97 |
| τ_{avg} , samples | 9.14 | 9.47 |
| σ_τ , samples | 4.32 | 1.96 |

PB: passband; SB: stopband; BP: beampattern

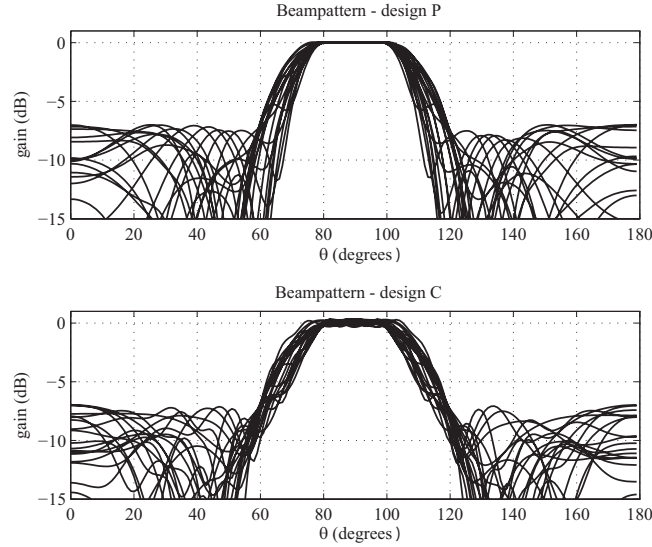


Figure 3: Beampattern plots for designs P and C for example 1. The plots are obtained by plotting the responses across 20 uniformly sampled frequency-points in the frequency band.

Table 3: Design results for example 2 of beamformers with symmetric beampatterns and perfect linear phase

| Parameters | P(LP) | C(LP) |
|--------------------------|-------------|-------|
| Max PB ripple A_p , dB | 0.72 | 2.02 |
| Min SB atten. A_a , dB | 6.99 | 6.87 |
| τ_{avg} , samples | 9.5 | 9.5 |
| σ_τ , samples | 0 | 0 |

PB: passband; SB: stopband; BP: beampattern

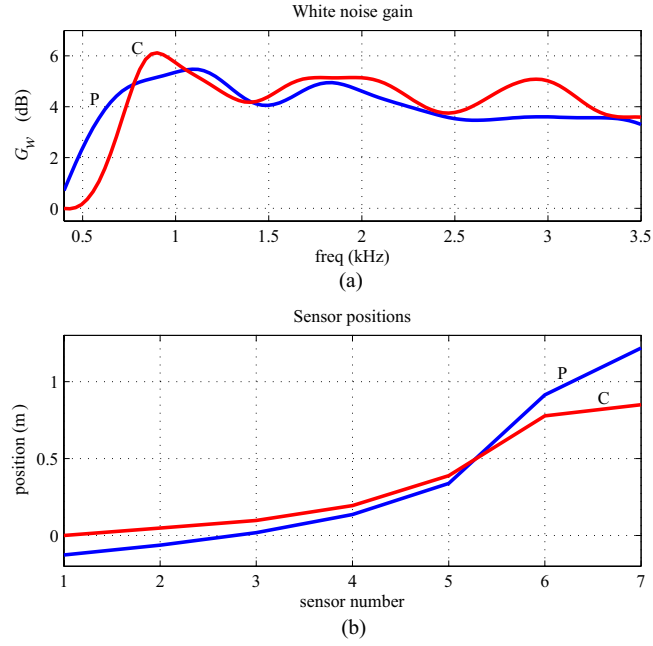


Figure 4: Plots of the white noise gain and sensor positions for designs P and C for example 1.

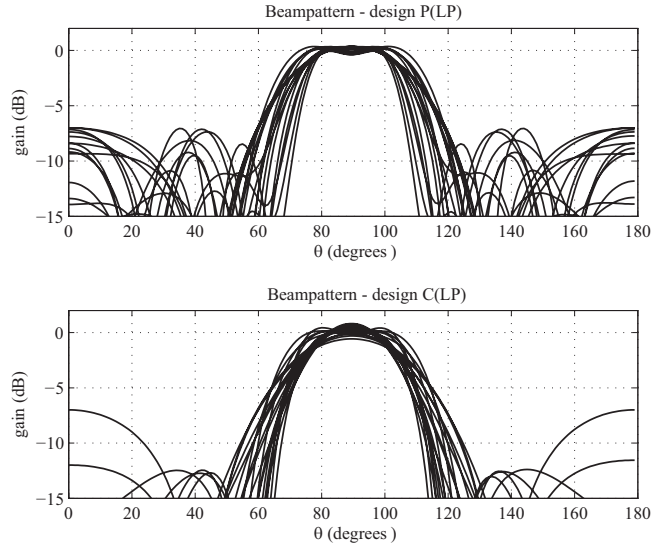


Figure 5: Beampattern plots for the design P(LP) and C(LP) for example 2. The plots are obtained by plotting the responses across 20 uniformly sampled frequency-points in the frequency band.

