Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. I. Prediction of the signal-to-noise-ratio improvement

Martin Kompis[a)]
University Clinic of ENT, Head and Neck Surgery, Inselspital, 3010 Bern, Switzerland

Norbert Dillier
ENT-Department, University Hospital of Zurich, 8091 Zurich, Switzerland

(Received 5 November 1999; revised 24 April 2000; revised 25 September 2000; accepted 14 November 2000)

Adaptive beamformers have been proposed as noise reduction schemes for conventional hearing aids and cochlear implants. A method to predict the amount of noise reduction that can be achieved by a two-microphone adaptive beamformer is presented. The prediction is based on a model of the acoustic environment in which the presence of one acoustic target-signal source and one acoustic noise source in a reverberant enclosure is assumed. The acoustic field is sampled using two omnidirectional microphones mounted close to the ears of a user. The model takes eleven different parameters into account, including reverberation time and size of the room, directionality of the acoustic sources, and design parameters of the beamformer itself, including length of the adaptive filter and delay in the target signal path. An approximation to predict the achievable signal-to-noise improvement based on the model is presented. Potential applications as well as limitations of the proposed prediction method are discussed and a FORTRAN subroutine to predict the achievable signal-to-noise improvement is provided. Experimental verification of the predictions is provided in a companion paper [J. Acoust. Soc. Am. **109**, 1134 (2001)]. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1338557]

PACS numbers: 43.66.Ts, 43.60.Lq, 43.60.Gk [RVS]

LIST OF SYMBOLS

A,B,C,D	models of impulse responses between acoustic sources and input of adaptive filter (cf. Fig. 2)
a_i	ith coefficient of filter A
b_i	ith coefficient of filter B
c	sound speed, m/s
d	sum of both microphone signals, delayed by Δ samples
d'	sum of both microphone signals
$E\{\ \}$	expected value
F_d	coefficient to scale the direct portion of the impulse responses A and B
F_σ	coefficient to scale the reverberant portion of the impulse responses A and B
F_{sample}	sampling rate $=1/T_{\text{sample}}$, Hz
G_0	magnitude of the first coefficient of the impulse responses A and B
G_1	magnitude of the second coefficient of the impulse responses A and B
G_{nR}	impulse response between noise source and output signal of right microphone
G_{nL}	impulse response between noise source and output signal of left microphone
G_{sR}	impulse response between target signal source and output signal of right microphone
G_{sL}	impulse response between target signal source and output signal of left microphone
$g_{\text{nR},i}$	ith coefficient of filter G_{nR}
$g_{\text{nL},i}$	ith coefficient of filter G_{nL}
h	noise reduction of the adaptive filter, defined as $E\{d^2\}-E\{\epsilon^2\}$
k	sample index
l_n	distance between noise source and center of listener's head, m
l_s	distance between target signal source and center of listener's head, m
n	signal emitted by the noise source
N	number of coefficients in the adaptive filter W
N_1,N_2,N_B,N_S	variances of noise signal at microphone 1, microphone 2, sum of microphone signals, and output of beamformer, respectively
P	cross-correlation vector
P_i	ith element of the cross-correlation vector P
$P_{\text{d/r}}$	direct-to-reverberant ratio of the noise signal at location of the listener
$Q_{\text{d/r}}$	direct-to-reverberant ratio of the target signal at location of the listener
r_c	critical distance, m
R	autocorrelation matrix
s	signal emitted by the target source
$S(\vartheta)$	ratio between rms value of a white noise signal in free field and on the surface of a rigid sphere
S_1,S_2,S_B,S_S	variances of target signal at microphone 1, microphone 2, sum of microphone signals, and output of beamformer, respectively

[a)]Electronic mail: martin.kompis@insel.ch

T	time constant for exponential decay of the filter coefficients modeling reverberation in impulse responses A and B, in multiples of the sampling period T_{sample}
T_r	reverberation time of room, s
T_{sample}	sampling period $=1/F_{\mathrm{sample}}$, s
V	volume of room or enclosure, m^3
W	vector representing coefficients of the adaptive filter
W^0	vector representing coefficients of the adaptive filter in the adapted state
x	reference signal (difference of microphone signals)
X	vector of last N values of signal x
y	output of the adaptive filter

Greek

α_n	azimuth of noise source
α_s	azimuth of target signal source
Δ	delay in target signal path between d' and d, in samples
ϵ	output signal of the adaptive beamformer
σ_i^2	variance of the ith coefficient in filters A and B
ϑ	angle between point on surface of a rigid sphere and direction of incidence of plane wave
γ_n	index of directionality of the noise source
γ_s	index of directionality of the target signal source

Note: All parameters are dimensionless, unless otherwise noted

I. INTRODUCTION

Many users of cochlear implants and conventional hearing aids complain about insufficient intelligibility of speech in noisy situations, even if the performance of their aid is satisfactory in quiet environments (Kochkin, 1993). As many hearing impaired listeners need significantly higher signal-to-noise ratios (SNR) for satisfactory communication than normal hearing listeners (Lurquin and Rafhay, 1996; Valente, 1998), numerous noise reduction methods for hearing aids and cochlear implants have been proposed (Lim and Oppenheim, 1979; Graupe *et al.*, 1987; Soede *et al.*, 1993; Bächler and Vonlanthen, 1995; Whitmal *et al.*, 1996; Vanden Berghe and Wouters, 1998). Some of the most promising noise reduction schemes assume that target signals are emitted in front of the listener, while signals arriving from other directions are considered to be noise (Peterson *et al.*, 1987; Soede *et al.*, 1993; Bächler and Vonlanthen, 1995). Directional noise reduction methods have been shown to improve SNR and to be of practical use for the hard-of-hearing (Peterson *et al.*, 1987; Greenberg and Zurek, 1992; Kompis and Dillier, 1994; Valente *et al.*, 1995; Kochkin, 1996; Cochlear Inc., 1997; Gravel *et al.*, 1999; Wouters *et al.*, 1999). Several methods are known to achieve spatial directionality. Besides the use of directional microphones, the output signals of several (omnidirectional or directional) microphones can be postprocessed using either fixed or adaptive postprocessing (Soede *et al.*, 1993; Kompis, 1998). In fixed postprocessing, all transfer functions between the microphone signals and the output are time independent. In adaptive postprocessing, the coefficients of at least one filter are continuously adjusted to optimize noise reduction in the given environment. In general, adaptive beamformers achieve higher noise reductions at the expense of higher computational loads and greater system complexity (DeBrunner and McKinney, 1995; Kates and Weiss, 1996; Kompis *et al.*, 1999; Kompis *et al.*, 2000).

