SUPERDIRECTIVE BEAMFORMING ROBUST AGAINST MICROPHONE MISMATCH

Simon Doclo, Marc Moonen

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

ABSTRACT

Fixed superdirective beamformers using small-size microphone arrays are known to be highly sensitive to errors in the assumed microphone array characteristics. This paper discusses the design of robust superdirective beamformers by taking into account the statistics of the microphone characteristics. Different design procedures are considered: applying a white noise gain constraint, trading off the mean noise and distortion energy, and maximizing the mean or the minimum directivity factor. When computational complexity is not important, maximizing the mean or the minimum directivity factor is the preferred design procedure. In addition, it is shown how to determine a suitable parameter range for the other design procedures.

1. INTRODUCTION

In many speech communication applications, the microphone signals are corrupted by background noise and reverberation. The objective of a fixed (data-independent) beamformer is to obtain spatial focusing on the speech source, thereby reducing noise and reverberation not coming from the same direction as the speech source. Different types of fixed beamformers are available, e.g. delay-and-sum beamforming, superdirective beamforming [1, 2], differential microphone arrays [3], and frequency-invariant beamforming.

It is well known that a superdirective beamformer, which maximizes the directivity factor of the array, is sensitive to uncorrelated noise, especially at low frequencies and for small-size arrays [1, 2]. In addition, superdirective beamformers are sensitive to deviations from the assumed microphone characteristics (gain, phase, and position). In many applications, these microphone array characteristics are not exactly known and can even change over time.

This paper discusses several design procedures for improving the robustness of superdirective beamformers against unknown microphone mismatch. A commonly used technique to limit the amplification of uncorrelated noise components, which also inherently increases the robustness against microphone mismatch, is to impose a white noise gain constraint [1, 2]. In addition, we discuss two design procedures that optimize a mean performance criterion, i.e. the weighted sum of the mean noise and distortion energy, and the mean (or the minimum) directivity factor. These design procedures obviously require knowledge of the gain, phase and position probability density functions and are related to [4, 5] where the design of robust beamformers with an arbitrary spatial directivity pattern has been discussed. When computational complexity is not an issue, maximizing the mean or the minimum directivity factor is the preferred design procedure. In addition, it is shown how to determine a suitable parameter range for the other design procedures.

Simon Doclo is a Postdoctoral Fellow of the Research Foundation - Flanders (FWO - Vlaanderen). This work was carried out at the ESAT-SCD laboratory, Katholieke Universiteit Leuven, in the frame of the FWO Project G.0233.01, the IWT Projects 020540 and 040803, the Concerted Research Action GOA-AMBIORICS, and the Interuniversity Attraction Pole IUAP P5-22.

2. CONFIGURATION AND DEFINITIONS

Consider the linear microphone array depicted in Fig. 1, consisting of N microphones and with d_n the distance between the nth microphone and the reference point, arbitrarily chosen here as the center of the microphone array. We assume that a noise field with spectral and spatial characteristics $\sigma_v^2(\omega, \phi, \theta)$ is present, where ϕ and θ represent the azimuthal and the elevation angle in spherical coordinates ($0 \leq \phi < 2\pi$, $0 \leq \theta \leq \pi$), and that a speech source $S(\omega)$ is located at an angle (ϕ_s, θ_s) in the far-field of the microphone array. The microphone characteristics of the nth microphone are described by

$$A_n(\omega, \phi, \theta) = a_n(\omega, \phi, \theta) e^{-j\psi_n(\omega, \phi, \theta)} , \quad (1)$$

where the gain $a_n(\omega, \phi, \theta)$ and phase $\psi_n(\omega, \phi, \theta)$ can be frequency- and angle-dependent. The nth microphone signal $Y_n(\omega)$ is equal to

$$Y_n(\omega) = g_n(\omega, \phi_s, \theta_s) S_r(\omega) + V_n(\omega) , \quad (2)$$