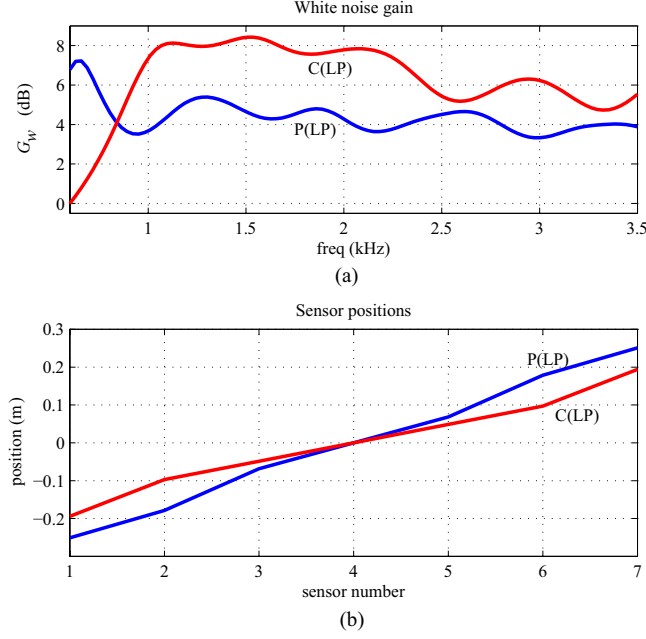


Figure 6: Plots of the white noise gain and sensor positions for design P(LP) and C(LP) for example 2.

can be seen in Tables 2 and 3 and in the beampattern plots in Figs. 3 and 5, the proposed designs in both the examples have the smallest passband ripple for nearly the same stopband attenuation values. From Figs. 4(a) and 6(a), we observe that the proposed designs have greater minimum-WNG in both the examples. In addition, the sensor-position plots in Figs. 4(b) and 6(b) reveal that the microphone positions of the proposed designs have expanded outward, resulting in a slight increase in their array lengths

Comparing design P in example 1 with design P(LP) in example 2, we observe that design P has a much smaller passband ripple despite having a wider bandwidth due to the lower cutoff frequency f_l . Therefore, this implies that the perfect linear-phase constraint severely restricts the degree of freedom and is attained at the cost of a larger passband ripple.

4.2. Examples 3 and 4: Nonsymmetric Beampatterns

In examples 3 and 4, we consider the design of beamformers with nonsymmetric beampatterns. The design specifications for the two examples are given in Table 4. Example 3 requires no phase linearity across the frequency

Table 4: Design specifications for a Beamformer with nonsymmetric beampattern for Examples 3 and 4

| Parameters | Values |
|--|---|
| No. of elements of beamformer, N | 7 |
| Sampling frequency, f_s , (Hz) | 8000 |
| FIR filter length | 20 |
| Passband region, Θ_{pb} , (deg) | $[80^\circ - 100^\circ]$ |
| Stopband region, Θ_{sb} , (deg) | $[0^\circ - 60^\circ] \cup [120^\circ - 180^\circ]$ |
| Frequency band, $[f_l - f_h]$, (Hz) | $[400 - 3500]$ (example 3) |
| | $[600 - 3500]$ (example 4) |
| Perfect Linear Phase | No (example 3) |
| | Yes (example 4) |
| Minimum stopband attenuation (dB) | 5.5 |
| Minimum WNG (dB) | 0 |

Table 5: Design results for example 3 of beamformers with nonsymmetric beampatterns and no linear-phase requirement

| Parameters | P | C |
|--------------------------|-------------|------|
| Max PB ripple A_p , dB | 0.88 | 1.04 |
| Min SB atten. A_a , dB | 6.97 | 6.9 |
| τ_{avg} , samples | 6.92 | 8.07 |
| σ_τ , samples | 10.49 | 8.71 |