While fixed beamformers have been theoretically analyzed and the achievable noise reduction can be predicted based on these theoretical considerations (Cox *et al.*, 1986; Stadler and Rabinowitz, 1993), predictions of the performance of adaptive systems are rare (Widrow *et al.*, 1975; DeBrunner and McKinney, 1995). To date, they do not take into account the length of the adaptive filter and reverberation time of the environment, two factors which have been found to be of major importance (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Kompis and Dillier, 1991; Greenberg and Zurek, 1992; Dillier *et al.*, 1993). Most reports on adaptive beamformer applications provide experimental data using either speech recognition tests with normal hearing or hearing impaired listeners (Peterson *et al.*, 1987; Kompis and Dillier, 1994; van Hoesel and Clark, 1995; Hamacher *et al.*, 1996; Welker *et al.*, 1997) or different measures related to signal-to-noise ratio improvement (Greenberg and Zurek, 1992; Greenberg *et al.*, 1993; Dillier *et al.*, 1993; Welker *et al.*, 1997; Kates, 1997). It is difficult to compare the results of these reports because of the numerous differences in the experimental setting, such as reverberation time, directionality of sound sources or filter adaptation. The effect of each difference is hard to estimate because of the lack of a theoretical background or sufficient experimental data. In this report, the noise reduction that can be achieved by a two-microphone adaptive beamformer (Griffiths and Jim, 1982; Peterson *et al.*, 1987) is analyzed. An approximate method to predict its noise reduction as a function of the design parameters of the beamformer and the acoustic parameters of the acoustic environment including the sound sources is derived. In Sec. II, the investigated adaptive beamformer is defined. In Sec. III, the assumptions for the theoretical analysis are discussed. Models of the impulse responses between the acoustic noise sources and the beamformer are presented in Sec. IV, and in Secs. V and VI, an approximation to predict the achievable improvement in signal-to-noise ratio is derived. Potential applications and limitations of the presented method to predict SNR improvements are discussed in Sec. VII. A short FORTRAN subroutine which performs the calculation to predict SNR improvement is included in the Appendix. Experimental verification of the predictions is provided in a companion paper (Kompis and Dillier, 2001).

II. THE ADAPTIVE BEAMFORMER

Figure 1 shows a schematic diagram of the two-microphone adaptive beamformer (Griffiths and Jim, 1982;

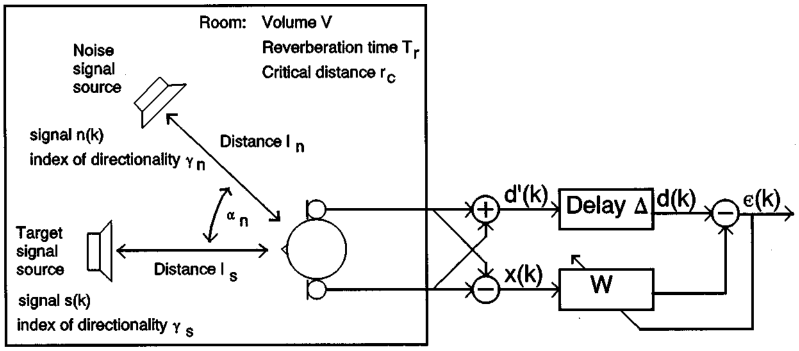

FIG. 1. Schematic diagram of the adaptive beamformer in the acoustic environment used to predict SNR improvements.

Peterson *et al.*, 1987) considered in this research. Note that some researchers prefer the term *Griffiths–Jim beamformer* to describe the same system.

Two omnidirectional microphones are mounted close to the ears of a user. The sum and the difference of the two microphone signals is calculated first. As the target signal source is assumed to lie in front of the listener, the sum d' will contain predominantly target signal, while the difference signal x will contain mainly noise, as noise is assumed to arrive from other directions. A finite-impulse response structured adaptive filter W transforms x in such a way that it can serve as a model of the remaining noise in d. The resulting signal y can then be directly subtracted from d, yielding the output ϵ. The coefficients of the adaptive filter are updated by a least-mean-squares (LMS) algorithm (Widrow *et al.*, 1975) which minimizes the total variance of the output signal. The LMS algorithm relies on the assumption that target and noise signals are uncorrelated. The delay in the target signal path between d' and d can be adjusted to optimize noise reduction. Typically, the length of the adaptive filter is chosen in the range of 10–50 ms, and delay is set to 25%–50% of the filter length (Peterson *et al.*, 1987; Kompis and Dillier, 1991; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; Kompis and Dillier, 1994).

The adaptive beamformer minimizes the variance of any signal of which a—possibly linearly transformed—copy is present in the reference signal x. Due to reverberation and misalignment of the target signal source with respect to the microphones, in most practical situations a part of the target signal will be present in the reference signal x. To prevent target signal cancellation, several algorithms, which stop filter adaptation when a target signal is detected, have been proposed (Van Compernolle, 1990; Greenberg and Zurek, 1992; Kompis and Dillier, 1994; van Hoesel and Clark, 1995; Kompis *et al.*, 1997). Using one of these algorithms, filter adaptation is limited to time segments in which no target signal is present, e.g., the numerous short pauses that occur in the running speech of a target speaker.

III. MODEL ASSUMPTIONS

To predict the SNR improvement that can be achieved by the adaptive beamformer, a simplified model of the acoustic setting is assumed as follows (cf. the left-hand side of Fig. 1 for a graphic representation). A listener in a reverberant room faces a single target signal source. A second acoustic source, emitting the noise signal, is placed at an azimuth α_n from the listener, where α_n is large enough to give rise to a difference in the time of arrival of the noise signal between the two microphones of at least one sampling period T_{sample}. No movement of either the listener or the sound sources is allowed. The directionality of the acoustic sources is described by the index of directionality γ_n for the noise source and γ_s for the target signal source, defined as the ratio between the signal intensity emitted in the direction of the listener to the intensity of a hypothetical omnidirectional source with the same total acoustic output power (DeBrunner and McKinney, 1995). The head of the listener is modeled as a rigid sphere of 9.3 cm in radius, as proposed by Kuhn (1977) and used in an earlier study (Kompis and Dillier, 1993). Two omnidirectional microphones are mounted on the surface of the rigid sphere opposite each other, serving as inputs to the adaptive beamformer. The acoustic properties of the room are defined by any two of the three parameters volume V, reverberation time T_r, and critical distance r_c. Reverberation time is defined as the time required for the reverberant signal to decay by 60 dB. The critical distance is defined as the distance from an omnidirectional acoustic source at which the direct-to-reverberant ratio is 1. The relationship between these parameters can be approximated by

$$r_c \approx \sqrt{\frac{6 \ln 10}{4\pi c}\,\frac{V}{T_r}}, \tag{1}$$

where c is the sound speed (Zwicker and Zollner, 1984). For the calculations in the Appendix, a sound speed of $c = 340\,\text{m/s}$ is assumed. Both the noise and the target signal source are assumed to emit white noise, with the signals of the two sources being uncorrelated. The adaptive beamformer processing the two microphone signals is configured as shown in Fig. 1 and defined by its sampling rate F_{sample}, the number of coefficients N of the adaptive filter, and the number of samples Δ of delay in the target signal path between d' and d. A perfectly adapted filter is assumed, i.e., it is assumed that filter adaptation took place in the absence of the target signal and the coefficients of the adaptive filter have converged to their optimal state. The state of the adaptive filter is assumed to be frozen at the end of adaptation, so that only the noise signal, but not the target signal, has had an influence on the filter coefficients.

