# Superdirective Beamforming Robust Against Microphone Mismatch

Simon Doclo, *Member, IEEE*, and Marc Moonen, *Senior Member, IEEE*

***Abstract*—Fixed superdirective beamformers using small-sized microphone arrays are known to be highly sensitive to errors in the assumed microphone array characteristics (gain, phase, position). 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, minimizing the mean deviation from the desired superdirective directivity pattern, and maximizing the mean or the worst case directivity factor. When computational complexity is not an issue, maximizing the mean or the worst case directivity factor is the preferred design procedure. In addition, it is shown how to determine a suitable parameter range for the other design procedures such that both a high directivity and a high level of robustness are obtained.**

***Index Terms*—Microphone arrays, microphone mismatch, robust design, superdirective beamformer.**

## I. Introduction

In many speech communication applications, such as hands-free mobile telephony, hearing aids, and voice-controlled systems, the recorded microphone signals are corrupted with background noise and reverberation. Background noise and reverberation cause a signal degradation which can lead to total unintelligibility of the speech and which decreases the performance of speech recognition and coding systems. Therefore, efficient signal enhancement algorithms are required.

The objective of a fixed (data-independent) beamformer is to obtain spatial focusing on the speech source, thereby reducing background noise and reverberation not coming from the same direction as the speech source. For the design of fixed beamformers, the direction of the speech source and the complete microphone configuration generally need to be known. Different types of fixed beamformers are available, e.g., delay-and-sum beamformers, superdirective beamformers [1]–[5], differential microphone arrays [6], and frequency-invariant beamformers [7], [8]. Fixed beamformers are frequently applied in, e.g., hearing aids [9]–[11].

A superdirective beamformer maximizes the directivity factor, i.e., the microphone array gain for a diffuse noise field. It is well known that superdirective beamformers are sensitive to uncorrelated noise, especially at low frequencies and for small-sized microphone arrays [1]–[3]. 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 [12].

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]–[3]. In addition, we discuss three design procedures that optimize a mean performance criterion, i.e., the weighted sum of the mean noise and distortion energy, the mean deviation from the desired superdirective directivity pattern, and the mean (or the worst case) directivity factor [13]. These design procedures obviously require knowledge of the microphone gain, phase, and position probability density functions and are related to [14], [15], 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 worst case directivity factor is the preferred design procedure. In addition, it is shown how to determine a suitable parameter range for the other design procedures such that both a high directivity and a high level of robustness are obtained.

The paper is organized as follows. Section II describes the used microphone array configuration and defines the spatial directivity pattern, the directivity factor, and the white noise gain. In Section III, the design of superdirective beamformers is discussed when the microphone characteristics are exactly known, and the use of a white noise gain for limiting the amplification of uncorrelated noise components is discussed. Section IV presents the design procedures for improving the robustness of superdirective beamformers against unknown microphone mismatch, by optimizing a mean performance criterion. In Section V, simulation results are presented for a small-sized microphone array.

## II. Configuration and Notation

### A. Microphone Array and Signals

Consider the linear microphone array depicted in Fig. 1, consisting of $N$ microphones and with $d_n$ the distance between the $n$th microphone and the reference point, arbitrarily chosen here

Manuscript received November 10, 2005; revised April 26, 2006. This work performed at the ESAT laboratory of the Katholieke Universiteit Leuven, and was supported in part by the IWT Project 020540 (Innovative Speech Processing Algorithms for Improved Performance of Cochlear Implants), in part by the IWT Project 040803 (Sound Management System for Public Address systems), in part by the FWO Research Project G.0504.04 (Design and analysis of signal processing procedures for objective audiometry in newborns), in part by the FWO Research Project G.0334.06 (Virtual acoustics for the design and evaluation of auditory prostheses), in part by the Concerted Research Action GOA-AMBIORICS (Algorithms for medical and biological research, integration, computation and software), in part by the K.U.Leuven Research Council CoE EF/05/006 (Optimization in Engineering), and in part by the Interuniversity Attraction Pole IUAP P5-22 (Dynamical Systems and Control: Computation, Identification and Modelling), initiated by the Belgian Federal Science Policy Office. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shoji Makino.

The authors are with the Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, B-3001 Leuven, Belgium (e-mail: simon.doclo@esat.kuleuven.be; marc.moonen@esat.kuleuven.be).

Digital Object Identifier 10.1109/TASL.2006.881676

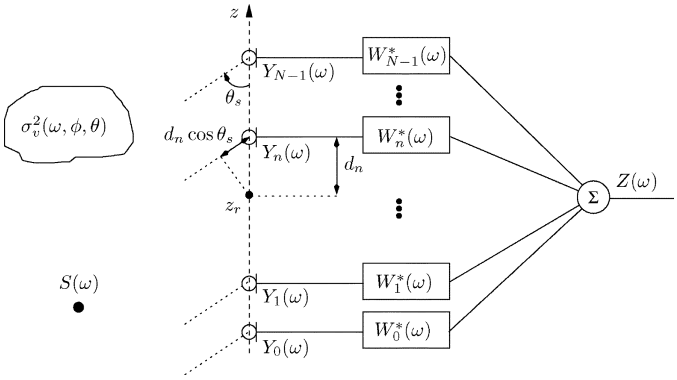



Fig. 1. Linear microphone array configuration.

as the center of the microphone array. Although we will assume a linear microphone array in this paper, all results can be readily extended for an arbitrary three-dimensional microphone configuration. All expressions are derived in the frequency-domain, using the normalized frequency $\omega$. We assume that a noise field with spectral and spatial characteristics $\sigma_v^2(\omega,\phi,\theta)$ is present, where $\phi$ and $\theta$ represent the azimuthal and the elevation angle ($0 \leq \phi < 2\pi$, $0 \leq \theta \leq \pi$), and that a speech source $S(\omega)$ is located at an angle $(\phi_s,\theta_s)$ in the far-field of the microphone array.[1]

The microphone characteristics of the $n$th microphone are described by

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

where both the gain $a_n(\omega,\phi,\theta)$ and the phase $\psi_n(\omega,\phi,\theta)$ can be frequency- and angle-dependent. In Sections II and III, we assume that the microphone characteristics are perfectly known (e.g., using a measurement or a calibration procedure), whereas in Section IV, the microphone characteristics are not assumed to be perfectly known. The $n$th microphone signal $Y_n(\omega)$ is equal to

$$Y_n(\omega) = g_n(\omega,\phi_s,\theta_s)S_r(\omega) + V_n(\omega) \tag{2}$$

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

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

where the delay $\tau_n(\phi,\theta)$ in number of samples is equal to $(d_n\cos\theta/c)f_s$, with $c$ the speed of sound propagation $c = 340m/s$ and $f_s$ the sampling frequency.[2] The stacked vector of microphone signals

$$\mathbf{Y}(\omega) = [Y_0(\omega) \quad Y_1(\omega) \quad \ldots, \quad Y_{N-1}(\omega)]^T \tag{4}$$

can be written as

$$\mathbf{Y}(\omega) = \mathbf{g}_s(\omega)S_r(\omega) + \mathbf{V}(\omega) \tag{5}$$

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

$$\mathbf{g}(\omega,\phi,\theta) = [g_0(\omega,\phi,\theta) \;\; g_1(\omega,\phi,\theta) \;\; \ldots, \;\; g_{N-1}(\omega,\phi,\theta)]^T \tag{6}$$

and $\mathbf{V}(\omega)$ is defined similarly as $\mathbf{Y}(\omega)$ in (4). The output signal $Z(\omega)$ is equal to

$$Z(\omega) = \sum_{n=0}^{N-1} W_n^*(\omega)Y_n(\omega) = \mathbf{W}^H(\omega)\mathbf{Y}(\omega) \tag{7}$$

$$= \mathbf{W}^H(\omega)\mathbf{g}_s(\omega)S_r(\omega) + \mathbf{W}^H(\omega)\mathbf{V}(\omega) \tag{8}$$

with $\mathbf{W}(\omega) = [W_0(\omega)\; W_1(\omega)\; \ldots,\; W_{N-1}(\omega)]^T$ the filter coefficients of the beamformer.

### B. Spatial Directivity Pattern and Array Gain

The *spatial directivity pattern* $H(\omega,\phi,\theta)$ is defined as the transfer function between (the reference signal corresponding to) a source at an angle $(\phi,\theta)$ and the output signal of the microphone array, i.e.,

$$H(\omega,\phi,\theta) = \mathbf{W}^H(\omega)\mathbf{g}(\omega,\phi,\theta). \tag{9}$$

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, i.e.,

$$G(\omega) = \frac{\mathrm{SNR}_{out}(\omega)}{\mathrm{SNR}_{in}(\omega)}. \tag{10}$$

The input SNR is equal to

$$\mathrm{SNR}_{in}(\omega) = \frac{\Phi_s(\omega)}{\Phi_v(\omega)} \tag{11}$$

with $\Phi_s(\omega) = \mathcal{E}\{|S_r(\omega)|^2\}$ the speech energy of the reference signal and $\Phi_v(\omega) = \mathcal{E}\{|V_r(\omega)|^2\}$ the noise energy of the reference signal. Using (8), the SNR of the output signal $Z(\omega)$ is equal to

$$\mathrm{SNR}_{out}(\omega) = \frac{\Phi_s(\omega)\left|\mathbf{W}^H(\omega)\mathbf{g}_s(\omega)\right|^2}{\mathbf{W}^H(\omega)\mathbf{\Phi}_{VV}(\omega)\mathbf{W}(\omega)} \tag{12}$$

with the noise correlation matrix $\mathbf{\Phi}_{VV}(\omega)$ equal to

$$\mathbf{\Phi}_{VV}(\omega) = \mathcal{E}\left\{\mathbf{V}(\omega)\mathbf{V}^H(\omega)\right\} \tag{13}$$

$$= \begin{bmatrix} \Phi_{V_0V_0}(\omega) & \ldots, & \Phi_{V_0V_{N-1}}(\omega) \\ \vdots & & \vdots \\ \Phi_{V_{N-1}V_0}(\omega) & \ldots, & \Phi_{V_{N-1}V_{N-1}}(\omega) \end{bmatrix}. \tag{14}$$

Hence, using (11) and (12), the array gain in (10) is equal to

$$G(\omega) = \frac{\left|\mathbf{W}^H(\omega)\mathbf{g}_s(\omega)\right|^2}{\mathbf{W}^H(\omega)\widetilde{\mathbf{\Phi}}_{VV}(\omega)\mathbf{W}(\omega)} \tag{15}$$

with the normalized noise correlation matrix $\widetilde{\mathbf{\Phi}}_{VV}(\omega) = \mathbf{\Phi}_{VV}(\omega)/\Phi_v(\omega)$. Note that for a homogeneous noise field, i.e., $\Phi_{V_nV_n}(\omega) = \Phi_v(\omega)$, $n = 0,\ldots,N-1$, the normalized noise correlation matrix is equal to the noise coherence matrix

[1]Although we assume that the speech source is located in the far-field of the microphone array, all results can be easily extended for a speech source in the near-field of the microphone array [4].

[2]For a linear microphone array, $\tau_n(\phi,\theta)$ is independent of the angle $\phi$. For an arbitrary three-dimensional microphone configuration, the delay is equal to $\tau_n(\phi,\theta) = (d_{xn}\cos\phi\sin\theta + d_{yn}\sin\phi\sin\theta + d_{zn}\cos\theta)f_s/c$, with $d_{xn}$, $d_{yn}$, and $d_{zn}$ the distance to the reference point in the $x$, $y$, and $z$-direction.