PB: passband; SB: stopband; BP: beampattern

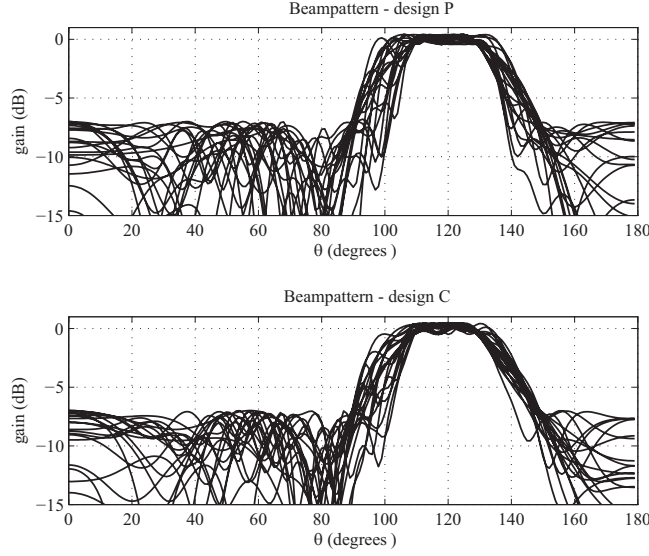


Figure 7: Beampattern plots for designs P and C for example 3. The plots are obtained by plotting the responses across 20 uniformly sampled frequency-points in the frequency band.

band and therefore we compare designs P and C. On the other hand, example 4 requires perfect linear phase and therefore designs P(LP) and C(LP) are compared.

The comparison results for examples 3 and 4 are summarized in Tables 5 and 6 and the beamformer response, white noise gain, and sensor positions for the two examples are plotted in Figs. 7 & 8 and Figs. 9 & 10, respectively. From Tables 5 and 6, it is apparent that the proposed designs in both examples 3 and 4 have smaller passband ripple than the competing designs, for similar stopband attenuation. From Figs. 8(a) and 10(a), we observe that the minimum-WNG of the proposed and competing designs in Examples 3 and 4 are at their prescribed lower limit of 0 dB. It is also interesting to observe that in example 3 the passband ripple improvement over the competing design design is not as high as in example 4. This implies that the microphone positions computed by the competing method in example 3 is close to the optimal solution; this can be confirmed from Fig. 8(b) where we observe that the sensor positions of the proposed and competing designs are almost identical, except for the 6th sensor. For example 4, we observe from

Table 6: Design results for example 4 of beamformers with nonsymmetric beampatterns and perfect linear phase

| Parameters | P(LP) | C(LP) |
|--------------------------|-------------|-------|
| Max PB ripple A_p , dB | 0.83 | 2.6 |
| Min SB atten. A_a , dB | 6.98 | 5.85 |
| τ_{avg} , samples | 9.5 | 9.5 |
| σ_τ , samples | 0 | 0 |

PB: passband; SB: stopband; BP: beampattern

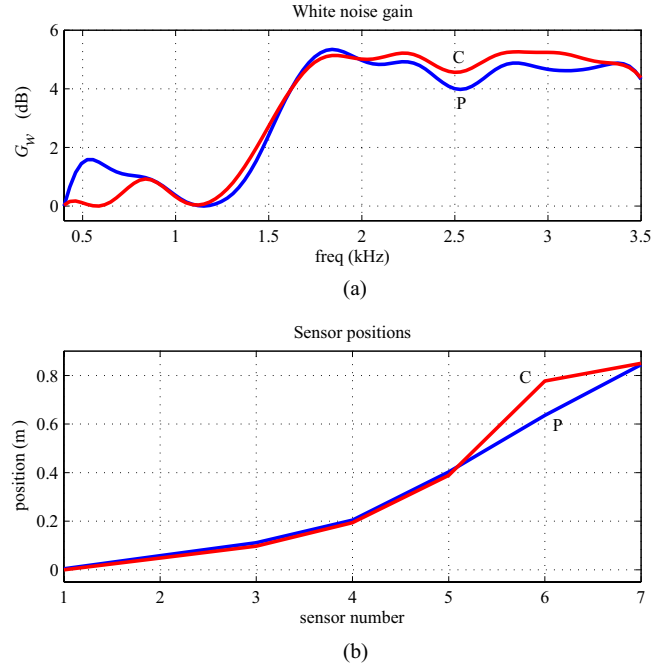


Figure 8: Plots of the white noise gain and sensor positions for designs P and C for example 3.

