DESIGN OF BROADBAND SPEECH BEAMFORMERS ROBUST AGAINST ERRORS IN THE MICROPHONE ARRAY CHARACTERISTICS

Simon Doclo, Marc Moonen

Katholieke Universiteit Leuven, Dept. of Elec. Engineering (ESAT - SISTA)
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
{simon.doclo,marc.moonen}@esat.kuleuven.ac.be

ABSTRACT

Fixed broadband beamformers for speech applications using small-sized microphone arrays are known to be highly sensitive to errors in the microphone array characteristics. This paper describes two procedures for designing broadband beamformers with an arbitrary spatial directivity pattern, which are robust against gain and phase errors. The first design procedure optimises the mean performance of the broadband beamformer and requires knowledge of the gain and phase probability density functions, whereas the second design procedure optimises the worst-case performance by using a minimax criterion. Simulations with a small-sized microphone array show the performance improvement that can be obtained by using a robust broadband beamformer design procedure.

1. INTRODUCTION

In many speech communication applications, such as hands-free telephony, hearing aids and voice-controlled systems, the microphone signals are corrupted by background noise and reverberation. Fixed and adaptive beamforming are well-known multi-microphone signal enhancement techniques for noise reduction and dereverberation [1]. Fixed beamformers are frequently used for creating the speech reference signal in a Generalised Sidelobe Canceller, for creating multiple beams [2], in highly reverberant acoustic environments and in applications where the position of the speech source is approximately known, e.g. hearing aids [3].

In [4][5] several procedures are presented for designing broadband beamformers with an arbitrary spatial directivity pattern using an FIR filter-and-sum structure. Several cost functions can be used, leading to e.g. weighted least-squares (LS) filter design, non-linear optimisation [6], a maximum energy array or eigenfilters. Whatever design procedure is used, fixed beamformers are known to be highly sensitive to errors in the microphone array characteristics (gain, phase, position), especially when using small-sized microphone arrays. The robustness against random errors can be improved by limiting the white noise gain [7] or by performing a calibration procedure with the used microphone array [8].

This paper discusses the design of broadband beamformers that are robust against unknown gain and phase errors in the microphone characteristics. In Section 2 the far-field broadband beamforming problem is introduced. Section 3 discusses the weighted LS cost function, which can be used for broadband beamformer design when the microphone characteristics are exactly known. Section 4 describes two procedures for designing robust broadband beamformers, without a need for calibration or measurement. The first procedure optimises the mean performance, whereas the second procedure optimises the worst-case performance. In Section 5 simulation results are presented and it is shown that robust broadband beamformer design gives rise to a significant performance improvement when gain and phase errors occur.

This work was carried out in the frame of the F.W.O. Project G.0233.01, *Signal Processing and Automatic Patient Fitting for Advanced Auditory Prostheses*, the Concerted Research Action *Mathematical Engineering Techniques for Information and Communication Systems* (MEFISTO-666) of the Flemish Government, the I.W.T.-project *Multi-microphone Signal Enhancement Techniques for handsfree telephony and voice controlled systems* (MUSETTE-II), IUAP P5-22, *Dynamical systems and control: computation, identification and modelling*, and P5-11, *Mobile multimedia communication systems and networks* and was partially sponsored by Cochlear.

2. BROADBAND BEAMFORMING: CONFIGURATION

Consider the linear microphone array depicted in Fig. 1, with N microphones, N L-taps FIR filters $\mathbf{w}_n$ (with real coefficients) and d_n the distance between the nth microphone and the centre of the array. The characteristics of the nth microphone are described by

$$A_n(\omega,\theta) = a_n(\omega,\theta)e^{-j\gamma_n(\omega,\theta)}, \quad n = 0 \dots N-1 , \tag{1}$$

where both the gain $a_n(\omega,\theta)$ and the phase $\gamma_n(\omega,\theta)$ can be frequency and angle-dependent. Assuming far-field conditions, the spatial directivity pattern $H(\omega,\theta)$ for a source $S(\omega)$ with frequency ω at an angle θ from the microphone array is defined as

$$H(\omega,\theta) = \mathbf{w}^T\bar{\mathbf{g}}(\omega,\theta) , \tag{2}$$

with $\mathbf{w}$ the real-valued M-dimensional vector ($M = LN$) of filter coefficients, $\mathbf{w} = \begin{bmatrix}\mathbf{w}_0^T & \dots & \mathbf{w}_{N-1}^T\end{bmatrix}^T$, and the steering vector $\bar{\mathbf{g}}(\omega,\theta)$ equal to

$$\bar{\mathbf{g}}(\omega,\theta) = \mathbf{A}(\omega,\theta)\cdot\mathbf{g}(\omega,\theta) . \tag{3}$$