with $S_r(\omega)$ the speech component of the signal received at the reference point, $V_n(\omega)$ the noise component of the nth microphone and

$$g_n(\omega, \phi, \theta) = A_n(\omega, \phi, \theta)\, e^{-j\omega\tau_n(\phi, \theta)} , \quad (3)$$

where the delay $\tau_n(\phi, \theta)$ in number of samples is equal to $\tau_n(\phi, \theta) = (d_n \cos\theta f_s)/c$, with c the speed of sound propagation and f_s the sampling frequency. The stacked vector of microphone signals $\mathbf{Y}(\omega) = \begin{bmatrix} Y_0(\omega) & Y_1(\omega) & \ldots & Y_{N-1}(\omega) \end{bmatrix}^T$ can be written as

$$\mathbf{Y}(\omega) = \mathbf{g}_s(\omega) S_r(\omega) + \mathbf{V}(\omega) , \quad (4)$$

with $\mathbf{g}_s(\omega) = \mathbf{g}(\omega, \phi_s, \theta_s)$, the *steering vector* $\mathbf{g}(\omega, \phi, \theta)$ equal to

$$\mathbf{g}(\omega, \phi, \theta) = \begin{bmatrix} g_0(\omega, \phi, \theta) & g_1(\omega, \phi, \theta) & \ldots & g_{N-1}(\omega, \phi, \theta) \end{bmatrix}^T , \quad (5)$$

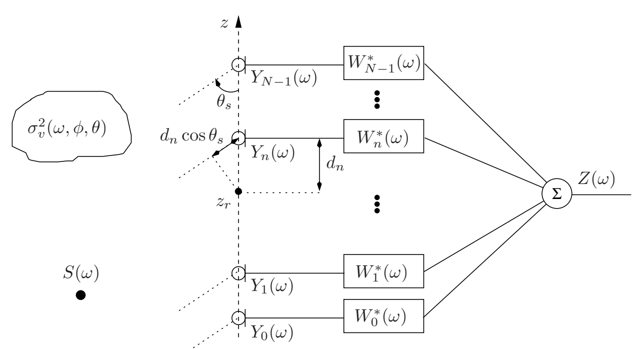

Fig. 1. Linear microphone array configuration

and $\mathbf{V}(\omega)$ defined similarly as $\mathbf{Y}(\omega)$. The output signal $Z(\omega)$ is

$$Z(\omega) = \mathbf{W}^H(\omega)\mathbf{Y}(\omega) = \mathbf{W}^H(\omega)\mathbf{g}_s(\omega)S_r(\omega) + \mathbf{W}^H(\omega)\mathbf{V}(\omega),$$

with $\mathbf{W}(\omega)$ the weight vector of the beamformer.

The *array gain* $G(\omega)$ is defined as the signal-to-noise ratio (SNR) improvement between the reference (input) signal and the microphone array output signal, and is equal to

$$G(\omega) = \frac{|\mathbf{W}(\omega)^H \mathbf{g}_s(\omega)|^2}{\mathbf{W}^H(\omega)\widetilde{\mathbf{\Phi}}_{VV}(\omega)\mathbf{W}(\omega)} , \qquad (6)$$

with $\widetilde{\mathbf{\Phi}}_{VV}(\omega)$ the normalized noise correlation matrix, i.e. $\widetilde{\mathbf{\Phi}}_{VV}(\omega) = \mathbf{\Phi}_{VV}(\omega)/\Phi_v(\omega) = \mathcal{E}\{\mathbf{V}(\omega)\mathbf{V}^H(\omega)\}/\Phi_v(\omega)$, with $\Phi_v(\omega)$ the noise energy of the reference signal. By spatially integrating the noise field $\sigma_v^2(\omega,\phi,\theta)$, the (n,p)-th element of $\widetilde{\mathbf{\Phi}}_{VV}(\omega)$ can be computed as

$$\widetilde{\Phi}_{V_nV_p}(\omega) = \frac{\int_0^{2\pi}\int_0^{\pi} g_n(\omega,\phi,\theta)g_p^*(\omega,\phi,\theta)\,\sigma_v^2(\omega,\phi,\theta)\,\sin\theta\,d\theta d\phi}{\int_0^{2\pi}\int_0^{\pi}\sigma_v^2(\omega,\phi,\theta)\,\sin\theta\,d\theta d\phi} . \qquad (7)$$