$\mathbf{\Gamma}_{VV}(\omega)$, with $\Gamma_{V_nV_p}(\omega) = \Phi_{V_nV_p}(\omega)/\sqrt{\Phi_{V_nV_n}(\omega)\Phi_{V_pV_p}(\omega)}$. By spatially integrating the noise field $\sigma_v^2(\omega,\phi,\theta)$ over all angles, the $(n,p)$th element of $\widetilde{\mathbf{\Phi}}_{VV}(\omega)$ can be computed as

$$\widetilde{\Phi}_{V_nV_p}(\omega) = \frac{\int_0^{2\pi}\int_0^{\pi} g_n(\omega,\phi,\theta)g_p^*(\omega,\phi,\theta)\sigma_v^2(\omega,\phi,\theta)\sin\theta d\theta d\phi}{\int_0^{2\pi}\int_0^{\pi}\sigma_v^2(\omega,\phi,\theta)\sin\theta d\theta d\phi}. \quad (16)$$

Two common quantities to describe the performance of a microphone array are the directivity factor and the white noise gain. The *directivity factor* (DF) is defined as the ability of the microphone array to suppress spherically isotropic noise (diffuse noise), i.e., independent noise sources uniformly distributed in all directions, for which $\sigma_v^2(\omega,\phi,\theta) = \sigma_v^2(\omega)$. Hence, using (16), the directivity factor is equal to

$$DF(\omega) = \frac{\left|\mathbf{W}^H(\omega)\mathbf{g}_s(\omega)\right|^2}{\mathbf{W}^H(\omega)\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}(\omega)\mathbf{W}(\omega)} \quad (17)$$

with

$$\widetilde{\Phi}_{V_nV_p}^{\text{diff}}(\omega) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} A_n(\omega,\phi,\theta)A_p^*(\omega,\phi,\theta) \times e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}\sin\theta d\theta d\phi. \quad (18)$$

When the microphone characteristics are independent of the angles $\phi$ and $\theta$, i.e., $A_n(\omega,\phi,\theta) = A_n(\omega), n = 0,\ldots,N-1$, this expression can be simplified as

$$\begin{aligned}\widetilde{\Phi}_{V_nV_p}^{\text{diff}}(\omega) &= A_n(\omega)A_p^*(\omega)\frac{1}{2}\int_0^{\pi} e^{-j\omega\frac{(d_n-d_p)\cos\theta}{c}f_s}\sin\theta d\theta \\ &= A_n(\omega)A_p^*(\omega)\text{sinc}\left(\frac{\omega(d_n-d_p)f_s}{c}\right) \quad (19)\end{aligned}$$

with $\text{sinc}(x) = \sin(x)/x$.

The *white noise gain* (WNG) is defined as the ability of the microphone array to suppress spatially uncorrelated noise, e.g., sensor noise of the microphones. Hence, the noise correlation matrix $\widetilde{\mathbf{\Phi}}_{VV}^{uncorr}(\omega) = \mathbf{I}_N$, with $\mathbf{I}_N$ the $N \times N$-dimensional identity matrix, such that the WNG is equal to

$$\text{WNG}(\omega) = \frac{\left|\mathbf{W}^H(\omega)\mathbf{g}_s(\omega)\right|^2}{\mathbf{W}^H(\omega)\mathbf{W}(\omega)}. \quad (20)$$

The WNG is a commonly used measure for robustness.

## III. Superdirective Beamforming

For the sake of conciseness, we will omit the variable $\omega$ where possible in the remainder of the paper.

### A. Optimization Criteria

The superdirective beamformer $\mathbf{W}_{sd}$ maximizes the array gain for diffuse noise, i.e., the directivity factor defined in (17)

$$\max_{\mathbf{W}} \frac{|\mathbf{W}^H\mathbf{g}_s|^2}{\mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\mathbf{W}}. \quad (21)$$

Obviously, the solution is given by the generalized eigenvector corresponding to the largest generalized eigenvalue of $\mathbf{g}_s\mathbf{g}_s^H$ and $\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}$, i.e., $\mathbf{W}_{sd} = \alpha(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}})^{-1}\mathbf{g}_s$, where $\alpha$ is an arbitrary constant. By imposing the constraint $\mathbf{W}^H\mathbf{g}_s = 1$, i.e., a unity response in the direction of the speech source, the superdirective beamformer $\mathbf{W}_{sd}$ can be computed as

$$\mathbf{W}_{sd} = \frac{\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}{\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}. \quad (22)$$

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

$$\min_{\mathbf{W}} \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\mathbf{W}, \quad \text{subject to } \mathbf{W}^H\mathbf{g}_s = 1. \quad (23)$$

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

$$J_t(\mathbf{W},\lambda) = J_v(\mathbf{W}) + \lambda J_d(\mathbf{W}) \quad (24)$$

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

$$J_v(\mathbf{W}) = \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\mathbf{W}, \quad J_d(\mathbf{W}) = |\mathbf{W}^H\mathbf{g}_s - 1|^2. \quad (25)$$

The cost function $J_t(\mathbf{W},\lambda)$ is minimized by setting the derivative $\partial J_t(\mathbf{W},\lambda)/\partial\mathbf{W}$ equal to zero, such that the filter $\mathbf{W}_t(\lambda)$ minimizing $J_t(\mathbf{W},\lambda)$ is equal to

$$\mathbf{W}_t(\lambda) = \left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}} + \lambda\mathbf{g}_s\mathbf{g}_s^H\right)^{-1}\lambda\mathbf{g}_s \quad (26)$$

$$= \frac{\lambda\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}{1+\lambda\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}. \quad (27)$$

For this filter, the noise energy and the distortion energy are equal to

$$J_v(\mathbf{W}_t(\lambda)) = \frac{\lambda^2\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}{\left[1+\lambda\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s\right]^2} \quad (28)$$

$$J_d(\mathbf{W}_t(\lambda)) = \frac{1}{\left[1+\lambda\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s\right]^2} \quad (29)$$

such that the larger $\lambda$, the larger the noise energy and the smaller the distortion energy. Note that the superdirective beamformer is equal to $\mathbf{W}_t(\lambda)$ when $\lambda$ approaches $\infty$, i.e., $\mathbf{W}_{sd} = \mathbf{W}_t(\infty)$. For the superdirective beamformer the distortion energy is equal to zero, and the noise energy is equal to

$$J_v(\mathbf{W}_{sd}) = \frac{1}{\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\right)^{-1}\mathbf{g}_s}. \quad (30)$$

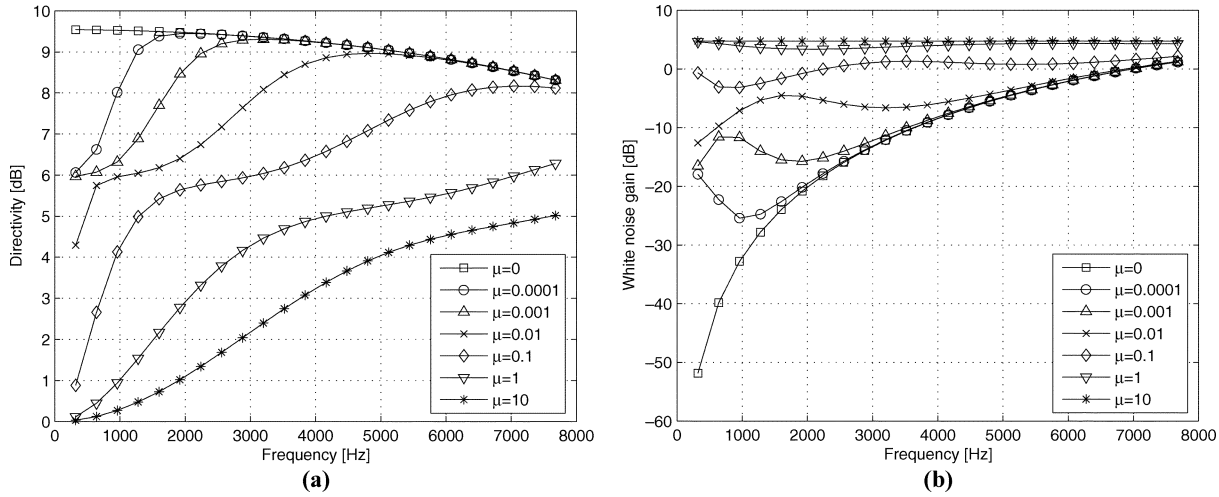



Fig. 2. (a) Directivity factor and (b) WNG of superdirective beamformer for different values of $\mu$ (microphone positions [0 0.01 0.025] m, $f_s = 16$ kHz, $\theta_s = 0°$).

### B. White Noise Gain (WNG) Constraint

It is well known that superdirective beamformers are sensitive to uncorrelated noise (i.e., the WNG is small), especially at low frequencies, such that uncorrelated noise components may even be amplified [1]–[3]. In addition, superdirective beamformers are sensitive to deviations from the assumed microphone characteristics (gain, phase, and position). 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 WNG constraint [1]–[3]. i.e.,

$$\frac{|\mathbf{W}^H\mathbf{g}_s|^2}{\mathbf{W}^H\mathbf{W}} \geq \frac{1}{\beta} \tag{31}$$

where $1/\beta$ represents the minimum desired WNG. The value of $\beta$ needs to be chosen in function of the amount of sensor noise present and/or the expected amount of microphone mismatch. Since for superdirective beamformers $\mathbf{W}^H\mathbf{g}_s = 1$, this inequality constraint (31) is equivalent to limiting the norm of the filter, i.e., $\mathbf{W}^H\mathbf{W} \leq \beta$. Hence, using (23), the optimization problem becomes

$$\min_{\mathbf{W}} \mathbf{W}^H\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}\mathbf{W}, \quad \text{subject to } \mathbf{W}^H\mathbf{g}_s = 1, \mathbf{W}^H\mathbf{W} \leq \beta. \tag{32}$$

Using the method of Lagrange multipliers, the solution of this optimization problem has the form

$$\mathbf{W}_{sd,\mu} = \frac{\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}} + \mu\mathbf{I}_N\right)^{-1}\mathbf{g}_s}{\mathbf{g}_s^H\left(\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}} + \mu\mathbf{I}_N\right)^{-1}\mathbf{g}_s} \tag{33}$$

which is equivalent to diagonal loading of the normalized noise correlation matrix $\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}$. The Lagrange multiplier $\mu$ needs to be determined such that the inequality constraint $\mathbf{W}_{sd,\mu}^H\mathbf{W}_{sd,\mu} \leq \beta$ is satisfied, e.g., using a multistep iterative procedure [2] or using a convex optimization approach via second-order cone programming [5]. The larger $\mu$, the larger the robustness of the beamformer, but the smaller its directivity factor. For $\mu = \infty$, the superdirective beamformer becomes equal to the delay-and-sum beamformer, i.e.,

$$\mathbf{W}_{sd,\infty} = \mathbf{W}_{ds} = \frac{\mathbf{g}_s}{\mathbf{g}_s^H\mathbf{g}_s} = \frac{\mathbf{g}_s}{N} \tag{34}$$

which is known to maximize the WNG in (20) and, hence, exhibits the largest robustness against uncorrelated noise.

*Example 1:* For a small-sized microphone array with $N = 3$ omnidirectional microphones at positions [0 0.01 0.025] m, sampling frequency $f_s = 16$ kHz, and direction of the speech source $\theta_s = 0°$, Fig. 2 depicts the directivity factor and the WNG for a superdirective beamformer designed using (33) for different values of $\mu$. When $\mu = 0$, the directivity factor is high (the maximum value is equal to $10\log_{10}(N^2) = 9.54$ dB), but the WNG is very poor, especially for low frequencies. When $\mu$ increases, the WNG improves, but the directivity factor decreases. When $\mu = 10$, the superdirective beamformer is practically equal to the delay-and-sum beamformer, i.e., the WNG is equal to $10\log_{10}(N) = 4.77$ dB for all frequencies, but the directivity factor is very poor, especially for low frequencies. This figure clearly illustrates the tradeoff between the directivity factor and the WNG.