In principle, no restrictions are imposed by the model on the variances of either the noise or the target signal. However, in order to simplify calculations and without loss of generality, it is assumed that the variance of the noise signal $n(k)$ equals 1, and room transfer functions are scaled in such a way as to let the variances of the noise signal equal 1 in both the sum signal $d(k)$ and the difference signal $x(k)$. Similarly, i.e., in order to simplify calculations and without loss of generality, the variance of the reverberant portion of the target signal at either microphone is assumed to be 1. Clearly, some of the above-mentioned assumptions are more limiting than others. The assumptions on the variances of the target and noise signals exclude situations without any reverberation. To generate a difference of at least one sampling period, at a sampling rate of, e.g., $F_{\text{sample}}=10\,\text{kHz}$, the minimum azimuth of the noise source must be roughly 10°, which does not seriously limit general applicability. The model requires that the signals of both acoustic sources are white noise. Furthermore, effects of the frequency dependence of the acoustic diffraction by the head of the listener of the directionality of the sound sources are not taken into account. While this is clearly unrealistic in light of the predominantly low-frequency speech and noise sounds, which are to be expected as input signals in a hearing aid application, this assumption becomes more acceptable when considering that the most frequently used adaptation algorithm, the LMS algorithm (Widrow *et al.*, 1975), minimizes total signal variance, i.e., the spectral components of a noise signal are reduced according to their relative power. Therefore, in numerous realizations of the adaptive beamformer, microphone signals are prewhitened by usually 6 dB per octave to account for the importance of the spectral components with respect to speech intelligibility (Peterson *et al.*, 1987; Dillier *et al.*, 1993; Kompis and Dillier, 1994; Welker *et al.*, 1997). Usually, changes introduced by these pre-emphasis filters are compensated by a de-emphasizing filter in the output path of the adaptive beamformer (Kompis, 1998). With these provisions, the spectra of the practically important speech signals actually being processed by the beamforming algorithm approach the white spectra of the model. Although it can be shown that broadband SNR improvement corresponds closely to an intelligibility-weighted measure of speech-to-interference ratio gain (Greenberg *et al.*, 1993) in numerous realistic experimental settings (Kompis and Dillier, 2001), the noninclusion of frequency dependence remains a limitation of the model. In the model of the listener, no pinnae or shoulders are accounted for. This simple model has been verified earlier and seems to be sufficient for a number of hearing aid applications (Kompis and Dillier, 1993). As there are several ways to mount hearing aid microphones with respect to the pinnae, and as the presented model does not generally take into account frequency dependence, the inclusion of pinnae or shoulder effects into the model does not seem to be justified. Again, however, the noninclusion of the alterations in the frequency spectra due to the head of the listener may be a limiting factor for a number of applications.

Although the two assumptions that (a) the filter has been adapted in the absence of the target signal and is (b) perfectly adapted cannot be expected to be met perfectly in real situations, these assumptions are reasonably realistic for many practical applications. Several target-signal detection/adaptation-inhibition algorithms have been proposed and used in experiments (Van Compernolle, 1990; Greenberg and Zurek, 1992; Kompis and Dillier, 1994; van Hoesel and Clark, 1995; Kompis *et al.*, 1997). Using one of these algorithms, it can be assumed that the target signal does not significantly influence filter adaptation and filter adaptation takes place in the presence of the noise signal only (Kompis *et al.*, 1997). At filter lengths of 10–50 ms, which are usually used for adaptive beamformers, short adaptation time constants on the order of magnitude of 0.1 s (Dillier *et al.*, 1993; Kompis and Dillier, 1994) can be combined with small convergence errors. Therefore, the coefficients of the adaptive filter can be reasonably expected to have converged, e.g., during the short pauses between the first words of an utterance of a target speaker.

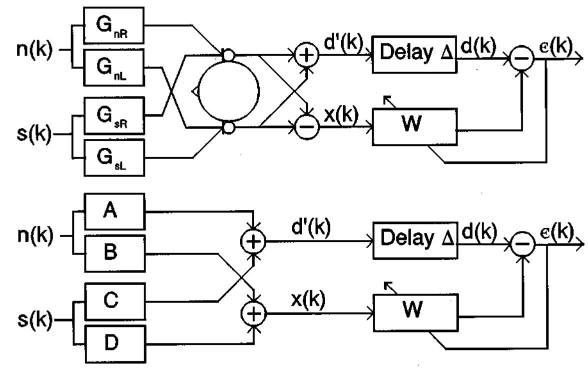

FIG. 2. Relationship between the transfer functions G_{nR}, G_{nL}, G_{sR}, and G_{sL} and *A, B, C,* and *D.*

IV. MODELING OF THE IMPULSE RESPONSES BETWEEN THE ACOUSTIC SOURCES AND THE MICROPHONES

The transfer functions between the two acoustic sources and the two microphones can be modeled as impulse responses G_{nR}, G_{nL}, G_{sR}, and G_{sL}, respectively. The first subscript (n or s) marks the source (noise or target signal), the second subscript (L or R) marks the left or right microphone. These impulse responses account for all effects of source directionality, room reverberation, and sound diffraction by the listener's head. For the analysis in Sec. V, it is convenient to convert these impulse responses into four slightly different impulse responses *A*, *B*, *C*, and *D* as follows:

$$\begin{aligned}
A&=G_{nR}+G_{nL}=(a_0,a_1,a_2,\dots),\\
B&=G_{nR}-G_{nL}=(b_0,b_1,b_2,\dots),\\
C&=G_{sR}+G_{sL},\\
D&=G_{sR}-G_{sL}.
\end{aligned}\tag{2}$$

Using this definition, the calculation of the sum and difference of the microphone signals at the first stage of the adaptive beamformer is already included in *A*, *B*, *C*, and *D*, as shown schematically in Fig. 2.