The *directivity factor* (DF) is defined as the ability to suppress spherically isotropic noise (diffuse noise), for which $\sigma_v^2(\omega,\phi,\theta) = \sigma_v^2(\omega)$. Hence, the directivity factor is equal to

$$\boxed{DF(\omega) = \frac{|\mathbf{W}(\omega)^H \mathbf{g}_s(\omega)|^2}{\mathbf{W}^H(\omega)\widetilde{\mathbf{\Phi}}_{VV}^{diff}(\omega)\mathbf{W}(\omega)}} \qquad (8)$$

where, using (3) and (7), the (n,p)-th element of $\widetilde{\mathbf{\Phi}}_{VV}^{diff}(\omega)$ is

$$\widetilde{\Phi}_{V_nV_p}^{diff}(\omega) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} A_n(\omega,\phi,\theta)A_p^*(\omega,\phi,\theta)\cdot e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}\sin\theta\,d\theta d\phi . \qquad (9)$$

The *white noise gain* (WNG) is defined as the ability to suppress spatially uncorrelated noise (e.g. internal noise of the microphones), for which the normalized noise correlation matrix $\widetilde{\mathbf{\Phi}}_{VV}^{unc}(\omega) = \mathbf{I}_N$, with $\mathbf{I}_N$ the identity matrix. Hence, the white noise gain is equal to

$$\boxed{WNG(\omega) = \frac{|\mathbf{W}(\omega)^H \mathbf{g}_s(\omega)|^2}{\mathbf{W}^H(\omega)\mathbf{W}(\omega)}} \qquad (10)$$

For conciseness, we will omit the frequency-domain variable ω where possible in the remainder of the paper.

3. SUPERDIRECTIVE BEAMFORMING

The superdirective beamformer $\mathbf{W}_{sd}$ maximizes the directivity factor defined in (8). By imposing a unit gain constraint in the direction of the speech source, i.e. $\mathbf{W}^H\mathbf{g}_s = 1$, the superdirective beamformer $\mathbf{W}_{sd}$ can be computed as

$$\boxed{\mathbf{W}_{sd} = \frac{\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff}\right)^{-1}\mathbf{g}_s}{\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff}\right)^{-1}\mathbf{g}_s}} \qquad (11)$$

The same solution is obtained by minimizing the normalized noise energy in the output signal, subject to a unit gain constraint in the direction of the speech source, i.e.

$$\min_{\mathbf{W}} \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{diff}\mathbf{W}, \quad \text{subject to } \mathbf{W}^H\mathbf{g}_s = 1 . \qquad (12)$$

Similarly, consider the weighted sum of the normalized noise energy $J_v(\mathbf{W})$ and distortion energy $J_d(\mathbf{W})$ in the output signal, i.e.

$$J_t(\mathbf{W},\lambda) = J_v(\mathbf{W}) + \lambda J_d(\mathbf{W}) , \qquad (13)$$

where $\lambda \geq 0$ is a weighting factor and

$$\boxed{J_v(\mathbf{W}) = \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{diff}\mathbf{W}, \quad J_d(\mathbf{W}) = |\mathbf{W}^H\mathbf{g}_s - 1|^2} \qquad (14)$$

The filter $\mathbf{W}_t(\lambda)$ minimizing $J_t(\mathbf{W},\lambda)$ is equal to

$$\mathbf{W}_t(\lambda) = \left(\widetilde{\mathbf{\Phi}}_{VV}^{diff} + \lambda\mathbf{g}_s\mathbf{g}_s^H\right)^{-1}\lambda\mathbf{g}_s = \frac{\lambda\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff}\right)^{-1}\mathbf{g}_s}{1+\lambda\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff}\right)^{-1}\mathbf{g}_s} . \qquad (15)$$

Note that $\mathbf{W}_{sd} = \mathbf{W}_t(\infty)$. It can be easily shown that the larger λ, the larger the noise energy and the smaller the distortion energy.