*Example 2:* For the same microphone configuration, Fig. 3 depicts the effect of a deviation from the assumed microphone characteristics for a superdirective beamformer designed using (33) for different values of $\mu$. The gain mismatch is [0 2 0] dB, the phase mismatch is [ –5 10 5]°, and the microphone position mismatch is [0.001 −0.001 0.001] m. Fig. 3(a) depicts the decrease of the directivity factor, Fig. 3(b) depicts the increase of the noise energy $J_v(\mathbf{W})$, and Fig. 3(c) depicts the distortion energy $J_d(\mathbf{W})$ when microphone mismatch is present. When $\mu = 0$, the superdirective beamformer is very sensitive to microphone mismatch, resulting in a large decrease of the directivity factor and a large increase of the noise energy and the distortion energy, especially for low frequencies. When $\mu$ increases, these effects become less pronounced. Note, however, that in comparison with the situation without microphone mismatch, the directivity factor always decreases and the noise energy and the distortion energy always increase (even for the delay-and-sum beamformer). These figures clearly illustrate that a larger value

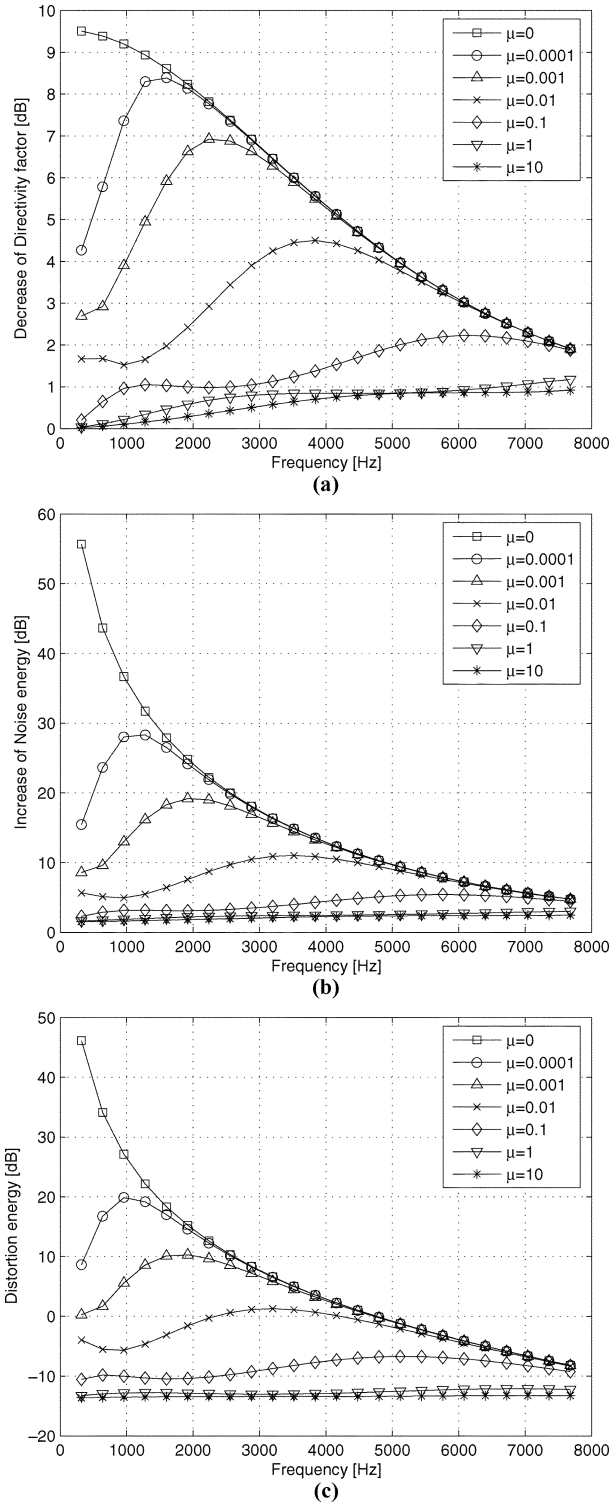



Fig. 3. (a) Decrease of directivity factor. (b) Increase of noise energy. (c) Distortion energy for different values of $\mu$ when microphone mismatch (gain, phase, and position) is present (nominal microphone positions [0 0.01 0.025] m, $f_s$ = 16 kHz, $\theta_s = 0°$).

of $\mu$ increases the robustness against microphone mismatch. Using these figures, it is possible to determine a (frequency-dependent) value for $\mu$ such that specific requirements regarding minimum directivity factor and maximum noise and distortion energy are satisfied *for this specific microphone mismatch*.

## IV. Robust Superdirective Beamforming Using Probability Density Function of Microphone Characteristics

Using the design procedures discussed in Section III, it is possible to design a superdirective beamformer when the microphone characteristics $A_n(\omega,\phi,\theta)$ and the microphone positions $d_n$ are exactly known. However, as has been shown in Section III-B, superdirective beamformers are highly sensitive to deviations from the assumed microphone characteristics. Since in practice it is difficult to manufacture microphones with exact predefined characteristics, it is practically impossible to exactly know the microphone characteristics without a measurement or a calibration procedure. Obviously, the cost of such a calibration procedure for every individual microphone array is objectionable. Moreover, after calibration the microphone characteristics can still drift over time.

In Section III-B, it has been shown that the robustness of superdirective beamformers against microphone mismatch 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 $\beta$ or $\mu$ that guarantees robustness for a range of microphone mismatches.

In this section, we present design procedures for improving the robustness of superdirective beamformers against unknown microphone mismatch by optimizing a *mean performance criterion*, i.e., a weighted sum over all feasible microphone characteristics using the probabilities of the microphone characteristics as weights. This procedure obviously requires knowledge of the microphone gain, phase, and position probability density functions (pdf) and is related to [14], [15], where the design of robust beamformers with an arbitrary spatial directivity pattern has been discussed. The three following design procedures will be discussed:

1) minimize the weighted sum of the mean noise and distortion energy (cf. Section IV-A);
2) minimize the mean deviation from the desired superdirective directivity pattern (cf. Section IV-B);
3) maximize the mean or the worst case directivity factor (cf. Section IV-C).

In order to be able to describe microphone position errors, we will incorporate these errors directly into the microphone characteristics defined in (1), i.e., we redefine $A_n(\omega,\phi,\theta)$ as

$$A_n(\omega,\phi,\theta) = a_n(\omega,\phi,\theta)e^{-j\psi_n(\omega,\phi,\theta)}e^{-j\omega\frac{\delta_n\cos\theta}{c}f_s} \qquad (35)$$

where $\delta_n$ represents the linear position error for the $n$th microphone. This position error in fact corresponds to a frequency- and angle-dependent phase error $\omega(\delta_n\cos\theta/c)f_s$. The probability density function $f_{\mathcal{A}}(A)$ describes the joint pdf of the stochastic variables $a$ (gain), $\psi$ (phase) and $\delta$ (position error). We assume that $a$, $\psi$ and $\delta$ are independent stochastic variables, such that the joint pdf is separable, i.e.,

$$f_{\mathcal{A}}(A) = f_\alpha(a)f_\Psi(\psi)f_\Delta(\delta) \qquad (36)$$

with $f_\alpha(a)$ the gain pdf, $f_\Psi(\psi)$ the phase pdf and $f_\Delta(\delta)$ the position error pdf. These pdfs are normalized such that the area under the pdfs is equal to 1.

### A. Mean Noise and Distortion Energy

Similar to (24), the weighted sum of the mean noise energy $J_{vm}(\mathbf{W})$ and the mean distortion energy $J_{dm}(\mathbf{W})$ is equal to

$$J_{tm}(\mathbf{W},\lambda) = J_{vm}(\mathbf{W}) + \lambda J_{dm}(\mathbf{W}) \tag{39}$$

with

$$J_{vm}(\mathbf{W}) = \int_{A_0}\dots\int_{A_{N-1}} \mathbf{W}^H \widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}(\mathbf{A})\mathbf{W} \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1} \tag{40}$$

$$J_{dm}(\mathbf{W}) = \int_{A_0}\dots\int_{A_{N-1}} \left|\mathbf{W}^H\mathbf{g}_s(\mathbf{A}) - 1\right|^2 \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1} \tag{41}$$

where $\widetilde{\mathbf{\Phi}}_{VV}^{\text{diff}}(\mathbf{A})$ denotes the normalized diffuse noise correlation matrix in (18) for the specific microphone array characteristic $\mathbf{A} = \{A_0,\dots,A_{N-1}\}$, and $\mathbf{g}_s(\mathbf{A})$ denotes the steering vector in (6) and (3) for the angle $(\phi_s,\theta_s)$ and the microphone array characteristic $\mathbf{A}$.

The mean noise energy $J_{vm}(\mathbf{W})$ can be written as

$$J_{vm}(\mathbf{W}) = \mathbf{W}^H\widetilde{\mathbf{\Phi}}_m^{\text{diff}}\mathbf{W} \tag{42}$$

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

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

Using (18), the $(n,p)$th element of $\widetilde{\mathbf{\Phi}}_m^{\text{diff}}$ is equal to

$$\begin{aligned}\widetilde{\Phi}_{m,np}^{\text{diff}} &= \int_{A_n}\int_{A_p}\left[\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} A_n(\omega,\phi,\theta)A_p^*(\omega,\phi,\theta) \times e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}\sin\theta d\theta d\phi\right] \times f_{\mathcal{A}}(A_n)f_{\mathcal{A}}(A_p)dA_n dA_p \\ &= \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}\sigma_{A,np}^2(\omega,\phi,\theta) \times e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}\sin\theta d\theta d\phi\end{aligned} \tag{37}$$

with

$$\sigma_{A,np}^2(\omega,\phi,\theta) = \int_{A_n}\int_{A_p} A_n(\omega,\phi,\theta)A_p^*(\omega,\phi,\theta)f_{\mathcal{A}}(A_n)f_{\mathcal{A}}(A_p)dA_n dA_p.$$

For different pdfs (uniform, log-uniform, normal, log-normal), the calculation of $\sigma_{A,np}^2(\omega,\phi,\theta)$ is discussed in Appendix I, and the calculation of $\widetilde{\Phi}_{m,np}^{\text{diff}}$ is discussed in Appendix II.

The mean distortion energy $J_{dm}(\mathbf{W})$ can be written as

$$J_{dm}(\mathbf{W}) = \mathbf{W}^H\mathbf{Q}_{sm}\mathbf{W} - \mathbf{W}^H\mathbf{q}_{sm} - \mathbf{q}_{sm}^H\mathbf{W} + 1 \tag{43}$$

with

$$\mathbf{Q}_{sm} = \int_{A_0}\dots\int_{A_{N-1}} \mathbf{g}_s(\mathbf{A})\mathbf{g}_s^H(\mathbf{A}) \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1} \tag{44}$$

$$\mathbf{q}_{sm} = \int_{A_0}\dots\int_{A_{N-1}} \mathbf{g}_s(\mathbf{A}) \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1}. \tag{45}$$

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

$$Q_{sm,np} = \int_{A_n}\int_{A_p} A_n(\omega,\phi_s,\theta_s)A_p^*(\omega,\phi_s,\theta_s) \times e^{-j\omega(\tau_n(\phi_s,\theta_s)-\tau_p(\phi_s,\theta_s))} \times f_{\mathcal{A}}(A_n)f_{\mathcal{A}}(A_p)dA_n dA_p \tag{38}$$

which can be written as

$$Q_{sm,np} = \sigma_{A,np}^2(\omega,\phi_s,\theta_s)e^{-j\omega\frac{(d_n-d_p)\cos\theta_s}{c}f_s}. \tag{46}$$

The $n$th element of $\mathbf{q}_{sm}$ is equal to

$$\begin{aligned}q_{sm,n} &= \int_{A_n} A_n(\omega,\phi_s,\theta_s)e^{-j\omega\tau_n(\phi_s,\theta_s)}f_{\mathcal{A}}(A_n)dA_n \\ &= \mu_{A,n}(\omega,\phi_s,\theta_s)e^{-j\omega\frac{d_n\cos\theta_s}{c}f_s}.\end{aligned} \tag{47}$$

For different pdfs, the calculation of $\mu_{A,n}(\omega,\phi,\theta)$ is discussed in Appendix I.

Similar to (27), the filter $\mathbf{W}_{tm,\lambda}$ minimizing $J_{tm}(\mathbf{W},\lambda)$ is equal to

$$\mathbf{W}_{tm,\lambda} = (\widetilde{\mathbf{\Phi}}_m + \lambda\mathbf{Q}_{sm})^{-1}\lambda\mathbf{q}_{sm}. \tag{48}$$

The larger $\lambda$, the larger the mean noise energy and the smaller the mean distortion energy.