While the impulse responses between the *target* sound source and the microphones do not influence filter adaptation and can therefore be handled in a simplified manner in Sec.

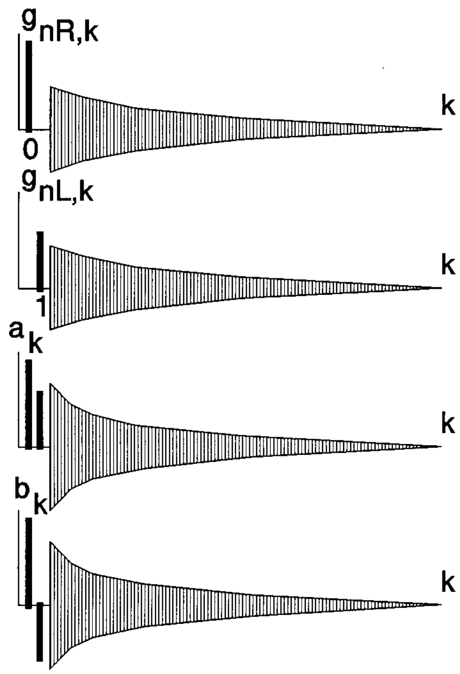

FIG. 3. Schematic representation of the model-transfer functions G_{nR}, G_{nL}, *A*, and *B* between noise source and the adaptive beamformer. The solid lines represent the directly incident portions of the noise signal, hatched areas represent the reverberant response.

VI, a more detailed model of the impulse responses between the *noise* source and the beamformer (i.e., G_{nR}, G_{nL}, *A*, and *B*) is required. These impulse responses are modeled by adding the direct response of the microphone which is closer to the noise source in coefficient 0, the direct response to the microphone farther away from it in coefficient 1, and the reverberation in coefficients 2 through ∞, as depicted in Fig. 3. In general, the difference in the time of arrival between the two microphones will not be exactly one sampling period T_{sample} as modeled, but usually larger, e.g., four samples at a sampling rate of $F_{\mathrm{sample}}=10\,\mathrm{kHz}$ and an azimuth of $\alpha_n=45°$ (differences smaller than T_{sample} are excluded by the model definitions in Sec. III). It was found that larger differences are negligible as long as the adaptive filter is much longer than the difference in the time of arrival. In most practical applications, filters are 10–100 times longer than the time-of-arrival difference of the noise sound and this prerequisite is met.

The size of the first two coefficients is a function of the angle of incidence of the direct, nonreverberated portion of the noise signal. The total rms value of a white noise signal at a point on the surface of a rigid sphere at an angle ϑ with respect to the angle of incidence and relative to the root-mean-square value of the same white noise in free field can be calculated from the formulas provided, e.g., by Schwarz (1943) or Morse (1983). Figure 4 shows the resulting function $S(\vartheta)$ for a rigid sphere with a radius of 9.3 cm for three different frequency bands of 0–2.5, 0–5, and 0–10 kHz, corresponding to sampling rates of 5, 10, and 20 kHz, if ideal nonaliasing filters are assumed. The differences between the three curves arise because of the more pronounced diffraction of the high frequency components of the signals.

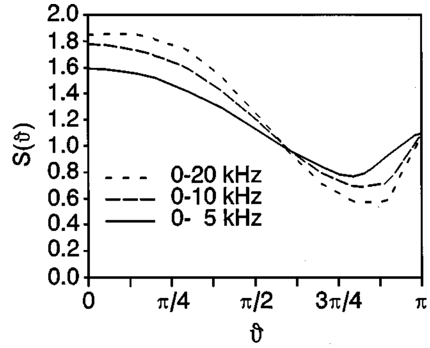

FIG. 4. Sound pressure at the surface of a head-sized ($r=9.3\,\mathrm{cm}$) rigid sphere as a function of the angle of sound incidence ϑ. $S(\vartheta)$ represents rms values relative to free field, for white noise processed by three different low-pass filters.

Using $S(\vartheta)$, the first two coefficients of *A* and *B* can be written as

$$\begin{aligned} a_0&=b_0=G_0=S(\pi/2-\alpha_n)F_d,\\ a_1&=-b_1=G_1=S(\pi/2+\alpha_n)/F_d, \end{aligned} \tag{3}$$

where F_d is a constant, the value of which will be determined shortly to account for the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise signal. All other coefficients, i.e., $a_i, b_i, i \geqslant 2$, representing the reverberant part of the room filter are modeled as a series of independent, normally distributed random variables, where

$$\begin{aligned} E\{a_i a_j\}&=E\{b_i b_j\}=\begin{cases} 0, & i\neq j\\ \sigma_i^2, & i=j \end{cases},\\ E\{a_i b_j\}&=0 \end{aligned} \tag{4}$$

holds for all *i* and *j*. Note that for any given acoustic setting, *A* and *B* are linear impulse responses with fixed, well-defined and time-independent values a_i and b_i for all *i*. However, as the exact values of every a_i and b_i for the reverberant part $(i\geqslant 2)$ are neither known nor required for the following computation, only some relevant statistical properties of the coefficients are used. Nevertheless, the underlying impulse responses are time invariant and linear. The variance σ_i^2 decreases exponentially with the index *i* as the reverberant portion of the signal decays exponentially:

$$\sigma_i^2=F_\sigma e^{-i/T}, \tag{5}$$

where F_σ is another newly introduced coefficient to account for the correct direct-to-reverberant ratio and *T* is a time constant (dimensionless, in multiples of the sampling period T_{sample}).