It is well known that superdirective beamformers are sensitive to uncorrelated noise, especially at low frequencies. A commonly used technique to limit the amplification of uncorrelated noise components, is to impose a WNG constraint [1, 2], such that the optimization problem in (12) becomes

$$\min_{\mathbf{W}} \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{diff}\mathbf{W}, \quad \text{subject to } \mathbf{W}^H\mathbf{g}_s = 1,\ \mathbf{W}^H\mathbf{W} \leq \beta . \qquad (16)$$

Using the method of Lagrange multipliers, it can be easily shown that the solution of this optimization problem has the form

$$\boxed{\mathbf{W}_{sd,\mu} = \frac{\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff} + \mu\mathbf{I}_N\right)^{-1}\mathbf{g}_s}{\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{diff} + \mu\mathbf{I}_N\right)^{-1}\mathbf{g}_s}} \qquad (17)$$

The Lagrange multiplier μ needs to be (iteratively) determined such that the inequality constraint $\mathbf{W}_{sd,\mu}^H\mathbf{W}_{sd,\mu} \leq \beta$ is satisfied [1, 2]. The larger μ, the larger the robustness of the beamformer, but the smaller its directivity factor.

4. ROBUST SUPERDIRECTIVE BEAMFORMING

Using the procedures in Section 3, it is possible to design a superdirective beamformer when the microphone characteristics and positions are exactly known. However, superdirective beamformers are highly sensitive to deviations from the assumed microphone characteristics, especially for small-size arrays and at low frequencies. In Section 3, it has been shown that robustness can be improved by imposing a WNG constraint. However, since the WNG is not directly related to microphone mismatch, it is quite difficult to choose a suitable value for β or μ that guarantees robustness for a range of microphone mismatches. In this section, we present design procedures for improving the robustness against unknown microphone mismatch by optimizing the *mean performance*, i.e. the weighted sum for all feasible microphone characteristics, using the probability of the microphone characteristics as weights. These procedures obviously require knowledge of the gain, phase and position probability density functions (pdf). We will discuss two performance criteria: the weighted sum of the mean noise and distortion energy, and the mean (or the minimum) directivity factor.

In order to be able to describe microphone position errors, we will incorporate them directly into the microphone characteristics, i.e.

$$A_n(\omega,\phi,\theta) = a_n(\omega,\phi,\theta)e^{-j\psi_n(\omega,\phi,\theta)}e^{-j\omega\frac{\delta_n\cos\theta}{c}f_s} , \qquad (18)$$

where δ_n represents the position error for the nth microphone. The pdf $f_{\mathcal{A}}(A)$ describes the joint pdf of the stochastic variables a (gain), ψ (phase) and δ (position error). We assume that a, ψ and δ are independent variables, such that the joint pdf is separable.