### B. Mean Deviation From the Desired Superdirective Directivity Pattern

In [14] and [15], design procedures have been discussed for designing beamformers with an arbitrary spatial directivity pattern that are robust against microphone mismatch. Consider the least-squares error between the spatial directivity pattern $H(\omega,\phi,\theta)$ and the desired spatial directivity pattern $D(\omega,\phi,\theta)$. The weighted least-squares cost function is then defined as

$$J_{LS}(\mathbf{W}) = \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\left|H(\omega,\phi,\theta) - D(\omega,\phi,\theta)\right|^2 d\theta d\phi \tag{49}$$

where $F(\omega,\phi,\theta)$ is a positive real weighting function, assigning more or less importance to certain angles. Here, we will define the desired spatial directivity pattern $D(\omega,\phi,\theta)$ to be equal to the spatial directivity pattern of the superdirective beamformer when no microphone mismatch occurs.

The weighted least-squares cost function can be written as the quadratic function

$$J_{LS}(\mathbf{W}) = \mathbf{W}^H\mathbf{Q}\mathbf{W} - \mathbf{W}^H\mathbf{q} - \mathbf{q}^H\mathbf{W} + d \tag{50}$$

with

$$\mathbf{Q} = \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\mathbf{g}(\omega,\phi,\theta)\mathbf{g}^H(\omega,\phi,\theta)d\theta d\phi \quad (51)$$

$$\mathbf{q} = \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\mathbf{g}(\omega,\phi,\theta)D^*(\omega,\phi,\theta)d\theta d\phi \quad (52)$$

$$d = \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\left|D(\omega,\phi,\theta)\right|^2 d\theta d\phi. \quad (53)$$

Robustness against microphone mismatch can be achieved by minimizing the mean weighted least-squares cost function over all feasible microphone characteristics, i.e.,

$$J_{LS,m}(\mathbf{W}) = \int_{A_0}\dots\int_{A_{N-1}} J_{LS}(\mathbf{W},\mathbf{A}) \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1} \quad (56)$$

where $J_{LS}(\mathbf{W},\mathbf{A})$ denotes the weighted least-squares cost function in (49) for the microphone array characteristic $\mathbf{A}$. This mean cost function can be written as

$$J_{LS,m}(\mathbf{W}) = \mathbf{W}^H\mathbf{Q}_m\mathbf{W} - \mathbf{W}^H\mathbf{q}_m - \mathbf{q}_m^H\mathbf{W} + d \quad (57)$$

where the $(n,p)$th element of $\mathbf{Q}_m$ is equal to

$$\begin{aligned} Q_{m,np} &= \int_{A_n}\int_{A_p}\left[\int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)A_n(\omega,\phi,\theta)A_p^*(\omega,\phi,\theta) \times e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}d\theta d\phi\right] \times f_{\mathcal{A}}(A_n)f_{\mathcal{A}}(A_p)dA_n dA_p \\ &= \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\sigma^2_{A,np}(\omega,\phi,\theta) \times e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}d\theta d\phi \end{aligned} \quad (54)$$

and the $n$th element of $\mathbf{q}_m$ is equal to

$$\begin{aligned} q_{m,n} &= \int_{A_n}\left[\int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)A_n(\omega,\phi,\theta) \times e^{-j\omega\tau_n(\phi,\theta)}D^*(\omega,\phi,\theta)d\theta d\phi\right] \times f_{\mathcal{A}}(A_n)dA_n \\ &= \int_0^{2\pi}\int_0^{\pi} F(\omega,\phi,\theta)\mu_{A,n}(\omega,\phi,\theta)e^{-j\omega\tau_n(\phi,\theta)} \times D^*(\omega,\phi,\theta)d\theta d\phi. \end{aligned} \quad (55)$$

For different pdfs, the calculation of $\sigma^2_{A,np}(\omega,\phi,\theta)$ and $\mu_{A,n}(\omega,\phi,\theta)$ is discussed in Appendix I. In general, the integrals in (54) and (55) need to be computed numerically.

The filter $\mathbf{W}_{LS,m}$ minimizing the mean cost function in (57) is equal to

$$\mathbf{W}_{LS,m} = \mathbf{Q}_m^{-1}\mathbf{q}_m. \quad (58)$$

### C. Mean and Worst-Case Directivity Factor

The *mean directivity factor* is defined as

$$DF_m(\mathbf{W}) = \int_{A_0}\dots\int_{A_{N-1}} DF(\mathbf{W},\mathbf{A}) \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})dA_0\dots dA_{N-1} \quad (59)$$

where $DF(\mathbf{W},\mathbf{A})$ denotes the directivity factor in (17) for the microphone array characteristic $\mathbf{A}$, i.e.,

$$DF(\mathbf{W},\mathbf{A}) = \frac{\left|\mathbf{W}^H\mathbf{g}_s(\mathbf{A})\right|^2}{\mathbf{W}^H\widetilde{\mathbf{\Phi}}^{\text{diff}}_{VV}(\mathbf{A})\mathbf{W}}. \quad (60)$$

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. Hence, we will approximate the integrals in (59) by a discrete sum, i.e.,

$$DF_m(\mathbf{W}) = \sum_{A_0}\dots\sum_{A_{N-1}} DF(\mathbf{W},\mathbf{A}) \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\Delta A_0\dots\Delta A_{N-1} \quad (61)$$

where $\Delta A_n$ denotes the grid spacing for the pdf describing the $n$th microphone characteristic. Obviously, the smaller the grid spacing, the more expensive the computation of this sum. For example, when only microphone gain deviations are considered and all microphone characteristics are assumed to be described by the same uniform pdf with minimum value $a_{\min}$ and maximum value $a_{\max}$ and $\Delta a$ is the used grid spacing, the sum in (61) consists of $((a_{\max}-a_{\min})/\Delta a)^N$ components.

Since no closed-form expression is available for the filter $\mathbf{W}_m$ maximizing (61), an iterative optimization technique will be used. The numerical robustness and the convergence speed of many unconstrained optimization techniques (e.g., quasi-Newton method [16]) can be improved by providing an analytical expression for the gradient, i.e.,

$$\frac{\partial DF_m(\mathbf{W},\mathbf{A})}{\partial\mathbf{W}} = \sum_{A_0}\dots\sum_{A_{N-1}} \frac{\partial DF(\mathbf{W},\mathbf{A})}{\partial\mathbf{W}} \times f_{\mathcal{A}}(A_0)\dots f_{\mathcal{A}}(A_{N-1})\Delta A_0\dots\Delta A_{N-1} \quad (62)$$

with $\partial DF(\mathbf{W},\mathbf{A})/\partial\mathbf{W}$ equal to

$$\frac{2}{\left(\mathbf{W}^H\widetilde{\mathbf{\Phi}}^{\text{diff}}_{VV}(\mathbf{A})\mathbf{W}\right)^2}\left[\left(\mathbf{W}^H\widetilde{\mathbf{\Phi}}^{\text{diff}}_{VV}(\mathbf{A})\mathbf{W}\right)\mathbf{g}_s(\mathbf{A})\mathbf{g}_s^H(\mathbf{A}) - \left|\mathbf{W}^H\mathbf{g}_s(\mathbf{A})\right|^2\widetilde{\mathbf{\Phi}}^{\text{diff}}_{VV}(\mathbf{A})\right]\mathbf{W}. \quad (63)$$

Although we cannot prove that the used optimization procedure converges to the global minimum, no problems with local minima have been observed in our simulations.

When maximizing the mean directivity factor, it is still possible that for some specific microphone deviation the directivity factor is quite small. 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, $K_\psi$ phase values and $K_\delta$ position error values), i.e., $\{a_{\min} = a_1, \ldots, a_{K_a} = a_{\max}\}$, $\{\psi_{\min} = \psi_1, \ldots, \psi_{K_\psi} = \psi_{\max}\}$, $\{\delta_{\min} = \delta_1, \ldots, \delta_{K_\delta} = \delta_{\max}\}$, as an approximation for the continuum of feasible microphone characteristics. We use this set to construct the $K_{\tot}$-dimensional vector $\mathbf{F}(\mathbf{W})$, with $K_{\tot} = (K_a K_\psi K_\delta)^N$, i.e.

$$\mathbf{F}(\mathbf{W}) = [DF_1(\mathbf{W}) \quad DF_2(\mathbf{W}) \quad \ldots \quad DF_{K_{\tot}}(\mathbf{W})]^T \tag{64}$$

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.,

$$\mathbf{W}_{\min} = \arg \max_{\mathbf{W}} \min_k DF_k(\mathbf{W}). \tag{65}$$

By considering the vector $-\mathbf{F}(\mathbf{W})$, this is equivalent to a minimax optimization problem that can be solved using a sequential quadratic programming method [16]. In order to improve the numerical robustness and the convergence speed, the gradient

$$\left[\frac{\partial DF_1(\mathbf{W})}{\partial \mathbf{W}} \quad \frac{\partial DF_2(\mathbf{W})}{\partial \mathbf{W}} \quad \ldots \quad \frac{\partial DF_{K_{\tot}}(\mathbf{W})}{\partial \mathbf{W}}\right] \tag{66}$$

which is an $N \times K_{\tot}$-dimensional matrix, can be supplied analytically. Obviously, the larger the values $K_a$, $K_\psi$, and $K_\delta$, the denser the grid of feasible microphone characteristics, and the higher the computational complexity for solving the minimax optimization problem.

Since both the mean directivity factor $DF_m(\mathbf{W})$ and the vector $\mathbf{F}(\mathbf{W})$ used in the minimax problem are scale-invariant, i.e.,

$$DF_m(\mathbf{W}) = DF_m(\alpha\mathbf{W}), \quad \mathbf{F}(\mathbf{W}) = \mathbf{F}(\alpha\mathbf{W}), \quad \forall \alpha \in \mathbb{C}$$

we can perform a normalization such that $\mathbf{W}_m^H \mathbf{g}_s = 1$ and $\mathbf{W}_{\min}^H \mathbf{g}_s = 1$, where $\mathbf{g}_s$ denotes the steering vector when no microphone deviation is present.

## V. Simulations

In this section, simulation results for a small-sized microphone array are presented. First, we describe the setup and the performance measures used. For the different beamformer design procedures we then compare the directivity factors, the mean noise and distortion energy, the spatial directivity pattern, and the required computation time. In addition, we investigate the effect of the number of microphones.

### A. Setup and Performance Measures

We use a linear nonuniform 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 behind-the-ear hearing aid. We assume that the microphone characteristics are independent of the angles $\phi$ and $\theta$, i.e., $A_n(\omega, \phi, \theta) = A_n(\omega)$, and that the nominal microphone characteristic $A_n(\omega) = 1$, $n = 0, \ldots, N-1$. Without loss of generality, we also assume that all microphone characteristics are described by the same probability density function $f_{\mathcal{A}}(A)$. The direction of the speech source is $\theta_s = 0^\circ$ (endfire), the sampling frequency $f_s = 16$ kHz and the design frequency is 1000 Hz.

TABLE I
Directivity Factor, Mean Directivity Factor, Worst-Case Directivity Factor, and Computation Time for Different Design Procedures ($N = 3$)

| Design procedure | $DF$ | $DF_m$ | $DF_{min}$ | Time [s] |
|---|---|---|---|---|
| $\mathbf{W}_{sd}$ ($\mu = 0$) | **9.52** | 1.33 | -28.12 | 1.6e-4 |
| $\mathbf{W}_{ds}$ ($\mu = \infty$) | 0.21 | 0.20 | 0.16 | 1.6e-4 |
| $\mathbf{W}_{sd,\mu}$ (max $DF_m$) | 5.97 | 4.88 | 1.82 | 1.6e-4 |
| $\mathbf{W}_{sd,\mu}$ (max $DF_{min}$) | 4.81 | 4.29 | 2.43 | 1.6e-4 |
| $\mathbf{W}_{tm,\lambda}$ (max $DF_m$) | 5.63 | 4.79 | 2.15 | 5.9e-4 |
| $\mathbf{W}_{tm,\lambda}$ (max $DF_{min}$) | 4.77 | 4.27 | 2.43 | 5.9e-4 |
| $\mathbf{W}_{LS,m}$ | 4.40 | 4.00 | 2.44 | 0.06 |
| $\mathbf{W}_m$ | 5.94 | **4.90** | 1.91 | 31.1 |
| $\mathbf{W}_{min}$ | 4.38 | 4.05 | **2.65** | 83.0 |

We compare the performance of the following *beamformer designs*, discussed in Sections III-B and IV:

1) $\mathbf{W}_{sd,\mu}$ using a WNG constraint, cf. (33), including the conventional superdirective beamformer $\mathbf{W}_{sd}(\mu = 0)$ and the delay-and-sum beamformer $\mathbf{W}_{ds}(\mu = \infty)$;
2) $\mathbf{W}_{tm,\lambda}$ minimizing the weighted sum of the mean noise and distortion energy, cf. (48);
3) $\mathbf{W}_{LS,m}$ minimizing the mean deviation from the desired superdirective directivity pattern, cf. (58) (We will assume that the weighting function $F(\omega, \phi, \theta)$ in (49) is equal to 1);
4) $\mathbf{W}_m$ maximizing the mean directivity factor;
5) $\mathbf{W}_{\min}$ maximizing the worst case directivity factor.