$\mathbf{A}(\omega,\theta)$ is an $M \times M$ diagonal matrix consisting of the microphone characteristics and $\mathbf{g}(\omega,\theta)$ is the steering vector assuming omni-directional microphones with a flat frequency response,

$$\mathbf{A}(\omega,\theta) = \begin{bmatrix} A_0(\omega,\theta)\mathbf{I}_L & & & \\ & A_1(\omega,\theta)\mathbf{I}_L & & \\ & & \ddots & \\ & & & A_{N-1}(\omega,\theta)\mathbf{I}_L \end{bmatrix} \tag{4}$$

$$\mathbf{g}(\omega,\theta) = \begin{bmatrix}\mathbf{e}^T(\omega)e^{-j\omega\tau_0(\theta)} & \dots & \mathbf{e}^T(\omega)e^{-j\omega\tau_{N-1}(\theta)}\end{bmatrix}^T , \tag{5}$$

with $\mathbf{e}(\omega) = \begin{bmatrix} 1 & e^{-j\omega} & \dots & e^{-j(L-1)\omega}\end{bmatrix}^T$ and $\mathbf{I}_L$ the $L\times L$ identity matrix. The delay $\tau_n(\theta)$ is equal to

$$\tau_n(\theta) = \frac{d_n\cos\theta}{c} f_s , \tag{6}$$

with c the speed of sound ($c = 340\,\frac{m}{s}$) and f_s the sampling frequency. The steering vector $\bar{\mathbf{g}}(\omega,\theta)$ can be decomposed into a real and an imaginary part, $\bar{\mathbf{g}}(\omega,\theta) = \bar{\mathbf{g}}_R(\omega,\theta) + j\bar{\mathbf{g}}_I(\omega,\theta)$, where

$$\bar{\mathbf{g}}_R(\omega,\theta) = \mathbf{A}_R(\omega,\theta)\mathbf{g}_R(\omega,\theta) - \mathbf{A}_I(\omega,\theta)\mathbf{g}_I(\omega,\theta) , \tag{7}$$

with $\mathbf{A}_R(\omega,\theta)$ and $\mathbf{A}_I(\omega,\theta)$ the real and imaginary part of $\mathbf{A}(\omega,\theta)$ and $\mathbf{g}_R(\omega,\theta)$ and $\mathbf{g}_I(\omega,\theta)$ the real and imaginary part of $\mathbf{g}(\omega,\theta)$.

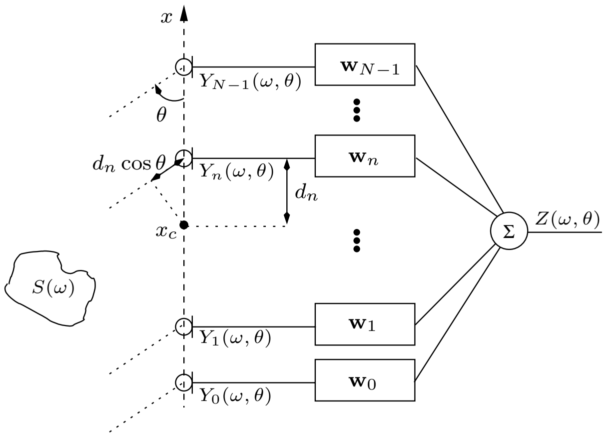

Fig. 1. Microphone array configuration (far-field assumption)