To complete the model of the impulse responses *A* and *B*, the three newly introduced variables *T*, F_d, and F_σ must be calculated first. To derive the value of the dimensionless time constant *T* from the reverberation time T_r and the sampling period T_{sample}, the definition of the reverberation time (i.e., time required for the reverberant signal to decay by 60 dB) can be used:

$$e^{-T_r/TT_{\mathrm{sample}}}=10^{-60/10} \tag{6}$$

from which *T* can be calculated as

$$T = \frac{T_r F_{\text{sample}}}{6 \ln(10)}. \tag{7}$$

The direct-to-reverberant ratio $P_{\text{d/r}}$ of the noise signal at the location of the listener can be estimated as

$$P_{\text{d/r}} = \left(\frac{r_c}{l_n}\right)^2 \gamma_n. \tag{8}$$

Using the two coefficients F_d and F_σ, it is possible to adjust the direct-to-reverberant ratio $P_{\text{d/r}}$ correctly,

$$\frac{G_0^2 + G_1^2}{\Sigma_{i=2}^{\infty} \sigma_i^2} = P_{\text{d/r}}, \tag{9}$$

and at the same time guarantee that

$$\sum_{i=0}^{\infty} a_i^2 = \sum_{i=0}^{\infty} b_i^2 = G_0^2 + G_1^2 + \sum_{i=2}^{\infty} \sigma_1^2 = 1 \tag{10}$$

as stated in Sec. III in order to keep calculations in the following sections as simple as possible. Using the identity

$$\sum_{i=M}^{N} e^{-i/T} = \frac{e^{-M/T} - e^{-(N+1)/T}}{1 - e^{-1/T}} \tag{11}$$

it can be found that

$$F_d = \sqrt{\frac{P_{\text{d/r}}}{(1+P_{\text{d/r}})(S^2(\pi/2 - \alpha_n) + S^2(\pi/2 + \alpha_n))}}, \tag{12}$$

$$F_\sigma = \frac{1 - e^{-1/T}}{(1+P_{\text{d/r}}) e^{-2/T}}. \tag{13}$$

V. APPROXIMATE SOLUTION FOR THE AMOUNT OF NOISE SUPPRESSION BY THE ADAPTIVE FILTER

In this section, an approximate solution for the amount of noise reduction h provided by the adaptive filter, defined as

$$h = E\{d^2\} - E\{\epsilon^2\}, \tag{14}$$

is derived. The noise reduction h for an ideally adapted filter can be calculated analytically if the delayed sum signal $d(k)$ and the reference signal $x(k)$ are known. The derivation of the corresponding equations can be found in standard textbooks (e.g., Widrow and Stearns, 1985) on adaptive filters and is not repeated here. To calculate the *approximate* noise reduction for the problem of the adaptive beamformer in a reverberant room, the following definitions are needed. Let X be a vector of the last N samples in the reference signal x, where N is the number of coefficients in the adaptive filter. Then an autocorrelation matrix R can be defined as

$$R = E\{X \cdot X^T\}, \tag{15}$$

where the superscript T stands for transposition and $E\{\ \}$ denotes the expected value over time. Similarly, let the cross-correlation vector P be

$$P = E\{X \cdot d\} = \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_{N-1} \end{bmatrix}. \tag{16}$$

Using these definitions, the vector W^0 containing the N filter coefficients of the ideally adapted filter for which the variance of the output signal $E\{\epsilon^2(k)\}$ becomes minimal can be written as

$$W^0 = R^{-1} P. \tag{17}$$

The noise reduction h can then be expressed as

$$h = E\{d^2\} - E\{\epsilon^2\} = W^{0T} P = P^T R^{-1} P. \tag{18}$$

For the investigated problem, signals x and d are not known. However, as the source signal n is known to be white noise signal with variance 1, the samples of n are known to be statistically independent. Using the coefficients a_i and b_i of the impulse responses A and B, the elements of the cross-correlation vector P can then be written as

$$P_i = \sum_{k=\max(0,\Delta-i)}^{\infty} a_k \cdot b_{k-i+\Delta}. \tag{19}$$

As long as the samples of the noise signal remain statistically independent in the reference signal $x(k)$, i.e., *after* modification by the impulse response B, the autocorrelation matrix R can be approximated by the identity matrix I,

$$R \approx I. \tag{20}$$

However, this approximation is reasonably accurate only for low direct-to-reverberant ratios $P_{\text{d/r}}$ of the noise signal, where the statistically independent coefficients of the reverberant response dominate the impulse response B. At high direct-to-reverberant ratios of the noise signal, B and therefore $x(k)$ are dominated by the directly incident noise portions and the assumption of statistically independent samples $x(k)$ is violated. It can be shown (Kompis and Dillier, 2001) that the given approximation is reasonably accurate for direct-to-reverberant ratios of the noise signal $P_{\text{d/r}} < +3$ dB. Using this approximation and Eqs. (18)+(19), the noise reduction h could be calculated if all model coefficients a_i and b_i were explicitly known. Except for a_0, a_1, b_0, and b_1 however, only the expected value, which is zero, and the expected variance, which is σ_i^2, are known. Therefore h cannot be calculated, but its expected value $E\{h\}$ can be approximated by

$$E\{h\} = E\{P^T R^{-1} P\} \approx \sum_{i=0}^{N-1} E\{P_i^2\}. \tag{21}$$

There is a meaningful interpretation of this equation. To simplify the discussion, let the delay in the target signal path Δ equal zero for this paragraph only. Equation (21) shows that in order to calculate the expected value of the noise reduction h, N positive values $E\{P_i^2\}$ are summed, thus increasing the noise reduction h with the length of the adaptive filter [P_i can never equal zero because of Eq. (19)]. Using the program in the Appendix it can even be shown that $E\{h\}$ approaches 1 (i.e., perfect noise cancellation) for any reverberation time with increasing filter lengths N, as long as the directly incident portion of the sound remains negligible. From the schematic representation of the impulse responses A and B and the definition of P_i in Eq. (19) it can be seen that in environments with short reverberation times T_r, only the first few coefficients a_i and b_i will contribute significantly to P_i^2, and P_i^2 will therefore only contribute significantly to $E\{h\}$ for small values of the index i. Calculating the contribution of the terms with large values of the index i is equivalent to shifting the impulse responses A and B significantly with respect to each other before multiplying and summing the corresponding coefficients in Eq. (19). Therefore, in situations with short reverberation times, after the first few terms in Eq. (21), $E\{h\}$ will increase only very slowly with N, meaning that already short adaptive filters can significantly reduce noise. For long reverberation times, the reverberant tails in Fig. 3 become long as well, but the first few coefficients a_i and b_i are smaller than for short reverberation times because of Eq. (10). This means that the contribution of the first few of the N filter coefficients of the adaptive beamformer are smaller than at short reverberation times, but the increase in noise reduction of the $N+1$st coefficient of the adaptive filter is larger for large N and longer filters will be needed to reach the same amount of noise reduction. At high direct-to-reverberant ratios $P_{\mathrm{d/r}}$ of the noise source, the first two coefficients in A and B (a_0, a_1, b_0, and b_1) representing the direct response are large, and the effect is similar to that of shortening reverberation time. Because of the approximation [Eq. (20)] used, Eq. (21) is only valid if the $P_{\mathrm{d/r}}$ is small, i.e., less than approximately $+3$ dB (Kompis and Dillier, 2001). This is a new assumption which was not discussed in Sec. III and which limits the range of applicability of the given analysis. As a consequence, achievable gains in signal-no-noise ratio will be underestimated for situations with high direct-to-reverberant ratios of the noise source. Consequences will be discussed in Sec. VII.