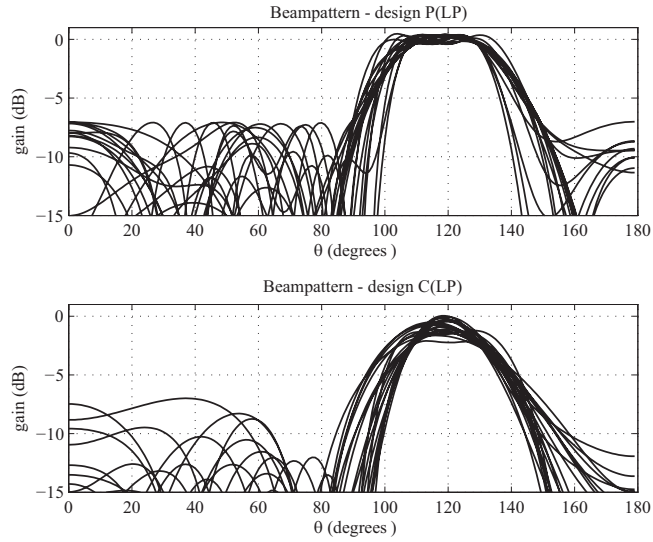


Figure 9: Beampattern plots for designs P(LP) and C(LP) for example 4. The plots are obtained by plotting the responses across 20 uniformly sampled frequency-points in the frequency band.

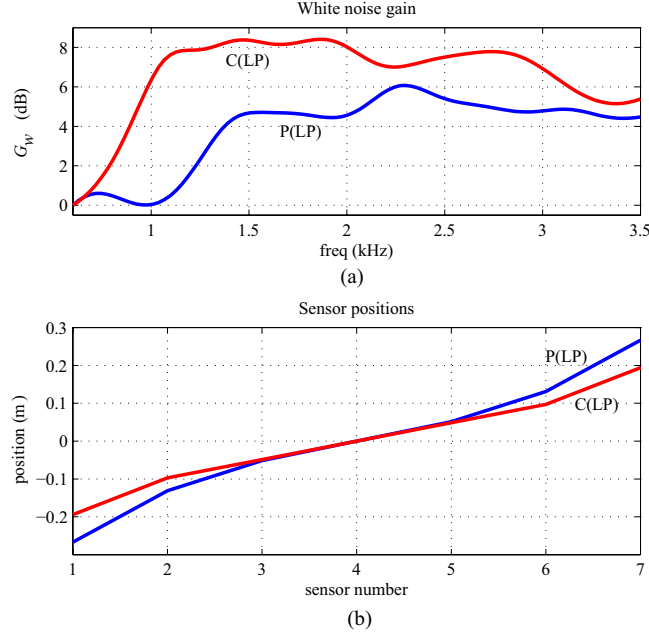


Figure 10: Plots of the white noise gain and sensor positions for designs P(LP) and C(LP) for example 4.

Fig. 10(b) that the sensor positions for the proposed design are more spread out thereby resulting in a slightly longer array length than the competing design.

The above design examples have shown that optimizing the microphone positions always result in beamformers with better performance. Note that in all the design examples in this paper, we have imposed the condition where the gain in the transition band cannot exceed the maximum gain in the passband; if this condition is relaxed, it may be possible to achieve smaller passband ripple in some of the examples. It should also be pointed out that the optimization of the microphone positions, passband-magnitude error, and WNG is a highly nonlinear and nonconvex problem and there is no guarantee that the optimized microphone positions and beamformer solution are globally optimal. In effect, the proposed method gives quality suboptimal designs that may sometime be globally optimal.

Due to the non-convex nature of the optimization problem, a good initialization point is important for fast convergence and for obtaining a good solution. For example, initializing with non-uniform sensor positions derived from the competing methods would, in general, yield better solutions than initializing with uniformly spaced sensors. This is because a good solution usually has sensor positions that bear greater resemblance to a non-uniformly spaced array and often much longer than a uniformly spaced array, where the adjacent sensor distance is constrained to be less than half the minimum wavelength to avoid spatial aliasing. Consequently, when initializing with uniformly spaced arrays, the optimization algorithm will take more iterations to converge to a good solution; furthermore, due to the highly non-linear and non-convex nature of the problem, there is also a much higher possibility that the algorithm will converge to a poorer suboptimal solution.