4.1. Mean noise and distortion energy

Similar to (13), the weighted sum of the mean noise energy $J_{vm}(\mathbf{W})$ and the mean distortion energy $J_{dm}(\mathbf{W})$ is equal to

$$\boxed{J_{tm}(\mathbf{W},\lambda) = J_{vm}(\mathbf{W}) + \lambda J_{dm}(\mathbf{W})\ ,} \tag{19}$$

with

$$J_{vm}(\mathbf{W}) = \int_{A_0} \cdots \int_{A_{N-1}} \mathbf{W}^H \widetilde{\mathbf{\Phi}}_{VV}^{diff}(\mathbf{A})\mathbf{W} \cdot f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1}\ , \tag{20}$$

$$J_{dm}(\mathbf{W}) = \int_{A_0} \cdots \int_{A_{N-1}} |\mathbf{W}^H \mathbf{g}_s(\mathbf{A}) - 1|^2 \cdot f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1}\ , \tag{21}$$

with $\widetilde{\mathbf{\Phi}}_{VV}^{diff}(\mathbf{A})$ the normalized noise correlation matrix in (9) for the specific microphone characteristic $\mathbf{A} = \{A_0, \dots, A_{N-1}\}$, and $\mathbf{g}_s(\mathbf{A})$ the steering vector in (5) and (3) for the angle (ϕ_s, θ_s) and the microphone characteristic $\mathbf{A}$.
The mean distortion energy $J_{dm}(\mathbf{W})$ can be written as

$$\boxed{J_{dm}(\mathbf{W}) = \mathbf{W}^H \mathbf{Q}_{sm}\mathbf{W} - \mathbf{W}^H \mathbf{q}_{sm} - \mathbf{q}_{sm}^H \mathbf{W} + 1} \tag{22}$$

with $\mathbf{Q}_{sm}$ and $\mathbf{q}_{sm}$ equal to

$$\int_{A_0} \cdots \int_{A_{N-1}} \mathbf{g}_s(\mathbf{A})\mathbf{g}_s^H(\mathbf{A})\, f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1},$$

$$\int_{A_0} \cdots \int_{A_{N-1}} \mathbf{g}_s(\mathbf{A})\, f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1}\ .$$

Using (3), the (n,p)-th element of $\mathbf{Q}_{sm}$ is equal to

$$Q_{sm,np} = \sigma^2_{A,np}(\omega,\phi_s,\theta_s)\, e^{-j\omega \frac{(d_n - d_p)\cos\theta_s}{c} f_s}\ , \tag{23}$$

with $\sigma^2_{A,np}(\omega,\phi_s,\theta_s)$ equal to

$$\int_{A_n}\int_{A_p} A_n(\omega,\phi_s,\theta_s) A_p^*(\omega,\phi_s,\theta_s) f_{\mathcal{A}}(A_n) f_{\mathcal{A}}(A_p) dA_n dA_p\ ,$$

and the nth element of $\mathbf{q}_{sm}$ is equal to

$$q_{sm,n} = \underbrace{\left[\int_{A_n} A_n(\omega,\phi_s,\theta_s) f_{\mathcal{A}}(A_n)\, dA_n\right]}_{\mu_{A,n}(\omega,\phi_s,\theta_s)} e^{-j\omega \frac{d_n \cos\theta_s}{c} f_s}\ . \tag{24}$$

The expressions $\sigma^2_{A,np}(\omega,\phi_s,\theta_s)$ and $\mu_{A,n}(\omega,\phi_s,\theta_s)$ can be easily calculated for e.g. uniform or Gaussian pdfs.
The mean noise energy $J_{vm}(\mathbf{W})$ can be written as

$$\boxed{J_{vm}(\mathbf{W}) = \mathbf{W}^H \widetilde{\mathbf{\Phi}}_m^{diff}\mathbf{W}} \tag{25}$$

with $\widetilde{\mathbf{\Phi}}_m^{diff}$ equal to

$$\int_{A_0} \cdots \int_{A_{N-1}} \widetilde{\mathbf{\Phi}}_{VV}^{diff}(\mathbf{A})\, f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1}\ .$$

Using (9), the (n,p)-th element of $\widetilde{\mathbf{\Phi}}_m$ is equal to

$$\widetilde{\Phi}_{m,np}^{diff} = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} \sigma^2_{A,np}(\omega,\phi,\theta) e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))} \sin\theta d\theta d\phi.$$

Similar to (15), the filter $\mathbf{W}_{tm,\lambda}$ minimizing $J_{tm}(\mathbf{W},\lambda)$ is equal to

$$\boxed{\mathbf{W}_{tm,\lambda} = \left(\widetilde{\mathbf{\Phi}}_m^{diff} + \lambda \mathbf{Q}_{sm}\right)^{-1} \lambda \mathbf{q}_{sm}} \tag{26}$$

The larger λ, the larger the mean noise energy and the smaller the mean distortion energy.