To estimate $E\{h\}$, each of the N terms of the sum in Eq. (23) must be calculated first. Each term is itself a sum, which can be conveniently split into three terms as follows:

$$\begin{aligned}
E\{P_i^2\} &= E\left\{\left(\sum_{k=\max(0,\Delta-i)}^{\infty} a_k \cdot b_{k-i+\Delta}\right)^2\right\} \\
&= E\left\{\left(\sum_{k=\max(0,\Delta-i)}^{\infty} a_k \cdot b_{k-i+\Delta}\Big|_{\substack{k<2\wedge\\ k-i+\Delta<2}}\right)^2\right\} \\
&\quad + E\left\{\left(\sum_{k=\max(0,\Delta-i)}^{\infty} a_k \cdot b_{k-i+\Delta}\Big|_{\substack{k<2\oplus\\ k-i+\Delta<2}}\right)^2\right\} \\
&\quad + E\left\{\left(\sum_{k=\max(0,\Delta-i)}^{\infty} a_k \cdot b_{k-i+\Delta}\Big|_{\substack{k\geqslant 2\wedge\\ k-i+\Delta\geqslant 2}}\right)^2\right\} \\
&= \varphi_{\mathrm{dd}}(i) + \varphi_{\mathrm{dr}}(i) + \varphi_{\mathrm{rr}}(i).
\end{aligned} \tag{22}$$

The three portions cover the terms concerning the directly incident portion of the noise only (φ_{dd}), the terms concerning the reverberant terms only (φ_{rr}), and the mixed terms (φ_{dr}). As a_0, a_1, b_0, and b_1 are explicitly known from Eq. (4), φ_{dd} can be directly calculated as follows:

$$\varphi_{\mathrm{dd}}(i) = \begin{cases} (G_0^2 - G_1^2)^2, & |i-\Delta| = 0 \\ (G_0 \cdot G_1)^2, & |i-\Delta| = 1 \\ 0, & |i-\Delta| \geqslant 2. \end{cases} \tag{23}$$

For the mixed term φ_{dr} the properties

$$\begin{aligned}
E\{(\xi_1+\xi_2)^2\} &= E\{\xi_1^2\} + E\{\xi_2^2\}, \\
E\{(\xi_1\cdot\xi_2)^2\} &= E\{\xi_1^2\} \cdot E\{\xi_2^2\},
\end{aligned} \tag{24}$$

of any two independent random variables ξ_1 and ξ_2 can be used, as all a_i and b_i are independent of each other for $i \geqslant 2$ and independent from G_0 and G_1. The result yields

$$\begin{aligned}
\varphi_{\mathrm{dr}}(i) &= \sum_{k=\max(0,\Delta-i)}^{\infty} E\{a_k^2\} \cdot E\{b_{k-i+\Delta}^2\}\Big|_{\substack{k<2\oplus\\ k-i+\Delta<2}} \\
&= \begin{cases} 0, & |i-\Delta| = 0 \\ G_1^2 \cdot \sigma_{|i-\Delta|+1}^2, & |i-\Delta| = 1 \\ G_0^2 \cdot \sigma_{|i-\Delta|}^2 + G_1^2 \cdot \sigma_{|i-\Delta|+1}^2, & |i-\Delta| \geqslant 2. \end{cases}
\end{aligned} \tag{25}$$

Similarly, using Eq. (11), the reverberant term φ_{rr} can be calculated as

$$\begin{aligned}
\varphi_{\mathrm{rr}}(i) &= \sum_{k=\max(0,\Delta-i)}^{\infty} E\{a_k^2\} \cdot E\{b_{k-i+\Delta}^2\}\Big|_{\substack{k\geqslant 2\wedge\\ k-i+\Delta\geqslant 2}} \\
&= \sum_{k=2}^{\infty} \sigma_k^2 \cdot \sigma_{k+|\Delta-i|}^2 \\
&= \frac{F_\sigma^2 \exp\left(-\dfrac{4+|\Delta-i|}{T}\right)}{1-e^{-2/T}}.
\end{aligned} \tag{26}$$

By substituting Eqs. (23), (25), and (26) into Eq. (22), using Eq. (21) an approximation for $E\{h\}$ can now be calculated.

VI. IMPROVEMENT OF THE SIGNAL-TO-NOISE RATIO

To estimate SNR improvement, the level of the target signal and of the noise signal will be compared at the following four different points of the signal processing chain (cf. Fig. 1) of the adaptive beamformer: (i) at the microphone with the less favorable SNR lying closer to the noise source (index 1), (ii) at the microphone with the more favorable SNR lying farther away from the noise source (index 2), (iii) after summation of both microphone signals, i.e., signal d' in Fig. 1 (index S), and (iv) at the output of the adaptive beamformer, i.e., signal ϵ in Fig. 1 (index B). By calculating the SNRs in those four signals, the SNR improvement of the adaptive beamformer can be related to either microphone signal or to the SNR gain of a simple fixed two-microphone beamformer (Kompis and Diller, 1994), in which both microphone signals are summed.

To calculate the level of the target signal in these four signals, the direct-to-reverberant ratio of the target signal $Q_{\mathrm{d/r}}$ at the location of the listener can be estimated—in analogy to Eq. (8)—as

$$Q_{\mathrm{d/r}} = \left(\frac{r_c}{l_s}\right)^2 \cdot \gamma_s . \tag{27}$$

As discussed in Sec. V, reverberation must be present for the approximation (20) to be valid. Without loss of generality, the variance of the reverberant portion of the target signal can therefore be set to 1, and the total variance (i.e., including the direct *and* reverberant portions) of the target signal in the two microphones becomes

$$S_1 = S_2 = 1 + Q_{\mathrm{d/r}} . \tag{28}$$

By adding both microphone signals, which corresponds to the signal processing of a part of the front end of the adaptive beamformer, the *variance* of the (uncorrelated) reverberant portion is doubled, while, assuming perfect alignment of the target source, the *amplitude* of the direct portion of the sound is doubled, and therefore its variance is multiplied by a factor of 4. However, this is only true for perfect alignment of the target signal source with respect to the microphones. In a realistic setting, e.g., for head-sized spacing between the microphones and for a sampling rate of, e.g., $F_{\mathrm{sample}} = 10\,\mathrm{kHz}$, this is valid for azimuths of the target signal source $\alpha_S = -3° \ldots +3°$. If the misalignment gives rise to a time difference of more than approximately T_{sample}, which in the above-mentioned example occurs at $\alpha_S > 10°$, uncorrelated samples of the white noise signal will add up and the variance of the direct portion of the signal is only doubled. To account for this effect, an alignment factor A is introduced, which can be assessed experimentally in anechoic environments and will, for white noise, yield values in the range of 4 (perfect alignment) down to approximately 2 (no alignment). The variance of the target signal portion in the sum d' can thereby be written as