The optimization problems in the examples were solved on a computer running an Intel Core i7-640LM processor using the SeDuMi optimization toolbox for MATLAB [28]. There were no significant differences in the running time between the optimization algorithms with optimized and fixed sensor positions in (37) and (59), respectively. In both algorithms, the running time for each iteration takes less than a minute to compute and the optimization usually converges to a good solution in less than 25 minutes.

5. Conclusions

A new method for the design of minimax robust far-field broadband beamformers with optimized microphone positions has been described. The method is formulated as an iterative optimization problem where the maximum

passband magnitude error is minimized and the microphone positions are optimized while ensuring that the minimum stopband attenuation is above a prescribed level. To maintain robustness, we constrain a sensitivity parameter, namely, the white noise gain, to be above prescribed levels across the frequency band. An additional feature of the method, which is very useful in certain applications, is the inherent capability of constraining the gain in transition bands to be below the maximum gain in the passband. This facilitates the elimination of transition-band anomalies which sometimes occur in beamformers designed by optimization. Numerical results have shown that beamformers designed by optimizing the microphone positions have much lower maximum passband ripple for the same stopband attenuation when compared with those where the microphone positions are fixed.

6. Appendix

6.1. Simplified optimization problem when the microphones positions are fixed

When the positions of the microphones are not required to be optimized we simply constrain the updates of the microphone positions to zero in (37). Additionally, the stopband and transition-band constraints also become convex and can be replaced by the convex formulations in (23) and (29). Consequently, the simplified optimization problem becomes

$$\begin{aligned}
& \text{minimize} && \| \mathbf{C}_k \boldsymbol{\delta} + \mathbf{d}_k \|_\infty + W \delta_{rlx} \\
& \text{subject to:} && \mathbf{Q}_k \boldsymbol{\delta} + \mathbf{g}_k \geq \mathbf{0} - \delta_{rlx} \\
& && \| \mathbf{U}_{sb}(\mathbf{x}_k + \boldsymbol{\delta}_x) \|_\infty \leq \Gamma_{sb} + \delta_{rlx} \\
& && \| \mathbf{U}_{tb}(\mathbf{x}_k + \boldsymbol{\delta}_x) \|_\infty \leq \Gamma_{pb}(k) + \delta_{rlx} \text{ (optional)} \\
& && \| \boldsymbol{\delta} \|_2 \leq \Gamma_\delta(k) + \delta_{rlx} \\
& && \delta_{d_n} = 0 \quad \forall n \in [0, N-1] \\
& && \delta_{rlx} \geq 0
\end{aligned} \tag{59}$$

where $\boldsymbol{\delta}_x \in \mathbf{R}^{LN}$ corresponds to the update for \mathbf{x} during the k th iteration and is part of variable $\boldsymbol{\delta}$ such that

$$\boldsymbol{\delta}^T = [\boldsymbol{\delta}_x^T \quad \delta_{d_0} \dots \delta_{d_{N-1}}] \tag{60}$$

Acknowledgment

I thank the anonymous reviewers for their contributions to improving the quality of this paper.