We will use the following *performance measures*:

1) the directivity factor $DF$ when no microphone deviations occur, cf. (17);
2) the mean directivity factor $DF_m$, cf. (61);
3) the worst case directivity factor $DF_{\min}$, cf. (65);
4) the mean noise energy $J_{vm}$, cf. (42);
5) the mean distortion energy $J_{dm}$, cf. (43).

Although these beamformers only need to be computed once during the design process, we will also compare the required computation time (AMD Opteron 250 2.4-GHz processor) to give an idea about the computational complexity for the different design procedures.

In the simulations, we will assume only gain deviations. We will use a uniform gain pdf with mean $\mu_{a,n} = 1$ and width $s_{a,n} = 0.3$, cf. Appendix I. The grid spacing used for computing $\mathbf{W}_m$, $DF_m$, $\mathbf{W}_{\min}$ and $DF_{\min}$ is $\Delta a = 0.02$, such that the sum in (61) and $\mathbf{F}(\mathbf{W})$ in (64) consist of $(2s_{a,n}/\Delta a)^N = 27\,000$ components.

### B. Directivity Factor

Table I summarizes the directivity factor $DF$, the mean directivity factor $DF_m$, the worst case directivity factor $DF_{\min}$,

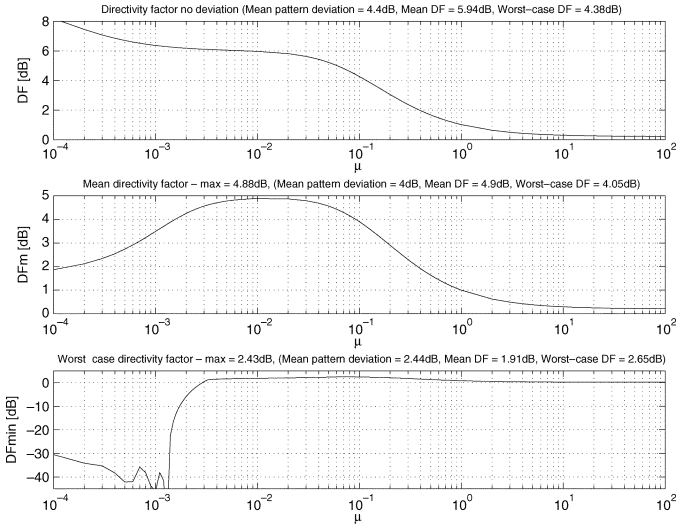



Fig. 4. Directivity factor, mean directivity factor and worst case directivity factor of $\mathbf{W}_{sd,\mu}$ as a function of $\mu$.

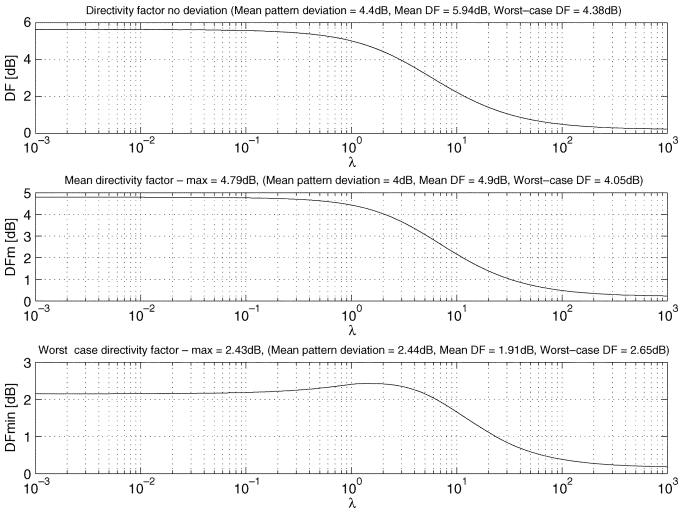



Fig. 5. Directivity factor, mean directivity factor and worst case directivity factor of $\mathbf{W}_{tm,\lambda}$ as a function of $\lambda$.

and the required computation time for the different design procedures. Obviously, the superdirective beamformer leads to the largest directivity factor when no microphone deviations occur ($DF = 9.52$ dB), the beamformer $\mathbf{W}_m$ leads to the largest mean directivity factor ($DF_m = 4.90$ dB), and the beamformer $\mathbf{W}_{\min}$ leads to the largest worst case directivity factor ($DF_{\min} = 2.65$ dB).

Fig. 4 plots the directivity factors for the beamformer $\mathbf{W}_{sd,\mu}$ as a function of the diagonal loading factor $\mu$. This factor provides a tradeoff between directivity and robustness: a small $\mu$ leads to a high directivity but a low robustness, while a high $\mu$ leads to a low directivity but a high robustness. Using this figure, it is possible to determine the values of $\mu$ for which the mean and the worst case directivity factor are maximized. For specific values, the directivity factors are summarized in Table I.

- The superdirective beamformer $\mathbf{W}_{sd}(\mu = 0)$ leads to the largest directivity factor when no deviations occur, but the mean directivity factor is only equal to $DF_m = 1.33$ dB, and the worst case directivity factor is even equal to $DF_{\min} = -28.12$ dB, illustrating the sensitivity of the superdirective beamformer to microphone deviations.
- The delay-and-sum beamformer $\mathbf{W}_{ds}(\mu = \infty)$ is very robust, but the directivity factor ($DF = 0.21$ dB), as well as the mean directivity factor ($DF_m = 0.20$ dB) and the worst case directivity factor ($DF_{\min} = 0.16$ dB), are all very small.
- For $\mu = 0.01$, the mean directivity factor is maximized ($DF_m = 4.88$ dB). This value is quite close to the maximum attainable value ($DF_m = 4.90$ dB), obtained by the beamformer $\mathbf{W}_m$.
- For $\mu = 0.07$, the worst case directivity factor is maximized ($DF_{\min} = 2.43$ dB). This value is quite close to the maximum attainable value ($DF_{\min} = 2.65$ dB), obtained by the beamformer $\mathbf{W}_{\min}$.

Fig. 5 plots the directivity factors for the beamformer $\mathbf{W}_{tm,\lambda}$ as a function of the weighting factor $\lambda$. Using this figure, it is possible to determine the values of $\lambda$ for which the mean and the worst case directivity factor are maximized. For these values of $\lambda$, the directivity factors are summarized in Table I.

- For $\lambda$ approaching 0, the mean directivity factor is maximized ($DF_m = 4.79$ dB). This value is quite close to the maximum attainable value ($DF_m = 4.90$ dB), obtained by the beamformer $\mathbf{W}_m$.
- For $\lambda = 1.4$, the worst case directivity factor is maximized ($DF_{\min} = 2.43$ dB). This value is quite close to the maximum attainable value ($DF_{\min} = 2.65$ dB), obtained by the beamformer $\mathbf{W}_{\min}$.

When comparing the computational complexity, it is obvious that the time required to compute the beamformers $\mathbf{W}_m$ and $\mathbf{W}_{\min}$ is much larger than the other design procedures.[3] Note that the required computation time for these two design procedures largely depends on the used grid spacing $\Delta a$.

Except for the superdirective beamformer $\mathbf{W}_{sd}$, which is very sensitive to deviations, and the delay-and-sum beamformer $\mathbf{W}_{ds}$, whose performance is very low, all other beamformer designs lead to a reasonable performance and robustness. Although it is hard to determine which design procedure is optimal, we can still make the following observations.

1) If computational complexity is not an issue, the beamformers $\mathbf{W}_m$ and $\mathbf{W}_{\min}$ are preferable, since they truly optimize the mean and the worst case directivity factor.
2) The performance of the beamformers $\mathbf{W}_{sd,\mu}$ and $\mathbf{W}_{tm,\lambda}$ is quite similar, where the parameters $\mu$ and $\lambda$ provide a tradeoff between directivity factor, mean directivity factor and worst case directivity factor. Using Figs. 4 and 5, it is possible to determine a suitable range for $\mu$ and $\lambda$. Note however that determining the specific values of $\mu$ and $\lambda$ that maximize the mean or the worst case directivity factor requires a multistep iterative procedure.
3) Although the beamformer $\mathbf{W}_{LS,m}$ leads to a large worst case directivity factor ($DF_{\min} = 2.44$ dB), its directivity factor and mean directivity factor are smaller than the other design procedures, making this the least preferable design procedure.

[3]The computation time required to determine the specific values of $\mu$ and $\lambda$ that maximize the mean or the worst case directivity factor has not been taken into account.

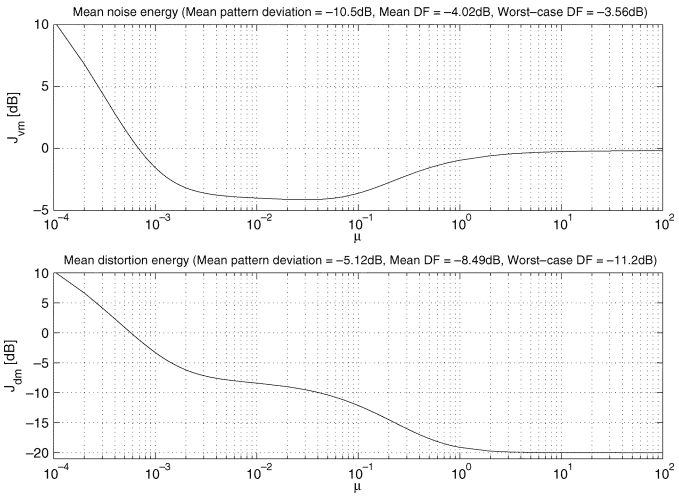



Fig. 6. Mean noise energy and mean distortion energy of $\mathbf{W}_{sd,\mu}$ as a function of $\mu$.

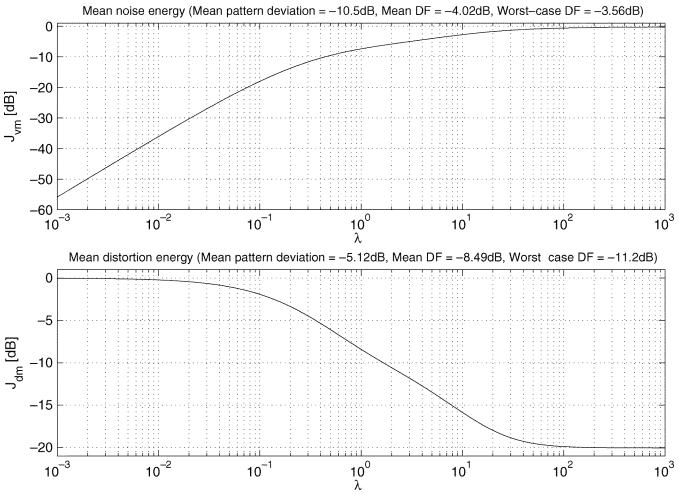



Fig. 7. Mean noise energy and mean distortion energy of $\mathbf{W}_{tm,\lambda}$ as a function of $\lambda$.

### C. Mean Noise and Distortion Energy

Fig. 6 plots the mean noise and distortion energy for the beamformer $\mathbf{W}_{sd,\mu}$ as a function of $\mu$. Fig. 7 plots the mean noise and distortion energy for the beamformer $\mathbf{W}_{tm,\lambda}$ as a function of $\lambda$. Since $\lambda$ provides a tradeoff between the mean noise and distortion energy, the mean distortion energy is a monotonically decreasing function, whereas the mean noise energy is a monotonically increasing function. Fig. 8 plots the mean distortion energy versus the mean noise energy for all discussed beamformers. From this figure, we can observe the following.