Using (2), the spatial directivity spectrum $|H(\omega,\theta)|^2$ is equal to

$$|H(\omega,\theta)|^2 = H(\omega,\theta)H^*(\omega,\theta) = \mathbf{w}^T\bar{\mathbf{G}}(\omega,\theta)\mathbf{w} \ , \tag{8}$$

with $\bar{\mathbf{G}}(\omega,\theta) = \bar{\mathbf{g}}(\omega,\theta)\bar{\mathbf{g}}^H(\omega,\theta)$, which can be written as

$$\bar{\mathbf{G}}(\omega,\theta) = \mathbf{A}(\omega,\theta)\cdot\mathbf{G}(\omega,\theta)\cdot\mathbf{A}^H(\omega,\theta) \ , \tag{9}$$

with $\mathbf{G}(\omega,\theta) = \mathbf{g}(\omega,\theta)\mathbf{g}^H(\omega,\theta)$. The matrix $\bar{\mathbf{G}}(\omega,\theta)$ can be decomposed into a real and an imaginary part $\bar{\mathbf{G}}_R(\omega,\theta)$ and $\bar{\mathbf{G}}_I(\omega,\theta)$. Since $\bar{\mathbf{G}}_I(\omega,\theta)$ is anti-symmetric, $|H(\omega,\theta)|^2$ is equal to

$$|H(\omega,\theta)|^2 = \mathbf{w}^T\bar{\mathbf{G}}_R(\omega,\theta)\mathbf{w} \ , \tag{10}$$

where the real part $\bar{\mathbf{G}}_R(\omega,\theta)$ can be written as

$$\mathbf{A}_R(\omega,\theta)\mathbf{G}_R(\omega,\theta)\mathbf{A}_R(\omega,\theta)+\mathbf{A}_I(\omega,\theta)\mathbf{G}_R(\omega,\theta)\mathbf{A}_I(\omega,\theta) - \mathbf{A}_I(\omega,\theta)\mathbf{G}_I(\omega,\theta)\mathbf{A}_R(\omega,\theta)+\mathbf{A}_R(\omega,\theta)\mathbf{G}_I(\omega,\theta)\mathbf{A}_I(\omega,\theta), \tag{11}$$

with $\mathbf{G}_R(\omega,\theta)$ and $\mathbf{G}_I(\omega,\theta)$ the real and imaginary part of $\mathbf{G}(\omega,\theta)$.

3. WEIGHTED LEAST-SQUARES COST FUNCTION

The design of a broadband beamformer consists of calculating the filter $\mathbf{w}$, such that $H(\omega,\theta)$ optimally fits the desired spatial directivity pattern $D(\omega,\theta)$, where $D(\omega,\theta)$ is an arbitrary 2-dimensional function. Several design procedures exist, depending on the specific cost function which is optimised. In this paper, we will only consider the weighted least-squares cost function. In [4][5], also eigenfilter-based and non-linear cost functions are discussed.

Considering the least-squares (LS) error $|H(\omega,\theta)-D(\omega,\theta)|^2$, the weighted LS cost function is defined as

$$J_{LS}(\mathbf{w}) = \int_\Theta\int_\Omega F(\omega,\theta)\big[H(\omega,\theta) - D(\omega,\theta)\big]^2 d\omega d\theta \ , \tag{12}$$

where $F(\omega,\theta)$ is a positive real weighting function, assigning more or less importance to certain frequencies and angles. This cost function can be written as the quadratic function

$$J_{LS}(\mathbf{w}) = \mathbf{w}^T\bar{\mathbf{Q}}_{LS}\mathbf{w} - 2\mathbf{w}^T\bar{\mathbf{a}} + d_{LS} \ , \tag{13}$$

with (assuming $D(\omega,\theta)$ to be real-valued)

$$\bar{\mathbf{Q}}_{LS} = \int_\Theta\int_\Omega F(\omega,\theta)\bar{\mathbf{G}}_R(\omega,\theta)d\omega d\theta \tag{14}$$

$$\bar{\mathbf{a}} = \int_\Theta\int_\Omega F(\omega,\theta)D(\omega,\theta)\bar{\mathbf{g}}_R(\omega,\theta)d\omega d\theta \tag{15}$$

$$d_{LS} = \int_\Theta\int_\Omega F(\omega,\theta)D^2(\omega,\theta)d\omega d\theta \ . \tag{16}$$