4.2. Mean and minimum directivity factor

The *mean directivity factor* is defined as

$$\boxed{DF_m(\mathbf{W}) = \int_{A_0} \cdots \int_{A_{N-1}} DF(\mathbf{W},\mathbf{A})\, f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, dA_0 \dots dA_{N-1}} \tag{27}$$

with $DF(\mathbf{W},\mathbf{A})$ the directivity factor defined in (8) for the microphone characteristic $\mathbf{A}$, i.e.

$$DF(\mathbf{W},\mathbf{A}) = \frac{|\mathbf{W}^H \mathbf{g}_s(\mathbf{A})|^2}{\mathbf{W}^H \widetilde{\mathbf{\Phi}}_{VV}^{diff}(\mathbf{A})\mathbf{W}}\ . \tag{28}$$

Since the filter $\mathbf{W}$ cannot be extracted from the integrals and the separability of the joint pdf $f_{\mathcal{A}}(A)$ cannot be exploited, computing and maximizing the mean directivity factor is computationally quite expensive. In general, we will approximate the integrals in (27) by a discrete (Riemann) sum, i.e.

$$DF_m(\mathbf{W}) \approx \sum_{A_0} \cdots \sum_{A_{N-1}} DF(\mathbf{W},\mathbf{A})\, f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\, \Delta A_0 \dots \Delta A_{N-1}\ , \tag{29}$$

with ΔA_n denoting the grid spacing for the pdf describing the nth microphone characteristic. Obviously, the smaller the grid spacing, the more expensive the computation of this sum. Since no closed-form expression is available for the filter $\mathbf{W}_m$ maximizing (29), an iterative optimization technique will be used.

When maximizing the mean directivity factor, it is still possible that for some specific microphone deviation the directivity factor is quite low. To overcome this problem, the *worst-case performance* can be optimized by maximizing the minimum directivity factor for all feasible microphone characteristics. We first define a (finite) grid of microphone characteristics (K_a gain values, K_ψ phase values and K_δ position error values), as an approximation for the continuum of feasible microphone characteristics. We use this set to construct the $(K_a K_\psi K_\delta)^N$-dimensional vector $\mathbf{F}(\mathbf{W})$, i.e.

$$\mathbf{F}(\mathbf{W}) = \begin{bmatrix} DF_1(\mathbf{W},\mathbf{A}) \\ DF_2(\mathbf{W},\mathbf{A}) \\ \vdots \\ DF_{(K_a K_\psi K_\delta)^N}(\mathbf{W},\mathbf{A}) \end{bmatrix}, \tag{30}$$

consisting of the directivity factor for each possible combination of gain, phase and position error values. The goal then is to maximize the minimum value of $\mathbf{F}(\mathbf{W})$, i.e.

$$\mathbf{W}_{min} = \arg\max_{\mathbf{W}} \min_k F_k(\mathbf{W})\ , \tag{31}$$

which can be solved using e.g. a sequential quadratic programming method. Obviously, the larger the values K_a, K_ψ and K_δ, the denser the grid of feasible microphone characteristics, and the higher the computational complexity for solving this minimax problem.

5. SIMULATIONS

We use a linear non-uniform microphone array consisting of $N = 3$ closely spaced microphones at nominal positions $[0\ \ 0.01\ \ 0.025]$ m, corresponding to a typical configuration for a multi-microphone BTE hearing aid. We assume that the microphone characteristics are independent of the angles ϕ and θ, i.e. $A_n(\omega,\phi,\theta) = A_n(\omega)$, and that the nominal microphone characteristic $A_n(\omega) = 1, n = 0 \dots N-1$.