- The superdirective beamformer $\mathbf{W}_{sd}$ leads to both a large mean noise energy and a large mean distortion energy, illustrating its sensitivity to microphone deviations. Of all beamformers $\mathbf{W}_{sd,\mu}$, the delay-and-sum beamformer $\mathbf{W}_{ds}$ produces the smallest mean distortion energy, but not the smallest mean noise energy.
- For every beamformer $\mathbf{W}_{sd,\mu}$, there exists a beamformer $\mathbf{W}_{tm,\lambda}$ for which both the mean noise energy and the mean distortion energy are smaller.

### D. Spatial Directivity Pattern

In this section, we discuss the spatial directivity pattern of the presented beamformers when no deviation is present and when a specific gain deviation [0.7 1.3 1.2] occurs.

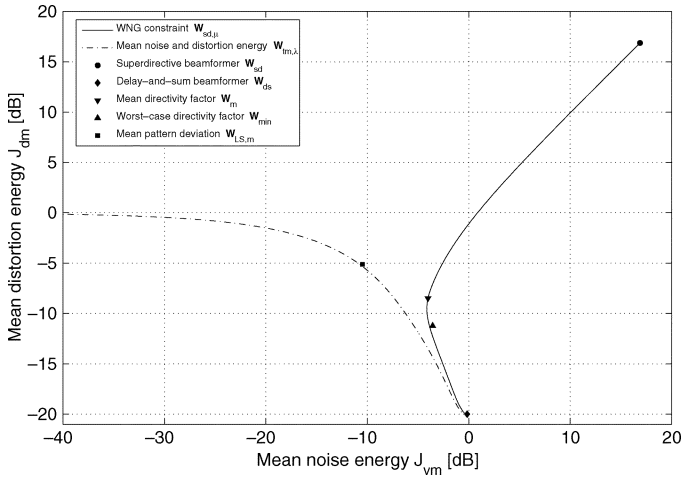



Fig. 8. Mean noise energy versus mean distortion energy for all discussed beamformers.

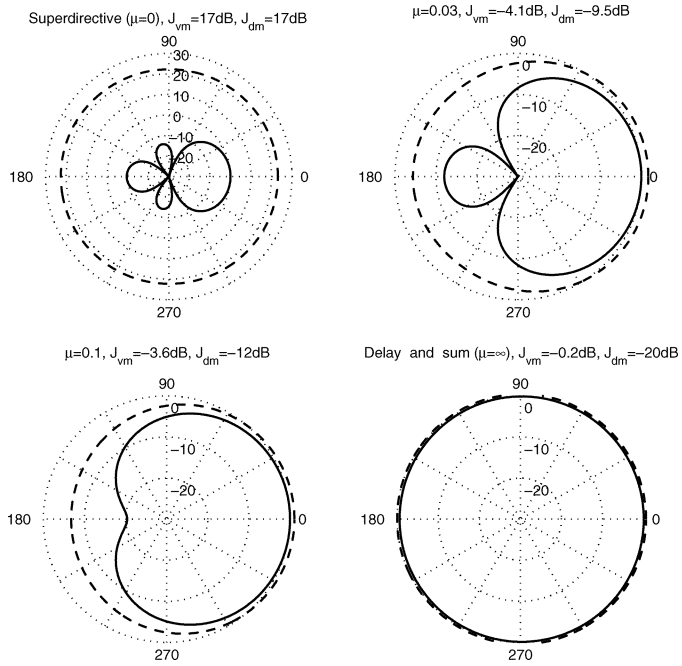



Fig. 9. Spatial directivity pattern of $\mathbf{W}_{sd,\mu}$ without deviation (solid line) and with deviation (dashed line) for different values of $\mu$.

Fig. 9 plots the spatial directivity pattern of the beamformer $\mathbf{W}_{sd,\mu}$ for different values of $\mu$. As can be seen from this figure, $\mu$ provides a tradeoff between directivity and robustness. The superdirective beamformer $(\mu = 0)$ exhibits a highly directional pattern when no deviation is present, but it is very sensitive to deviations. On the other hand, the delay-and-sum beamformer $(\mu = \infty)$ is very robust to deviations, but its spatial directivity pattern is almost omnidirectional.

Fig. 10 plots the spatial directivity pattern of the beamformer $\mathbf{W}_{tm,\lambda}$ for different values of $\lambda$. As can be seen from this figure, $\lambda$ also provides a tradeoff between directivity and robustness.

Fig. 11 plots the spatial directivity pattern for the beamformer $\mathbf{W}_{LS,m}$ minimizing the mean deviation from the desired superdirective directivity pattern, the beamformer $\mathbf{W}_m$ maximizing the mean directivity factor, and the beamformer $\mathbf{W}_{\min}$ maximizing the worst case directivity factor.

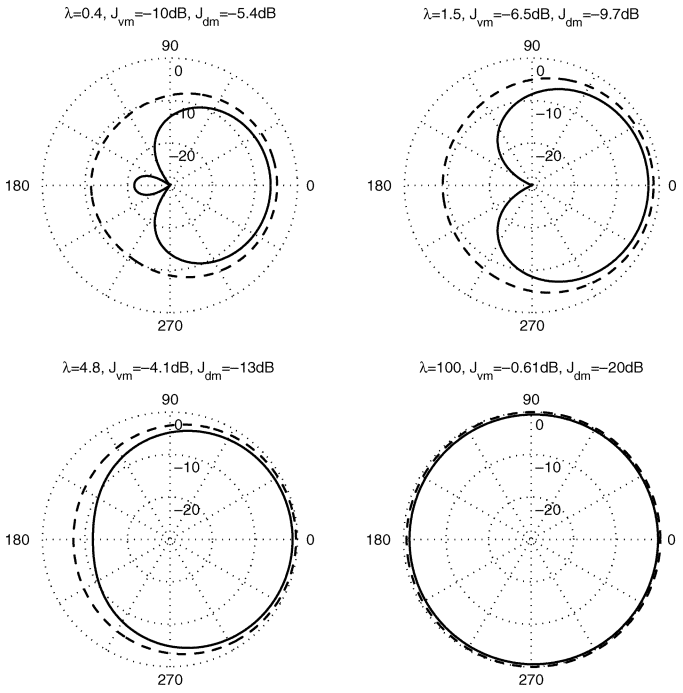



Fig. 10. Spatial directivity pattern of $\mathbf{W}_{tm,\lambda}$ without deviation (solid line) and with deviation (dashed line) for different values of $\lambda$.

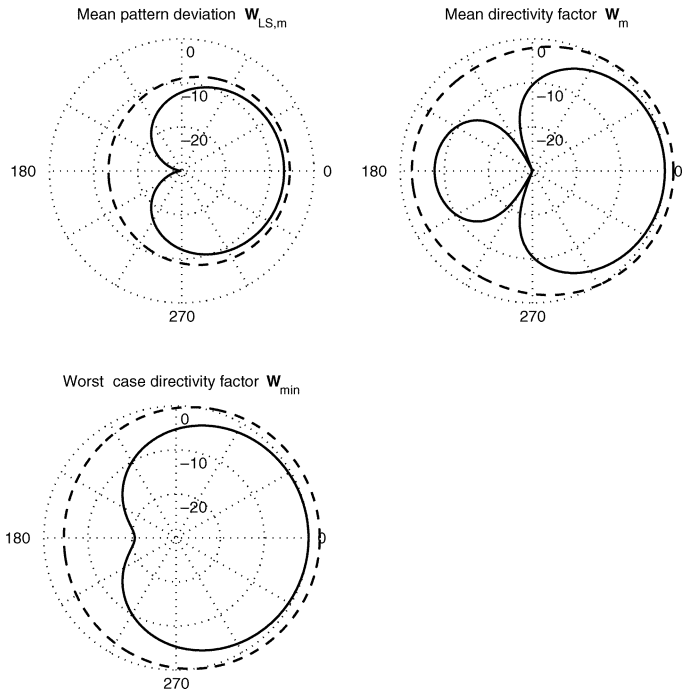



Fig. 11. Spatial directivity pattern of $\mathbf{W}_{LS,m}$, $\mathbf{W}_m$, and $\mathbf{W}_{\min}$ without deviation (solid line) and with deviation (dashed line).

### *E. Effect of the Number of Microphones*

In this section, we investigate the effect of the number of microphones on the performance, the robustness and the computation time for the different beamformer design procedures. Table II summarizes the directivity factors and the required computation time when using two and four microphones. For $N = 2$, the microphone positions are [0 0.01] m, and for $N = 4$ the microphone positions are [0 0.01 0.025 0.04] m. Apart from the number of microphones, we have used the same setup as described in Section V-A.

For $N = 2$ and $N = 4$, similar conclusions can be drawn as in Section V-B, i.e., the superdirective beamformer $\mathbf{W}_{sd}(\mu = 0)$ leads to the largest directivity factor when no microphone deviations occur, but is quite sensitive to microphone mismatch, whereas the delay-and-sum beamformer $\mathbf{W}_{ds}(\mu = \infty)$ is very robust, but the directivity factor is small. For specific values of the parameters $\mu$ and $\lambda$ maximizing the mean or the worst case directivity factor, the performance of the beamformers $\mathbf{W}_{sd,\mu}$ and $\mathbf{W}_{tm,\lambda}$ is comparable to the performance of the beamformers $\mathbf{W}_m$ and $\mathbf{W}_{\min}$ which truly optimize the mean and the worst case directivity factor.

As can be seen from Tables I and II, the directivity factor $DF$ of the superdirective beamformer $\mathbf{W}_{sd}(\mu = 0)$ increases logarithmically with the number of microphones $N$, whereas the mean directivity factor $DF_m$ and in particular the worst case directivity factor $DF_{\min}$ decrease quite substantially. For $N = 2$, the robustness of the superdirective beamformer is still quite reasonable, i.e., $DF_m = 3.49$ dB compared to the maximum attainable value 3.59 dB, and $DF_{\min} = 0.45$ dB compared to the maximum attainable value 1.15 dB, but this is definitely not the case for $N = 3$ and $N = 4$. Hence, the superdirective beamformer becomes more sensitive to microphone mismatch as the number of microphones increases, making a robust design more imperative. For the robust design procedures, all directivity factors increase as the number of microphones increases, illustrating their robustness against microphone mismatch. For the beamformers $\mathbf{W}_m$ and $\mathbf{W}_{\min}$, the computation time, however, grows exponentially with the number of microphones.

## VI. Conclusion

In this paper, we have presented several design procedures for improving the robustness of superdirective beamformers against unknown microphone mismatch by taking into account the statistics of the microphone characteristics. We consider minimizing the weighted sum of the mean noise and distortion energy, minimizing the mean deviation from the desired superdirective directivity pattern, and maximizing the mean or the worst case directivity factor. When computational complexity is not an issue, maximizing the mean or the worst case directivity factor is the preferred design procedure. In addition, it has been shown how to determine a suitable parameter range for $\mathbf{W}_{sd,\mu}$ and $\mathbf{W}_{tm,\lambda}$ such that both a high directivity and a high level of robustness are obtained.