The filter $\mathbf{w}_{LS}$, minimising the weighted LS cost function, is

$$\mathbf{w}_{LS} = \bar{\mathbf{Q}}_{LS}^{-1}\,\bar{\mathbf{a}} \ . \tag{17}$$

When the microphone characteristics are independent of ω and θ (i.e. for omni-directional, frequency-flat microphones), $\bar{\mathbf{a}}$ and $\bar{\mathbf{Q}}_{LS}$ are equal to

$$\bar{\mathbf{a}} = \mathbf{A}_R\,\mathbf{a} - \mathbf{A}_I\,\mathbf{a}^\circ \tag{18}$$

$$\bar{\mathbf{Q}}_{LS} = \mathbf{A}_R\mathbf{Q}_{LS}\mathbf{A}_R + \mathbf{A}_I\mathbf{Q}_{LS}\mathbf{A}_I - \mathbf{A}_I\mathbf{Q}_{LS}^\circ\mathbf{A}_R + \mathbf{A}_R\mathbf{Q}_{LS}^\circ\mathbf{A}_I$$

with

$$\mathbf{a} = \int_\Theta\int_\Omega F(\omega,\theta)D(\omega,\theta)\mathbf{g}_R(\omega,\theta)d\omega d\theta \tag{19}$$

$$\mathbf{a}^\circ = \int_\Theta\int_\Omega F(\omega,\theta)D(\omega,\theta)\mathbf{g}_I(\omega,\theta)d\omega d\theta \tag{20}$$

$$\mathbf{Q}_{LS} = \int_\Theta\int_\Omega F(\omega,\theta)\mathbf{G}_R(\omega,\theta)d\omega d\theta \tag{21}$$

$$\mathbf{Q}_{LS}^\circ = \int_\Theta\int_\Omega F(\omega,\theta)\mathbf{G}_I(\omega,\theta)d\omega d\theta \ . \tag{22}$$

The ith element of $\bar{\mathbf{a}}$ and the (i,j)-th element of $\bar{\mathbf{Q}}_{LS}$ are equal to

$$\bar{\mathbf{a}}^i = a_n\big(\cos\gamma_n\,\mathbf{a}^i + \sin\gamma_n\,\mathbf{a}^{\circ,i}\big) \tag{23}$$

$$\bar{\mathbf{Q}}_{LS}^{ij} = a_n a_m\Big(\cos\big(\gamma_n-\gamma_m\big)\,\mathbf{Q}_{LS}^{ij} + \sin\big(\gamma_n-\gamma_m\big)\,\mathbf{Q}_{LS}^{\circ,ij}\Big), \tag{24}$$

with $n = \lfloor\frac{i-1}{L}\rfloor$ and $m = \lfloor\frac{j-1}{L}\rfloor$.

4. ROBUST BROADBAND BEAMFORMING

Using the cost function in Section 3, it is possible to design beamformers when the microphone characteristics are exactly known. However, small deviations from the assumed microphone characteristics can lead to large deviations from the desired spatial directivity pattern [7][9]. Since in practice it is difficult to manufacture microphones with the same nominal characteristics, a measurement or calibration procedure is required in order to obtain the true microphone characteristics. However, after calibration the microphone characteristics can still drift over time. Instead of measuring or calibrating every individual microphone array, it is better to consider all feasible microphone characteristics and to either optimise:

- the *mean performance*, i.e. the weighted sum of the cost functions, using the probability of the microphone characteristics as weights (cfr. Section 4.1).
- the *worst-case performance*, i.e. the maximum cost function, leading to a minimax criterion (cfr. Section 4.2).

The same problem of gain and phase errors has been studied in [9]. However, in [9] only the narrowband case for a specific directivity pattern and a uniform pdf has been considered. The approach presented here is more general because we consider broadband beamformers with an arbitrary spatial directivity pattern, arbitrary probability density functions and several cost functions [4]. However, in this paper we will only use the weighted LS cost function. We refer to [4] for robust design based on other cost functions.

4.1. Weighted sum using probability density functions

The total cost function $J_{LS}^{tot}(\mathbf{w})$ is defined as the weighted sum of the cost functions for all feasible microphone characteristics, using the probability of the microphone characteristics as weights, i.e.

$$J_{LS}^{tot}(\mathbf{w}) = \int_{A_0}\dots\int_{A_{N-1}} J_{LS}(\mathbf{w})\,f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\,dA_0\dots dA_{N-1} \ , \tag{25}$$