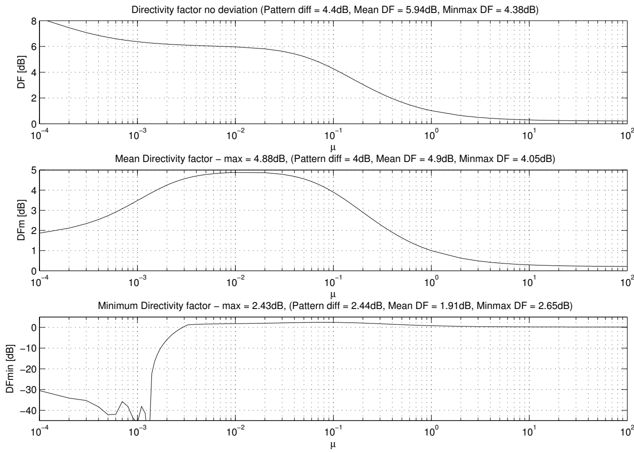

Fig. 2. Directivity factor, mean directivity factor and minimum directivity factor of $\mathbf{W}_{sd,\mu}$ as a function of μ

Design	DF	DF_m	DF_{min}
$\mathbf{W}_m$	5.94	**4.90**	1.91
$\mathbf{W}_{min}$	4.38	4.05	**2.65**
$\mathbf{W}_{sd}$ ($\mu = 0$)	**9.52**	1.33	-28.12
$\mathbf{W}_{ds}$ ($\mu = \infty$)	0.21	0.20	0.16
$\mathbf{W}_{sd,\mu}$ (max DF_m, $\mu = 0.01$)	5.97	4.88	1.82
$\mathbf{W}_{sd,\mu}$ (max DF_{min}, $\mu = 0.07$)	4.81	4.29	2.43
$\mathbf{W}_{tm,\lambda}$ (max DF_m, $\lambda \to 0$)	5.63	4.79	2.15
$\mathbf{W}_{tm,\lambda}$ (max DF_{min}, $\lambda = 1.4$)	4.77	4.26	2.43

Table 1. Directivity factor, mean and minimum directivity factor for different design procedures

Without loss of generality, we also assume that all microphone characteristics are described by the same pdf $f_{\mathcal{A}}(A)$. The direction of the speech source is $\theta_s = 0°$, the sampling frequency is $f_s = 16$ kHz and the design frequency is 1000 Hz. We will assume only gain deviations, mainly in order to limit the computational complexity for computing $\mathbf{W}_m$ and $\mathbf{W}_{min}$. We will use a uniform gain pdf with mean $\mu_{a,n} = 1$ and width $s_{a,n} = 0.3$. The grid spacing required for the design procedures in Section 4.2 is $\Delta a = 0.02$, such that the sum in (29) and $\mathbf{F}(\mathbf{W})$ in (30) consist of 27000 components.

Table 1 summarizes the directivity factor DF, the mean directivity factor DF_m, and the minimum directivity factor DF_{min} for different procedures. Obviously, the superdirective beamformer leads to the highest directivity factor when no microphone deviations occur ($DF = 9.52$ dB), the beamformer $\mathbf{W}_m$ leads to the highest mean directivity factor ($DF_m = 4.90$ dB), and the beamformer $\mathbf{W}_{min}$ leads to the highest minimum directivity factor ($DF_{min} = 2.65$ dB).