$$S_S = 2 + A \cdot Q_{\mathrm{d/r}} . \tag{29}$$

Similarly, the variance of the target signal in the reference path x becomes

$$S_D = 2 + (4 - A) \cdot Q_{\mathrm{d/r}} . \tag{30}$$

As, according to the model assumptions, noise and target signal are uncorrelated and as the filter W was adapted in the absence of the target signal, the variance of the target signal portion in the reference signal x will increase by the factor of $W^{0T} W^0$ at the output of the adaptive filter (signal y). Using Eq. (17) and approximations (20) and (21), this factor can be shown to be equal to $E\{h\}$. The variance of the target signal at the output ϵ of the adaptive beamformer can now be written as the sum of the variances of the filtered reference signal y and the delayed sum signal d,

$$S_B = S_S + E\{h\} \cdot S_D . \tag{31}$$

So much for the target signal. As to the signal of the noise source, its variance in the sum signal d' can be set to 1 without loss of generality:

$$N_S = 1. \tag{32}$$

The variance of the noise signal at the output of the beamformer can then be written as

$$N_B = 1 - E\{h\}. \tag{33}$$

The variance of the reverberant portion of the noise in the microphone signals is on average $\frac{1}{2}$ of that of the sum signal, the direct portion of the noise is not changed, thus

$$N_1 = \frac{1}{2} \cdot \frac{1}{P_{\mathrm{d/r}} + 1} + G_0^2, \quad N_2 = \frac{1}{2} \cdot \frac{1}{P_{\mathrm{d/r}} + 1} + G_1^2. \tag{34}$$

Now the improvement in SNR at the output of the adaptive beamformer, when compared to the SNR the microphone with the less favorable SNR (V_1), to the microphone with the more favorable SNR (V_2), or when compared to the two microphone fixed beamformer (V_S) can be calculated as follows:

$$\begin{aligned} V_1 &= 10 \log_{10} \frac{S_B \cdot N_1}{N_B \cdot S_1}, \\ V_2 &= 10 \log_{10} \frac{S_B \cdot N_2}{N_B \cdot S_2}, \\ V_S &= 10 \log_{10} \frac{S_B \cdot N_S}{N_B \cdot S_S}. \end{aligned} \tag{35}$$

The FORTRAN subroutine provided in the Appendix performs all computations necessary to determine all three SNR improvements in Eq. (35).

VII. DISCUSSION

The presented procedure used to estimate the SNR improvement of an adaptive beamformer in the given model setting is based on a number of assumptions and approximations. Its applications are therefore limited. A set of underlying assumptions have been listed and discussed in Sec. III. One additional limitation concerning the range of validity of the predictions is not listed in Sec. III, as it is not a consequence of the underlying model but rather of the approximation used in Eq. (20). For this approximation to be applicable, the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise source must be small, as stated in Sec. V. This limits the predictions to situations with at least a small level of reverberation. It can be shown experimentally (Kompis and Dillier, 2001) that, for realistic sets of parameter values, it is sufficient for $P_{\mathrm{d/r}}$ to be below approximately $+3$ dB for reasonably accurate predictions. For higher $P_{\mathrm{d/r}}$, SNR improvement will be systematically underestimated. However, for many applications, this is not a serious limitation. As the model is limited to low direct-to-reverberant ratios of the *noise* source only, predictions for high direct-to-reverberant ratios of the *target* signal source are not affected by this limitation. Although as a side effect of the precedence effect it may not always be easy to appreciate the amount of reverberation subjectively, in many acoustic settings in rooms with realistic amounts of reverberation direct-to-reverberant ratios are below $+3$ dB even at distances well below 1 m (Kompis and Dillier, 2001), and users of the system will probably tend to keep away from disturbing noise sources, thus further decreasing direct-to-reverberant ratio. Mainly in anechoic environments, however, where the adaptive beamformer is known for its excellent performance (Peterson *et al.*, 1987), the presented method does not adequately predict SNR improvement.

For hearing aid applications, the primary goal is improved speech intelligibility and not improved SNR, as predicted by the presented method. Because some frequency bands contribute more to speech intelligibility than others, SNR improvement may correlate poorly with improvement in speech recognition, if substantial differences between SNR improvements in different frequency bands exist. However, it can be shown that in the present context, SNR and intelligibility-weighted gain (Greenberg *et al.*, 1993) agree reasonably for a wide range of relevant experimental conditions (Kompis and Dillier, 2001).

The validation of the predicted SNR improvements is of major importance. Validation of the prediction procedure by comparisons to published experimental data is complicated by several factors. Comparisons are limited to experiments which meet or at least approach the model assumptions listed in Sec. III. Comparisons are not possible if different numbers or arrangements of microphones or several noise sources are used (e.g., Peterson *et al.*, 1990; Greenberg and Zurek, 1992). As the proposed prediction method is limited to reverberant conditions, comparisons with experiments in anechoic environments (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Greenberg and Zurek, 1992) are not meaningful. Some of the results reported in the literature list the improvement in terms of speech recognition scores rather than SNR improvement, and in some instances it is not possible to extract the latter information from these data (Kompis and Dillier, 1994; van Hoesel and Clark, 1995). In some reports (van Hoesel and Clark, 1995; Hamacher *et al.*, 1996), no data on the directionality of the sound sources are given. Directionality of the sound sources are required input parameters to calculate the predicted SNR improvement using the presented method. For these reasons, a series of 92 experiments using the adaptive beamformer was performed and experimental results were compared to the predicted SNR improvements. These data are reported separately (Kompis and Dillier, 2001).

Despite some limitations, the presented prediction method offers several advantages over actual experiments in real or simulated environments. Results for a wide range of acoustic settings can be obtained in a fraction of the time required for actual experiments. Results are substantially less prone to errors and problems in the experimental setting such as programming errors, inadvertently wrong entry of simulation data, wiring or microphone problems, etc. Furthermore, predictions are not influenced by technical limitations of experimental settings such as limited resolution of analog-to-digital converters, nonideal adaptation of the adaptive filter, effects of electrical or acoustic noise, etc. Therefore, the predictions offer a unique method to differentiate between implementational and/or experimental limitations and limitations of the adaptive beamforming method per se. Even if the prediction method is not used, it may be helpful for experiments by providing a list of parameters which have to be controlled in every experiment.