## Appendix I
## Calculation of Mean nd Variance Expressions

This appendix describes the calculation of the mean and variance expressions

$$\mu_{A,n}(\omega,\phi,\theta) = \int_{A_n} A_n(\omega,\phi,\theta) f_{\mathcal{A}}(A_n) dA_n$$

$$\sigma^2_{A,np}(\omega,\phi,\theta) = \int_{A_n}\int_{A_p} A_n(\omega,\phi,\theta) A_p^*(\omega,\phi,\theta) \times f_{\mathcal{A}}(A_n) f_{\mathcal{A}}(A_p) dA_n dA_p$$

TABLE II
DIRECTIVITY FACTOR, MEAN DIRECTIVITY FACTOR, WORST-CASE DIRECTIVITY FACTOR, AND COMPUTATION TIME FOR DIFFERENT DESIGN PROCEDURES AND NUMBER OF MICROPHONES

| | $N=2$ | | | | $N=4$ | | | |
|---|---|---|---|---|---|---|---|---|
| **Design procedure** | $DF$ | $DF_m$ | $DF_{min}$ | Time [s] | $DF$ | $DF_m$ | $DF_{min}$ | Time [s] |
| $\mathbf{W}_{sd}$ ($\mu=0$) | **6.01** | 3.49 | 0.45 | 1.1e-4 | **12.02** | 0.96 | -26.34 | 2.4e-4 |
| $\mathbf{W}_{ds}$ ($\mu=\infty$) | 0.05 | 0.05 | 0.04 | 1.1e-4 | 0.44 | 0.44 | 0.35 | 2.4e-4 |
| $\mathbf{W}_{sd,\mu}$ (max $DF_m$) | 5.81 | 3.59 | 0.57 | 1.1e-4 | 6.07 | 5.44 | 2.90 | 2.4e-4 |
| $\mathbf{W}_{sd,\mu}$ (max $DF_{min}$) | 2.21 | 1.96 | 1.14 | 1.1e-4 | 5.17 | 4.87 | 3.10 | 2.4e-4 |
| $\mathbf{W}_{tm,\lambda}$ (max $DF_m$) | 3.09 | 2.53 | 1.09 | 5.6e-4 | 5.99 | 5.43 | 2.92 | 6.6e-4 |
| $\mathbf{W}_{tm,\lambda}$ (max $DF_{min}$) | 2.36 | 2.06 | 1.15 | 5.6e-4 | 5.13 | 4.84 | 3.10 | 6.6e-4 |
| $\mathbf{W}_{LS,m}$ | 3.85 | 2.94 | 0.97 | 0.02 | 5.88 | 5.39 | 2.97 | 0.11 |
| $\mathbf{W}_m$ | 5.81 | **3.59** | 0.57 | 2.32 | 6.05 | **5.44** | 2.83 | 754 |
| $\mathbf{W}_{min}$ | 2.35 | 2.06 | **1.15** | 2.44 | 5.31 | 4.96 | **3.66** | 3129 |

for different probability density functions (uniform, log-uniform, normal, log-normal). Since the joint pdf $f_{\mathcal{A}}(A)$ is separable, the *mean* $\mu_{A,n}(\omega,\phi,\theta)$ can be written as

$$\mu_{A,n}(\omega,\phi,\theta)=\mu_{a,n}(\omega,\phi,\theta)\mu_{\psi,n}(\omega,\phi,\theta)\mu_{\delta,n}(\omega,\phi,\theta) \quad (67)$$

with

$$\mu_{a,n}(\omega,\phi,\theta)=\int_{a_n} a_n(\omega,\phi,\theta)f_\alpha(a_n)da_n$$
$$\mu_{\psi,n}(\omega,\phi,\theta)=\int_{\psi_n} e^{-j\psi_n(\omega,\phi,\theta)}f_\Psi(\psi_n)d\psi_n$$
$$\mu_{\delta,n}(\omega,\phi,\theta)=\int_{\delta_n} e^{-j\omega\frac{\delta_n\cos\theta}{c}f_s}f_\Delta(\delta_n)d\delta_n.$$

The calculation of $\mu_{a,n}(\omega,\phi,\theta)$ for different pdfs is discussed in Appendix I-B, while the calculation of $\mu_{\psi,n}(\omega,\phi,\theta)$ and $\mu_{\delta,n}(\omega,\phi,\theta)$ is discussed in Appendix I-C. The *variance* $\sigma^2_{A,np}(\omega,\phi,\theta)$ is equal to

$$\begin{aligned}\sigma^2_{A,np}(\omega,\phi,\theta)&=\mu_{A,n}(\omega,\phi,\theta)\mu^*_{A,p}(\omega,\phi,\theta), \quad p\neq n\\ &=\sigma^2_{a,n}(\omega,\phi,\theta), \quad p=n\end{aligned} \quad (68)$$

with

$$\sigma^2_{a,n}(\omega,\phi,\theta)=\int_{a_n} a_n^2(\omega,\phi,\theta)f_\alpha(a_n)da_n.$$

The calculation of $\sigma^2_{a,n}(\omega,\phi,\theta)$ is discussed in Appendix I-B.

### A. Probability Density Functions

For the gain pdf, we will consider four different pdfs: uniform, log-uniform, normal and log-normal. For the phase and the microphone position error pdf, we will only consider the uniform and the normal pdf.

*1) Uniform and Log-Uniform pdf:* The uniform pdf $f(x)$ with mean $u_x$ and width $s_x$ is described by

$$f(x)=\begin{cases}\frac{1}{2s_x}, & x_{\min}\le x\le x_{\max}\\ 0, & x<x_{\min}, x>x_{\max}\end{cases} \quad (69)$$

with $x_{\min}=u_x-s_x$ and $x_{\max}=u_x+s_x$. In the logarithmic domain, the uniform pdf $h(X)$ with mean $u_X$ and width $s_X$ (in decibels) is depicted in Fig. 12(a). Using $X=10\log_{10}(x)$, the uniform pdf $h(X)$ can be transformed [17] to the log-uniform pdf $f(x)$, i.e.,

$$f(x)=\begin{cases}\frac{5}{x\log(10)s_X}, & x_{\min}\le x\le x_{\max}\\ 0, & x<x_{\min}, x>x_{\max}\end{cases} \quad (70)$$

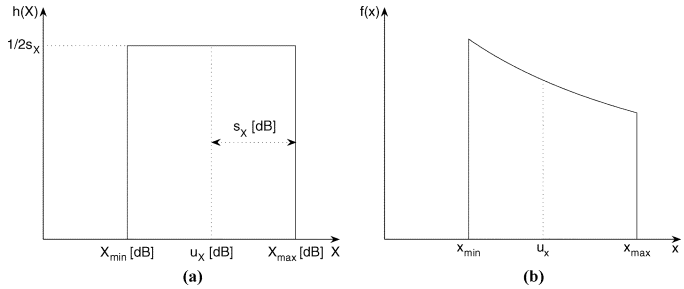



Fig. 12. Log-uniform pdf (mean $u_X$, width $s_X$).

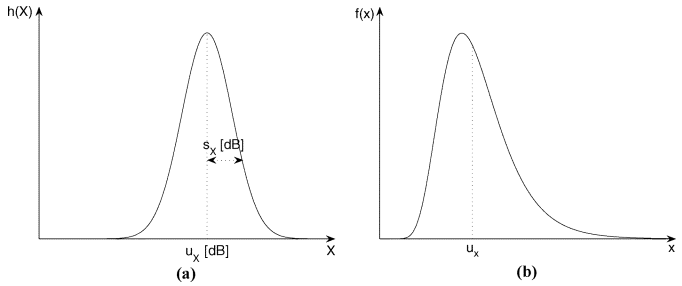



Fig. 13. Log-normal pdf (mean $u_X$, variance $s_X$).

with

$$x_{\min}=10^{X_{\min}/10}=10^{(u_X-s_X)/10}$$
$$x_{\max}=10^{X_{\max}/10}=10^{(u_X+s_X)/10}.$$

The log-uniform pdf $f(x)$ is depicted in Fig. 12(b).

*2) Normal and Log-Normal pdf:* The normal pdf $f(x)$ with mean $u_x$ and width $s_x$ is described by

$$f(x)=\frac{1}{\sqrt{2\pi s_x^2}}\exp\left[-\frac{(x-u_x)^2}{2s_x^2}\right]. \quad (71)$$

In the logarithmic domain, the normal pdf $h(X)$ with mean $u_X$ and width $s_X$ (in decibels) is depicted in Fig. 13(a). The normal pdf $h(X)$ can be transformed to the log-normal pdf $f(x)$, i.e.,

$$f(x)=\frac{10\exp\left[-\frac{(10\log_{10}(x)-u_X)^2}{2s_X^2}\right]}{x\log(10)\sqrt{2\pi s_X^2}} \quad (72)$$

which is depicted in Fig. 13(b).

### B. Calculation of $\mu_{a,n}(\omega,\phi,\theta)$ and $\sigma^2_{a,n}(\omega,\phi,\theta)$

In this section, the mean and variance expressions are calculated for all pdfs discussed in Appendix I-A. For the sake of conciseness, we will omit the variables $(\omega,\phi,\theta)$ in this section.

*1) Uniform pdf:* Using (69), the mean and variance for a uniform gain pdf with mean $u_{an}$ and width $s_{an}$ are equal to

$$\mu_{a,n} = u_{an}, \quad \sigma^2_{a,n} = u^2_{an} + \frac{s^2_{an}}{3}.$$

*2) Log-Uniform pdf:* Using (70), the mean for a log-uniform gain pdf with mean $u_{An}$ and width $s_{An}$ is equal to

$$\begin{aligned}\mu_{a,n} &= \int_{a_{n,\min}}^{a_{n,\max}} \frac{5}{\log(10) s_{An}} da_n \\ &= \frac{5}{\log(10) s_{An}} \left[10^{(u_{An}+s_{An})/10} - 10^{(u_{An}-s_{An})/10}\right],\end{aligned}$$

and the variance is equal to

$$\begin{aligned}\sigma^2_{a,n} &= \int_{a_{n,\min}}^{a_{n,\max}} \frac{5a_n}{\log(10) s_{An}} da_n \\ &= \frac{5}{2\log(10) s_{An}} \left[10^{(u_{An}+s_{An})/5} - 10^{(u_{An}-s_{An})/5}\right].\end{aligned}$$

*3) Normal pdf:* Using (71), the mean and variance for a normal gain pdf with mean $u_{an}$ and width $s_{an}$ are equal to

$$\mu_{a,n} = u_{an}, \quad \sigma^2_{a,n} = s^2_{an} + u^2_{an}.$$

*4) Log-Normal pdf:* Using (72), the mean for a log-normal gain pdf with mean $u_{An}$ and width $s_{An}$ is equal to

$$\mu_{a,n} = \int_0^\infty \frac{10 \exp\left[-\frac{(10\log_{10}(a_n)-u_{An})^2}{2s^2_{An}}\right]}{\log(10)\sqrt{2\pi s^2_{An}}} da_n.$$

Using the substitution $z = 10\log_{10}(a_n) - u_{An}$, we obtain

$$\mu_{a,n} = \int_{-\infty}^\infty \frac{\exp\left[-\frac{z^2}{2s^2_{An}}\right]}{\sqrt{2\pi s^2_{An}}} \exp\left[\frac{z+u_{An}}{10\log_{10} e}\right] dz.$$

Using [18]

$$\int_{-\infty}^\infty e^{-(ax^2+bx+c)} dx = \sqrt{\frac{\pi}{a}} \exp\left[\frac{b^2}{4a} - c\right]$$

the mean $\mu_{a,n}$ is equal to

$$\mu_{a,n} = 10^{u_{An}/10} \exp\left[\frac{s^2_{An}}{2(10\log_{10} e)^2}\right].$$

Using (72), the variance for a log-normal gain pdf is equal to

$$\begin{aligned}\sigma^2_{a,n} &= \int_0^\infty \frac{10a_n \exp\left[-\frac{(10\log_{10}(a_n)-u_{An})^2}{2s^2_{An}}\right]}{\log(10)\sqrt{2\pi s^2_{An}}} da_n \\ &= \int_{-\infty}^\infty \frac{\exp\left[-\frac{z^2}{2s^2_{An}}\right]}{\sqrt{2\pi s^2_{An}}} \exp\left[\frac{2(z+u_{An})}{10\log_{10} e}\right] dz \\ &= 10^{u_{An}/5} \exp\left[\frac{2s^2_{An}}{(10\log_{10} e)^2}\right].\end{aligned}$$

### C. Calculation of $\mu_{\psi,n}(\omega,\phi,\theta)$ and $\mu_{\delta,n}(\omega,\phi,\theta)$

In this section, the mean and variance expressions $\mu_{\psi,n}(\omega,\phi,\theta)$ and $\mu_{\delta,n}(\omega,\phi,\theta)$ are calculated for the uniform and normal pdfs. For the sake of conciseness, we will omit the variables $(\omega,\phi,\theta)$ in this section.