with $f_{\mathcal{A}}(A)$ the pdf of the stochastic variable $A = ae^{-j\gamma}$, i.e. the joint pdf of the stochastic variables a (gain) and γ (phase). We assume that $f_{\mathcal{A}}(A)$ is independent of frequency and angle, or that $f_{\mathcal{A}}(A)$ is available for different frequency-angle regions. Without loss of generality, we also assume that all microphone characteristics are described by the same pdf $f_{\mathcal{A}}(A)$. Furthermore, we assume that a and γ are independent variables, such that the joint pdf is separable, i.e. $f_{\mathcal{A}}(A) = f_\alpha(a) f_{\mathcal{G}}(\gamma)$, with $f_\alpha(a)$ the pdf of the gain a and $f_{\mathcal{G}}(\gamma)$ the pdf of the phase γ. These pdfs are normalised such that $\int_a f_\alpha(a)\, da = 1$ and $\int_\gamma f_{\mathcal{G}}(\gamma)\, d\gamma = 1$.
By combining (13) and (25), the total cost function is equal to

$$J_{LS}^{tot}(\mathbf{w}) = \mathbf{w}^T \bar{\mathbf{Q}}_{tot} \mathbf{w} - 2\mathbf{w}^T \bar{\mathbf{a}}_{tot} + d_{LS} \ , \qquad (26)$$

which has the same form as (13), with

$$\bar{\mathbf{a}}_{tot} = \int_{A_0} \ldots \int_{A_{N-1}} \bar{\mathbf{a}}\, f_{\mathcal{A}}(A_0) \ldots f_{\mathcal{A}}(A_{N-1})\, dA_0 \ldots dA_{N-1}$$

$$\bar{\mathbf{Q}}_{tot} = \int_{A_0} \ldots \int_{A_{N-1}} \bar{\mathbf{Q}}_{LS}\, f_{\mathcal{A}}(A_0) \ldots f_{\mathcal{A}}(A_{N-1})\, dA_0 \ldots dA_{N-1} \ .$$

Using (23), the ith element of $\bar{\mathbf{a}}_{tot}$ is equal to

$$\bar{\mathbf{a}}_{tot}^i = \int_{a_n}\int_{\gamma_n} a_n \big(\cos\gamma_n\, \mathbf{a}^i + \sin\gamma_n\, \mathbf{a}^{\circ,i} \big) f_\alpha(a_n) f_{\mathcal{G}}(\gamma_n) da_n d\gamma_n,$$

such that

$$\bar{\mathbf{a}}_{tot} = \mu_a \mu_\gamma^c\, \mathbf{a} + \mu_a \mu_\gamma^s\, \mathbf{a}^\circ \ , \qquad (27)$$

with μ_a the mean of the gain pdf, $\mu_a = \int_a a f_\alpha(a) da$, and μ_γ^c and μ_γ^s equal to

$$\mu_\gamma^c = \int_\gamma \cos\gamma f_{\mathcal{G}}(\gamma) d\gamma, \quad \mu_\gamma^s = \int_\gamma \sin\gamma f_{\mathcal{G}}(\gamma) d\gamma \ . \qquad (28)$$

Using (24), the (i,j)-th element of $\bar{\mathbf{Q}}_{tot}$ is equal to

$$\int_{a_n}\int_{a_m}\int_{\gamma_n}\int_{\gamma_m} a_n a_m \Big(\cos(\gamma_n - \gamma_m)\, \mathbf{Q}_{LS}^{ij} + \sin(\gamma_n - \gamma_m)\, \mathbf{Q}_{LS}^{\circ,ij} \Big)$$
$$f_\alpha(a_n) f_\alpha(a_m) f_{\mathcal{G}}(\gamma_n) f_{\mathcal{G}}(\gamma_m) da_n da_m d\gamma_n d\gamma_m \ . \qquad (29)$$

If $n = m$, $\bar{\mathbf{Q}}_{tot}^{ij}$ is equal to

$$\bar{\mathbf{Q}}_{tot}^{ij} = \int_{a_n} a_n^2 f_\alpha(a_n) da_n\, \mathbf{Q}_{LS}^{ij} = \sigma_a^2\, \mathbf{Q}_{LS}^{ij} \ , \qquad (30)$$

with σ_a^2 the variance of the gain pdf, i.e. $\sigma_a^2 = \int_a a^2 f_\alpha(a) da$.
If $n \neq m$, $\bar{\mathbf{Q}}_{tot}^{ij}$ is equal to

$$\bar{\mathbf{Q}}_{tot}^{ij} = \mu_a^2 \left[\sigma_\gamma^c \mathbf{Q}_{LS}^{ij} + \sigma_\gamma^s \mathbf{Q}_{LS}^{\circ,ij} \right] \ , \qquad (31)$$