Figure 2 plots the directivity factors for the beamformer $\mathbf{W}_{sd,\mu}$ as a function of the factor μ. This factor provides a trade-off between directivity and robustness: the superdirective beamformer $\mathbf{W}_{sd}$ ($\mu = 0$) leads to the highest directivity factor when no deviations occur, but the mean directivity factor is only equal to $DF_m = 1.33$ dB, and the minimum directivity factor is equal to $DF_{min} = -28.12$ dB, illustrating the sensitivity of the superdirective beamformer to microphone deviations. On the other hand, the delay-and-sum beamformer $\mathbf{W}_{ds}$ ($\mu = \infty$) is very robust, but the directivity factor is very low. For $\mu = 0.01$, the mean directivity factor is maximized ($DF_m = 4.88$ dB), while for $\mu = 0.07$, the minimum directivity factor is maximized ($DF_{min} = 2.43$ dB). These values are quite close to the maximum attainable values.

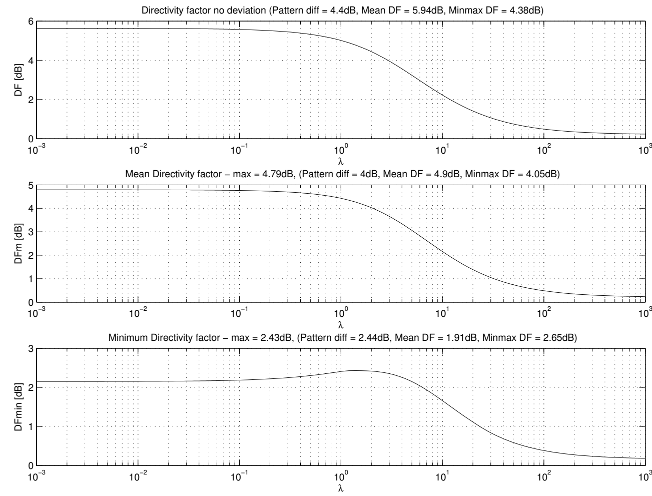

Fig. 3. Directivity factor, mean directivity factor and minimum directivity factor of $\mathbf{W}_{tm,\lambda}$ as a function of λ

Figure 3 plots the directivity factors for the beamformer $\mathbf{W}_{tm,\lambda}$ as a function of the factor λ. Using this figure, it is possible to determine the values of λ for which the mean and the minimum directivity factor are maximized. For λ approaching 0, the mean directivity factor is maximized ($DF_m = 4.79$ dB), while for $\lambda = 1.4$, the minimum directivity factor is maximized ($DF_{min} = 2.43$ dB). These values are again quite close to the maximum attainable values.

Except for the superdirective beamformer, which is very sensitive to deviations, and the delay-and-sum beamformer, whose performance is very low, all other beamformer designs may lead to a reasonable performance and robustness. Although it is hard to determine which design procedure is the optimal one, we can make the following conclusions:

1. If computational complexity is not important, the beamformers $\mathbf{W}_m$ and $\mathbf{W}_{min}$ are preferable, since they respectively optimize the mean or the worst-case directivity factor.
2. The performance of the beamformers $\mathbf{W}_{sd,\mu}$ and $\mathbf{W}_{tm,\lambda}$ is quite similar, where respectively the parameters μ and λ provide a trade-off between directivity factor, mean directivity factor and minimum directivity factor. Using Figures 2 and 3, it is possible to determine a suitable range for μ and λ.

6. REFERENCES

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

[2] J. Bitzer and K. U. Simmer, *Superdirective Microphone Arrays*, chapter 2 in "Microphone Arrays: Signal Processing Techniques and Applications", pp. 19–38, Springer-Verlag, May 2001.

[3] G. Elko, *Superdirectional Microphone Arrays*, chapter 10 in "Acoustic Signal Processing for Telecommunication", pp. 181–237, Kluwer Academic Publishers, Boston, 2000.

[4] S. Doclo and M. Moonen, "Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2511–2526, Oct. 2003.

[5] S. Doclo and M. Moonen, "Design of broadband beamformers robust against microphone position errors," in *Proc. IWAENC*, Kyoto, Japan, Sept. 2003, pp. 267–270.