*1) Uniform pdf:* Using (69), the mean for a uniform phase pdf with mean $u_{\psi n}$ and width $s_{\psi n}$ is equal to

$$\mu_{\psi,n} = \int_{u_{\psi n}-s_{\psi n}}^{u_{\psi n}+s_{\psi n}} \frac{e^{-j\psi_n}}{2s_{\psi n}} d\psi_n = e^{-ju_{\psi n}} \frac{\sin(s_{\psi n})}{s_{\psi n}}.$$

Since a microphone position error $\delta_n$ corresponds to a frequency- and angle-dependent phase error $\omega(\delta_n \cos\theta/c) f_s$, the mean for a uniform position error pdf with mean $u_{\delta n}$ and width $s_{\delta n}$ is equal to

$$\mu_{\delta,n}(\omega,\phi,\theta) = e^{-j\omega\frac{u_{\delta n}\cos\theta}{c} f_s} \frac{\sin\left(\omega\frac{s_{\delta n}\cos\theta}{c} f_s\right)}{\omega\frac{s_{\delta n}\cos\theta}{c} f_s}.$$

*2) Normal pdf:* Using (71), the mean for a normal phase pdf with mean $u_{\psi n}$ and width $s_{\psi n}$ is equal to

$$\begin{aligned}\mu_{\psi,n} &= \int_{-\infty}^\infty \frac{e^{-j\psi_n}}{\sqrt{2\pi s^2_{\psi n}}} \exp\left[-\frac{(\psi_n - u_{\psi n})^2}{2s^2_{\psi n}}\right] d\psi_n \\ &= e^{-ju_{\psi n}} \int_{-\infty}^\infty \frac{e^{-j\psi_n}}{\sqrt{2\pi s^2_{\psi n}}} \exp\left[-\frac{\psi_n^2}{2s^2_{\psi n}}\right] d\psi_n.\end{aligned}$$

Using [18]

$$\int_{-\infty}^\infty e^{-jbx} e^{-ax^2} dx = \int_{-\infty}^\infty \cos(bx) e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}}$$

the mean $\mu_{\psi,n}$ is equal to

$$\mu_{\psi,n} = e^{-ju_{\psi n}} e^{-\frac{s^2_{\psi n}}{2}}.$$

Similarly, the mean for a normal position error pdf with mean $u_{\delta n}$ and width $s_{\delta n}$ is equal to

$$\mu_{\delta,n}(\omega,\phi,\theta) = e^{-j\omega\frac{u_{\delta n}\cos\theta}{c} f_s} \exp\left[-\frac{1}{2}\left(\omega\frac{s_{\delta n}\cos\theta}{c} f_s\right)^2\right].$$

## APPENDIX II

This appendix discusses the calculation of the mean diffuse noise correlation matrix $\widetilde{\Phi}^{\text{diff}}_{m,np}(\omega)$ in (37), i.e.,

$$\frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \sigma^2_{A,np}(\omega,\phi,\theta) e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))} \sin\theta d\theta d\phi \quad (73)$$

for different probability density functions.

Using (68), if $p = n$, (73) is equal to

$$\widetilde{\Phi}^{\text{diff}}_{m,nn}(\omega) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \sigma^2_{a,n}(\omega,\phi,\theta) \sin\theta d\theta d\phi.$$

$$\widetilde{\Phi}^{\text{diff}}_{m,np}(\omega) = \mu_{a,n}(\omega)\mu_{a,p}(\omega)\mu_{\psi,n}(\omega)\mu^{*}_{\psi,p}(\omega) \times \underbrace{\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} \mu_{\delta,n}(\omega,\phi,\theta)\mu^{*}_{\delta,p}(\omega,\phi,\theta)\cdot e^{-j\omega(\tau_n(\phi,\theta)-\tau_p(\phi,\theta))}\sin\theta d\theta d\phi}_{I_{\delta,np}(\omega)} \tag{74}$$

$$I_{\delta,np}(\omega) = \frac{1}{2}\int_0^{\pi} \frac{\sin\left(\omega\frac{s_{\delta n}\cos\theta}{c}f_s\right)}{\omega\frac{s_{\delta n}\cos\theta}{c}f_s}\frac{\sin\left(\omega\frac{s_{\delta p}\cos\theta}{c}f_s\right)}{\omega\frac{s_{\delta p}\cos\theta}{c}f_s}e^{-j\omega\frac{(d_n-d_p)\cos\theta}{c}f_s}\sin\theta d\theta$$

$$= \int_0^{1} \frac{\sin\left(\omega\frac{s_{\delta n}u}{c}f_s\right)}{\omega\frac{s_{\delta n}u}{c}f_s}\frac{\sin\left(\omega\frac{s_{\delta p}u}{c}f_s\right)}{\omega\frac{s_{\delta p}u}{c}f_s}\cos\left[\omega\frac{(d_n-d_p)u}{c}f_s\right]du \tag{75}$$

$$I_{\delta,np}(\omega) = \frac{1}{2}\int_0^{\pi} \exp\left[-\frac{1}{2}\left(\omega\frac{s_{\delta n}\cos\theta}{c}f_s\right)^2\right]\exp\left[-\frac{1}{2}\left(\omega\frac{s_{\delta p}\cos\theta}{c}f_s\right)^2\right]\cdot e^{-j\omega\frac{(d_n-d_p)\cos\theta}{c}f_s}\sin\theta d\theta$$

$$= \int_0^{1} \exp\left[-\frac{1}{2}\left(\omega\frac{s_{\delta n}u}{c}f_s\right)^2-\frac{1}{2}\left(\omega\frac{s_{\delta p}u}{c}f_s\right)^2\right]\cos\left[\omega\frac{(d_n-d_p)u}{c}f_s\right]du \tag{76}$$

Assuming that $\sigma^2_{a,n}(\omega,\phi,\theta)$ is independent of $\phi$ and $\theta$, i.e., $\sigma^2_{a,n}(\omega,\phi,\theta)=\sigma^2_{a,n}(\omega)$

$$\widetilde{\Phi}^{\text{diff}}_{m,nn}(\omega) = \sigma^2_{a,n}(\omega)$$

where expressions for $\sigma^2_{a,n}(\omega)$ for different pdfs have been calculated in Appendix I.

Using (68), if $p \neq n$, (73) is equal to

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

Assuming that $\mu_{a,n}(\omega,\phi,\theta)$ and $\mu_{\psi,n}(\omega,\phi,\theta)$ are independent of the angles $\phi$ and $\theta$, i.e., $\mu_{a,n}(\omega,\phi,\theta)=\mu_{a,n}(\omega)$ and $\mu_{\psi,n}(\omega,\phi,\theta)=\mu_{\psi,n}(\omega)$, $\widetilde{\Phi}^{\text{diff}}_{m,np}(\omega)$ is equal to (74), shown at the top of the page, where expressions for $\mu_{a,n}(\omega)$, $\mu_{\psi,n}(\omega)$ and $\mu_{\delta,n}(\omega,\phi,\theta)$ for different pdfs have been calculated in Appendix I.

- For *uniform* position error pdfs $\mu_{\delta,n}(\omega,\phi,\theta)$ and $\mu_{\delta,p}(\omega,\phi,\theta)$ with $u_{\delta n}=u_{\delta p}=0$, $I_{\delta,np}(\omega)$ is equal to (75), which needs to be computed numerically.
- For *normal* position error pdfs $\mu_{\delta,n}(\omega,\phi,\theta)$ and $\mu_{\delta,p}(\omega,\phi,\theta)$ with $u_{\delta n}=u_{\delta p}=0$, $I_{\delta,np}(\omega)$ is equal to (76), which needs to be computed numerically.

## REFERENCES


[1] E. N. Gilbert and S. P. Morgan, "Optimum design of directive antenna arrays subject to random deviations," *Bell Syst. Tech. J.*, vol. 34, pp. 637–663, May 1955.

[2] H. Cox, R. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 3, pp. 393–398, Jun. 1986.

[3] 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, May 2001, ch. 2, pp. 19–38.

[4] J. G. Ryan and R. A. Goubran, "Array optimization applied in the near field of a microphone array," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 2, pp. 173–176, Mar. 2000.

[5] S. Yan and Y. Ma, "Robust supergain beamforming for circular array via second-order cone programming," *Appl. Acoust.*, vol. 66, no. 9, pp. 1018–1032, Sep. 2005.

[6] G. Elko, "Superdirectional Microphone Arrays," in *Acoustic signal processing for telecommunication*, S. L. Gay and J. Benesty, Eds. Boston, MA: Kluwer, 2000, ch. 10, pp. 181–237.

[7] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Amer.*, vol. 97, no. 2, pp. 91–95, Feb. 1995.

[8] ——, "Constant Directivity Beamforming," in *Microphone arrays: Signal processing techniques and applications,* M. S. Brandstein and D. B. Ward, Eds. New York: Springer-Verlag, May 2001, ch. 1, pp. 3–17.

[9] J. M. Kates, "Superdirective arrays for hearing aids," *J. Acoust. Soc. Amer.*, vol. 94, no. 4, pp. 1930–1933, Oct. 1993.

[10] W. Soede, A. J. Berkhout, and F. A. Bilsen, "Development of a directional hearing instrument based on array technology," *J. Acoust. Soc. Amer.*, vol. 94, no. 2, pp. 785–798, Aug. 1993.

[11] R. W. Stadler and W. M. Rabinowitz, "On the potential of fixed arrays for hearing aids," *J. Acoust. Soc. Amer.*, vol. 94, no. 3, pp. 1332–1342, Sep. 1993.

[12] L. B. Jensen, "Hearing aid with adaptive matching of input transducers," U.S. Patent 6,741,714, May 25, 2004.

[13] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, pp. 41–44.

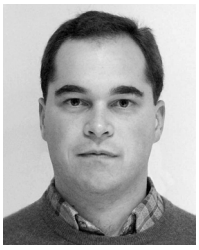
[14] ——, "Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2511–2526, Oct. 2003.

[15] ——, "Design of broadband beamformers robust against microphone position errors," in *Proc. Int. Workshop Acoust. Echo and Noise Control (IWAENC)*, Kyoto, Japan, Sep. 2003, pp. 267–270.

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

[17] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill Education, 1991.

[18] M. R. Spiegel and J. Liu, *Mathematical Handbook of Formulas and Tables*, 2nd ed. New York: McGraw-Hill, 1999.



**Simon Doclo** (S'95–M'03) was born in Wilrijk, Belgium, in 1974. He received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1997 and 2003, respectively.

Currently, he is a Postdoctoral Fellow of the Fund for Scientific Research–Flanders, affiliated with the Electrical Engineering Department of the Katholieke Universiteit Leuven. In 2005, he was a Visiting Postdoctoral Fellow at the Adaptive Systems Laboratory, McMaster University, Hamilton, ON, Canada. His research interests are in microphone array processing for acoustic noise reduction, dereverberation and sound localization, adaptive filtering, speech enhancement, and hearing aid processing.

Dr. Doclo received the First Prize "KVIV-Studentenprijzen" (with E. De Clippel) for the best M.Sc. engineering thesis in Flanders in 1997, a Best Student Paper Award at the International Workshop on Acoustic Echo and Noise Control in 2001, and the EURASIP Signal Processing Best Paper Award 2003 (with M. Moonen). He was secretary of the IEEE Benelux Signal Processing Chapter from 1997 to 2002, and serves as Guest Editor for the *EURASIP Journal on Applied Signal Processing*.

**Marc Moonen** (M'94–SM'06) received the electrical engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1986 and 1990, respectively.

Since 2004, he has been a Full Professor with the Electrical Engineering Department, Katholieke Universiteit Leuven, where he is currently heading a research team of 16 Ph.D. candidates and postdocs, working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL, and audio signal processing.

Dr. Moonen received the 1994 K.U. Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with P. Vandaele), the 2004 Alcatel Bell (Belgium) Award (with R. Cendrillon), and was a 1997 "Laureate of the Belgium Royal Academy of Science." He received a journal best paper award from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with G. Leus) and from Elsevier Signal Processing (with S. Doclo). He was Chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), and is currently a EURASIP AdCom Member (European Association for Signal, Speech, and Image Processing, 2000) and a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications. He served as Editor-in-Chief for the EURASIP Journal on Applied Signal Processing (2003–2005), and was a member of the editorial board of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (2002-2003) and the IEEE *Signal Processing Magazine* (2003 2005). He is currently a member of the editorial board of *Integration*, the *VLSI Journal*, the *EURASIP Journal on Applied Signal Processing*, the *EURASIP Journal on Wireless Communications and Networking*, and *Signal Processing*.