References

- [1] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*, Springer-Verlag, May 2001.
- [2] E. Hansler and G. Schmidt, *Acoustic echo and noise control - A practical approach*, Wiley-Interscience 2004.
- [3] R. C. Nongpiur and D. J. Shpak, L -infinity norm design of linear-phase robust broadband beamformers using constrained optimization, *IEEE Trans. Signal Process.*, 61 (23) (2013) 6034-6046.
- [4] R. Smith, Constant beamwidth receiving arrays for broad band sonar systems, *Acustica*, 23 (1970) 21-26.
- [5] M. Goodwin, Frequency-independent beamforming, *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, 1993.
- [6] F. Khalil, J.P. Jullien, A. Gilioire, Microphone array for sound pickup in teleconference systems, *J. Audio Eng. Soc.*, 42 (1994) 691-700.
- [7] Y. Mahieux, G. Le Tourneur, and A. Saliou, A microphone array for multimedia workstations, *J. Audio Eng. Soc.*, 44 (1996) 365-372.
- [8] C. Marro, Y. Mahieux, and K. U. Simmer, Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering, *IEEE Trans. Speech Audio Processing*, 6 (1998) 240-259.
- [9] D. B. Ward, R. A. Kennedy, and R. C. Williamson, Constant directivity beamforming, chapter 1 in *Microphone Arrays: Signal Processing Techniques and Applications* (Brandstein, M. S. and Ward, D. B., Eds.), pp. 1-15, Springer-Verlag, May 2001.
- [10] D. B. Ward, R. A. Kennedy, and R. C. Williamson, FIR filter design for frequency-invariant beamformers, *IEEE Signal Process. Lett.*, 3 (1996) 69-71.
- [11] D. B. Ward, R. A. Kennedy, and R. C. Williamson, Theory and design of broadband sensor arrays with frequency-invariant far-field beam pattern, *J. Acoust. Soc. Amer.*, 97 (1995) 1023-1034.
- [12] H. Cox, R. Zeskind, and T. Kooij, Practical supergain, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-34 (3) (1996) 393-398.
- [13] J. M. Kates, Superdirective arrays for hearing aids, *J. Acoust. Soc. Amer.*, 94 (1993) 1930-1933.

- [14] J. Bitzer and K. U. Simmer, Superdirective microphone arrays, in *Microphone Arrays: Signal Processing Techniques and Applications*, M. S. Brandstein and D. B. Ward, Eds. New York: Springer-Verlag, ch. 2, pp. 19-38, May 2001.
- [15] D. J. Shpak and A. Antoniou, A flexible optimization method for the pattern synthesis of equispaced linear arrays with equiphase excitation, *IEEE Trans. Antennas Propagat.*, 40 (1992) 1113-1120.
- [16] D. J. Shpak, A method for the optimal pattern synthesis of linear arrays with prescribed nulls, *IEEE Trans. Antennas Propagat.*, 44 (1996) 286-294.
- [17] S. Doclo and M. Moonen, Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics, *IEEE Trans. Signal Process.*, 51 (10) (2003) 2511-2526.
- [18] H. Chen, W. Ser, and Z. L. Yu, Optimal design of nearfield wideband beamformers robust against errors in microphone array characteristics, *IEEE Trans. Circuit Syst.*, 54 (2007) 1950-1959.
- [19] H. Chen and W. Ser, Design of robust broadband beamformers with passband shaping characteristics using Tikhonov regularization, *IEEE Trans. Audio, Speech, Lang. Process.*, 17 (4) (2009) 665-681.
- [20] M. Crocco and A. Trucco, A computationally efficient procedure for the design of robust broadband beamformers, *IEEE Trans. Signal Process.*, 58 (10) (2010) 5420-5424.
- [21] M. Crocco and A. Trucco, Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance, *IEEE Trans. Signal Process.*, 59 (5) (2011) 2169-2181.
- [22] E. Mabande, A. Schad, and W. Kellermann, Design of robust superdirective beamformer as a convex optimization problem, *Int. Conf. Acoust., Speech, Signal Process.*, Taipei, Taiwan, Apr. 2009.
- [23] E. Mabande, A. Schad, and W. Kellermann, A time-domain implementation of data-independent robust broadband beamformers with low filter order, 2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays, Edinburgh, May 2011.
- [24] M. Crocco, A. Trucco, Stochastic and analytic optimization of sparse aperiodic arrays and broadband beamformers with robust superdirective patterns, *IEEE Trans. Audio, Speech, Lang. Process.*, 20 (9) (2012) 2433-2447.
- [25] W.-S. Lu and T. Hinamoto, Optimal design of IIR digital filters with robust stability using conic-quadratic-programming updates, *IEEE Trans. Signal Process.*, 51 (6) (2003) 1581-1592.
- [26] R. C. Nongpiur, D. J. Shpak and A. Antoniou, Improved design method for nearly linear-phase IIR filters using constrained optimization, *IEEE Trans. Signal Process.*, 61 (4) (2013) 895-906.
- [27] A. Antoniou, W.-S. Lu, *Practical Optimization: Algorithms And Engineering Applications*, Springer 2007.
- [28] J. F. Sturm, Using SeDuMi1.02, a MATLAB toolbox for optimization over symmetric cones, *Optim. Methods Softw.*, 11-12, pp. 625-653, 1999.