with

$$\sigma_\gamma^c = \int_{\gamma_1}\int_{\gamma_2} \cos(\gamma_1 - \gamma_2) f_{\mathcal{G}}(\gamma_1) f_{\mathcal{G}}(\gamma_2) d\gamma_1 d\gamma_2 \qquad (32)$$

$$\sigma_\gamma^s = \int_{\gamma_1}\int_{\gamma_2} \sin(\gamma_1 - \gamma_2) f_{\mathcal{G}}(\gamma_1) f_{\mathcal{G}}(\gamma_2) d\gamma_1 d\gamma_2 \ , \qquad (33)$$

such that

$$\sigma_\gamma^c = (\mu_\gamma^c)^2 + (\mu_\gamma^s)^2, \qquad \sigma_\gamma^s = \mu_\gamma^s \mu_\gamma^c - \mu_\gamma^c \mu_\gamma^s = 0 \ . \qquad (34)$$

The matrix $\bar{\mathbf{Q}}_{tot}$ can now be computed as

$$\bar{\mathbf{Q}}_{tot} = \begin{bmatrix} \sigma_a^2\, \mathbf{1}_L & \mu_a^2 \sigma_\gamma^c\, \mathbf{1}_L & \ldots & \mu_a^2 \sigma_\gamma^c\, \mathbf{1}_L \\ \mu_a^2 \sigma_\gamma^c\, \mathbf{1}_L & \sigma_a^2 \mathbf{1}_L & \ldots & \mu_a^2 \sigma_\gamma^c\, \mathbf{1}_L \\ \vdots & \vdots & & \vdots \\ \mu_a^2 \sigma_\gamma^c\, \mathbf{1}_L & \mu_a^2 \sigma_\gamma^c & \ldots & \sigma_a^2\, \mathbf{1}_L \end{bmatrix} \odot \mathbf{Q}_{LS} \ ,$$

where $\mathbf{1}_L$ is an $L \times L$-matrix with all elements equal to 1 and $\odot$ denoting element-wise multiplication. As can be seen, we only need the mean and the variance of the gain pdf $f_\alpha(a)$, whereas in general complete knowledge of the phase pdf $f_{\mathcal{G}}(\gamma)$ is required.

When optimising the mean performance, it is however still possible - although typically with a low probability - that for some specific gain/phase combination, the cost function is quite high. If this is considered to be a problem, the worst-case performance should be optimised (cfr. Section 4.2).

4.2. Minimax criterion

For the minimax criterion, we first have to define a (finite) set of microphone characteristics (K_a gain values and K_γ phase values),

$$\{a_{min} = a_1, \ldots, a_{K_a} = a_{max}\}, \ \{\gamma_{min} = \gamma_1, \ldots, \gamma_{K_\gamma} = \gamma_{max}\},$$

as an approximation for the continuum of feasible microphone characteristics, and use this set of gain and phase values to construct the $(K_a K_\gamma)^N$-dimensional vector $\mathbf{F}(\mathbf{w})$,

$$\mathbf{F}(\mathbf{w}) = \left[\ F_1(\mathbf{w}) \quad F_2(\mathbf{w}) \quad \ldots \quad F_{(K_a K_\gamma)^N}(\mathbf{w}) \ \right]^T \ , \qquad (35)$$

which consists of the used cost function (weighted LS or any other cost function) at each possible combination of gain and phase values. The goal then is to minimise the L_∞-norm of $\mathbf{F}(\mathbf{w})$, i.e. the maximum value of the elements $F_k(\mathbf{w})$,

$$\min_{\mathbf{w}} \|\mathbf{F}(\mathbf{w})\|_\infty = \min_{\mathbf{w}} \max_k F_k(\mathbf{w}) \ , \qquad (36)$$

which can be done using a sequential quadratic programming (SQP) method [10]. In order to improve the numerical robustness and the convergence speed, the gradient

$$\left[\ \frac{\partial F_1(\mathbf{w})}{\partial \mathbf{w}} \quad \frac{\partial F_2(\mathbf{w})}{\partial \mathbf{w}} \quad \ldots \quad \frac{\partial F_{(K_a K_\gamma)^N}(\mathbf{w})}{\partial \mathbf{w}} \ \right] \ , \qquad (37)$$

which is an $M \times (K_a K_\gamma)^N$-dimensional matrix, can be supplied analytically. As can be seen, the larger K_a and K_γ, the denser the grid of feasible microphone characteristics, and the higher the computational complexity for solving the minimax problem.

When only considering gain errors and using the weighted LS cost function, it can be proven that *for any* $\mathbf{w}$*, the maximum value of* $\mathbf{F}(\mathbf{w})$ *occurs on a boundary point of an* N*-dimensional hypercube* [4], i.e. $a_n = a_{min}$ or $a_n = a_{max}$, $n = 0 \ldots N-1$. This implies that $K_a = 2$ suffices and $\mathbf{F}(\mathbf{w})$ consists of 2^N elements.

5. SIMULATIONS

We have performed simulations using a small-sized non-uniform linear microphone array consisting of $N = 3$ microphones at positions $\left[-0.01 \quad 0 \quad 0.015 \right]$ m. We have designed an end-fire beamformer with passband specifications $(\Omega_p, \Theta_p) = (300\text{–}4000\,\text{Hz}, 0°\text{–}60°)$ and stopband specifications $(\Omega_s, \Theta_s) = (300\text{–}4000\,\text{Hz}, 80°\text{–}180°)$ and $f_s = 8\,\text{kHz}$. The filter length $L = 20$ and the weighting function $F(\omega, \theta) = 1$. We have designed several types of broadband beamformers using the weighted LS cost function:

1. a non-robust beamformer (i.e. assuming $a_n = 1, \gamma_n = 0°$)
2. a robust beamformer using a uniform gain pdf $(0.85, 1.15)$
3. a robust beamformer using a uniform phase pdf $(-5°, 10°)$
4. a robust beamformer using a uniform gain/phase pdf
5. a robust beamformer using the minimax criterion (only gain errors, $a_{min} = 0.85$, $a_{max} = 1.15$, $K_a = 2$)

For all beamformers, we have computed the following cost functions:

1. the cost function J without phase and gain errors ($A_n = 1$)

2. the cost function J_{dev} for microphone gains $[\,0.9 \quad 1.1 \quad 1.05\,]$
3. the mean cost function J_a^{tot} for the uniform gain pdf
4. the mean cost function J_γ^{tot} for the uniform phase pdf
5. the mean cost function J_A^{tot} for the uniform gain/phase pdf
6. the maximum cost function J_{max} when the gain varies between $a_{min} = 0.85$ and $a_{max} = 1.15$

Table 1 summarises the different cost functions. Obviously, the design procedure optimising a specific cost function leads to the best value for this cost function (bold values). This implies that when no errors occur, the robust design procedures give rise to a higher cost function J than the non-robust design procedure. However, when gain and/or phase errors occur, the non-robust design procedure produces very bad results (e.g. compare J_{max} for all design procedures and see Figure 3), whereas all robust design procedures produce satisfactory results.

Figure 2 shows the spatial directivity plots of the non-robust, the gain/phase-robust and the minimax beamformer for several frequencies, when no gain and phase errors occur. As can be seen, the performance of the non-robust beamformer is the best, but the performance of the robust beamformers is certainly acceptable.

Figure 3 shows the spatial directivity plots in case of (small) gain and phase errors (microphone gains = $[\,0.9 \quad 1.1 \quad 1.05\,]$ and phases = $[\,5^\circ \quad -2^\circ \quad 5^\circ\,]$). As can be seen, the performance of the non-robust beamformer deteriorates considerably. Certainly for the low frequencies, the spatial directivity pattern is almost omni-directional and the amplification is very high. On the other hand, the robust beamformers retain the desired spatial directivity pattern, even when gain and phase errors occur.

Design	J	J_{dev}	J_a^{tot}	J_γ^{tot}	J_A^{tot}	J_{max}
Non-robust	**0.313**	220.1	123.3	62.67	185.7	961.3
Gain	0.474	0.685	**0.642**	0.576	0.744	1.441
Phase	0.431	0.700	0.666	**0.557**	0.791	1.749
Gain-phase	0.518	**0.652**	0.653	0.596	**0.732**	1.368
Minimax	0.747	0.843	0.803	0.792	0.849	**1.035**

Table 1. Different cost functions for robust beamformer design

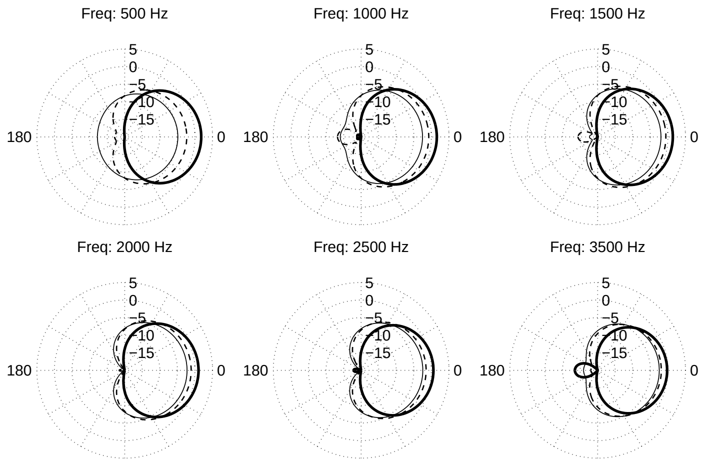

Fig. 2. Spatial directivity plots, no gain and phase errors (non-robust: thick solid, gain/phase-robust: dashed, minimax: solid)

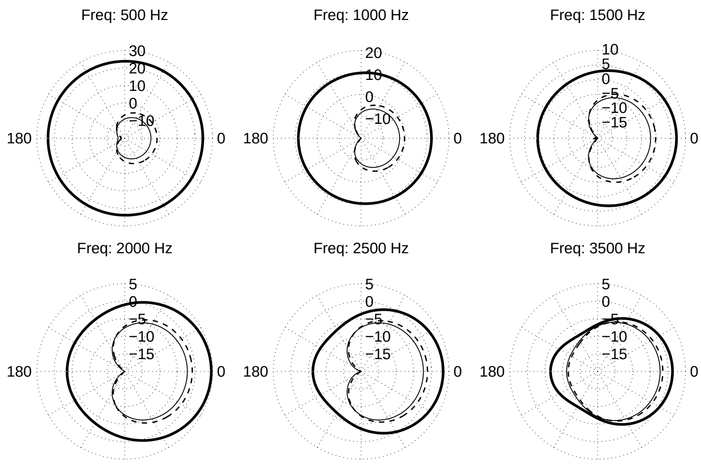

Fig. 3. Spatial directivity plots, gain and phase errors (non-robust: thick solid, gain/phase-robust: dashed, minimax: solid)

6. CONCLUSION

In this paper we have described two procedures for designing broadband speech beamformers that are robust against gain and phase errors. The first design procedure optimises the mean performance using gain and phase pdfs, whereas the second design procedure optimises the worst-case performance using a minimax criterion. Simulations for both design procedures show the performance improvement that is obtained when gain and/or phase errors occur.

7. REFERENCES

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[2] S. Van Gerven, D. Van Compernolle, P. Wauters, W. Verstraeten, K. Eneman, and K. Delaet, "Multiple beam broadband beamforming: Filter design and real-time implementation," in *Proc. IEEE Workshop Appl. of Signal Proc. to Audio and Acoustics*, New Paltz, USA, Oct. 1995, pp. 173–176.

[3] R. W. Stadler and W. M. Rabinowitz, "On the potential of fixed arrays for hearing aids," *Journal of the Acoustical Society of America*, vol. 94, no. 3, pp. 1332–1342, Sept. 1993.

[4] S. Doclo and M. Moonen, "Far-field and near-field broadband beamformer design," Tech. Rep. ESAT-SISTA/TR 2002-109, ESAT, K.U.Leuven, Belgium, July 2002.

[5] S. Doclo and M. Moonen, "Design of far-field broadband beamformers using eigenfilters," in *Proc. European Signal Proc. Conf.*, Toulouse, France, Sept. 2002, pp. III 237-240.

[6] M. Kajala and M. Hämäläinen, "Broadband beamforming optimization for speech enhancement in noisy environments," in *Proc. IEEE Workshop Appl. of Signal Proc. to Audio and Acoust.*, New Paltz, USA, Oct. 1999, pp. 19–22.

[7] H. Cox, R. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 3, pp. 393–398, June 1986.

[8] C. Sydow, "Broadband beamforming for a microphone array," *Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 845–849, Aug. 1994.

[9] M. H. Er, "A robust formulation for an optimum beamformer subject to amplitude and phase perturbations," *Signal Processing*, vol. 19, no. 1, pp. 17–26, 1990.

[10] R. Fletcher, *Practical Methods of Optimization*, Wiley, New York, 1987.