The presented prediction method cannot be expected to replace experiments completely, but experiments and predictions can complement each other favorably. One potential application of the presented algorithm is to enable a validation of experimental data, e.g., if experimental results are either unexpectedly favorable or unexpectedly poor. If the predictions are sufficiently verified experimentally, many time-consuming experiments can be even omitted completely in the early stages of the development of a practical adaptive beamforming noise reduction system.

Probably the most interesting application is the study of the complex behavior of the adaptive beamformer in a wide variety of acoustic situations within a reasonable time span. A first effort in this direction is presented in a companion paper (Kompis and Dillier, 2001).

Because of the numerous underlying assumptions and the approximation used, there is considerable room for improvement for the presented prediction algorithm. Extension to situations with higher $P_{\mathrm{d/r}}$, to frequency-dependent predictions of the SNR improvement, or extensions to cases using other numbers or arrangements of microphones (Peterson *et al.*, 1990; Greenberg and Zurek, 1992; Kates and Weiss, 1996) or directional microphones (Kompis and Dillier, 1994; DeBrunner and McKinney, 1995) might prove to be very useful.

To perform the relatively complex calculations to predict SNR improvements, a FORTRAN subroutine is provided in the Appendix. FORTRAN was chosen as it is still one of the most widely used programming languages among scientists and engineers (Kornbluh, 1999) and its code can be easily translated to other programming languages.

Despite the above-discussed drawbacks and limitations, the presented method to predict the SNR improvement of adaptive beamformer may be a useful tool in the design and further development of adaptive multimicrophone noise reduction systems for conventional hearing aids and cochlear implants. With its unique possibility to preliminarily evaluate different adaptive beamformers in a wide range of acoustic settings, it may help to point to new directions in research by showing where inherent limitations of the current adaptive beamformer design need to be overcome by innovative concepts.

VIII. SUMMARY

A method to predict the SNR improvement of a two-microphone adaptive beamformer in a reverberant environment has been presented. Predictions are limited to static situations with one noise and one target signal source and perfect adaptation of the adaptive filter is assumed. A FORTRAN subroutine to perform the necessary calculations has been provided. A systematic validation study of the predictions is provided in a separate text (Kompis and Dillier, 2001).

ACKNOWLEDGMENTS

This work was supported by the Swiss National Research Foundation, Grant Nos. 4018-10864 and 3238-56352.99, and Ascom Tech Ltd.

APPENDIX: FORTRAN SUBROUTINE

```
      SUBROUTINE BEAMF(V,Tr,Fls,GammaS,A,Fln,
     1     GammaN,AlphaN,Fsampl,N,ID,V1,V2,Vs)
C Input:
C  V  Volume of room (m3)
C  Tr     Reverberation time Tr (s)
C  Fls    Distance listener - target signal
C          source (m)
C  GammaS Index of directionality of target
C          signal source
C  A      Alignment factor for target signal
C          source (range 2 to 4)
C  Fln    Distance listener-noise source (m)
C  GammaN Index of directionality of target
C          signal source
C  AlphaN Azimuth of noise source (degrees)
C  Fsampl Sampling rate (Hz)
C  N      Number of coefficients in the
C          adaptive filter
C  ID     Delay in target signal path (in
C          multiples of 1/Fsampl)
C Output:
C  V1     SNR-improvement vs.
C          signal of microphone 1 (dB)
C  V2     SNR-improvement vs.
C          signal of microphone 2 (dB)
C  Vs     SNR-improvement vs.
C          sum of microphone signals (dB)
C Array for function S(theta) for 3 sampling
C  rates Fsampl=5,10,and 20kHz
      DIMENSION S(3,11)
      DATA(S(1,I),I=1,4)/1.59,1.57,1.52,1.42/
      DATA(S(1,I),I=5,8)/1.29,1.13,0.97,0.84/
      DATA(S(1,I),I=9,11)/0.76,0.93,1.10/
      DATA(S(2,I),I=1,4)/1.78,1.75,1.69,1.59/
      DATA(S(2,I),I=5,8)/1.42,1.21,0.98,0.79/
      DATA(S(2,I),I=9,11)/0.68,0.72,1.09/
      DATA(S(3,I),I=1,4)/1.85,1.86,1.81,1.73/
      DATA(S(3,I),I=5,8)/1.53,1.26,0.96,0.71/
      DATA(S(3,I),I=9,11)/0.57,0.58,1.06/
      rc=0.057*SQRT(V/Tr) ! critical distance
      Pdr=SQR(rc/Fln)*GammaN
      Qdr=SQR(rc/Fls)*GammaS
      T=Tr*Fsampl/13.815    ! 13.815=6*ln 10
      IFS=2
      IF(Fsampl.LT. 7500.) IFS=1
      IF(Fsampl.GT.15000.) IFS=3
C limit AlphaN to values of 0...90 degrees
C  because of symmetry
      AlphaN=ABS(AlphaN)
      IF(AlphaN.GT.90.) AlphaN=180.-AlphaN
C S(theta) for position of each microphone
C   (ST0,ST1) by interpolation
      RINDEX=(90.-AlphaN)/18.+1.
      INDEX=INT(RINDEX)
      RINDEX=RINDEX-INDEX
      ST0=S(IFS,INDEX)*(1.-RINDEX)+
     1 S(IFS,INDEX+1)*RINDEX
      RINDEX=AlphaN/18.+6.
      INDEX=INT(RINDEX)
      RINDEX=RINDEX-INDEX
      ST1=S(IFS,INDEX)*
     1  (1.-RINDEX)+S(IFS,INDEX+1)*RINDEX
      Fd=SQRT(Pdr/((1.+Pdr)*
     1  (SQR(ST0)+SQR(ST1))))
      Fsigma=(1.-EXP(-1./T))/
     1  ((1+Pdr)*EXP(-2./T))
      G0=Fd*ST0
      G1=Fd*ST1
      h=0.
      DO 100 I=0,N-1
      PHIdd=0.
      IF(IABS(I-ID).EQ.1) PHIdd=SQR(G0*G1)
      IF(I.EQ.ID) PHIdd=SQR(G0*G0-G1*G1)
      PHIdr=0.
      IF(IABS(I-ID).EQ.1) PHIdr=G1*G1*
     1  SIGMA2(IABS(I-ID)+1,Fsigma,T)
      IF(IABS(I-ID).GE.2) PHIdr =
     1  G0*G0*SIGMA2(IABS(I-ID)   ,Fsigma,T)+
     1  G1*G1*SIGMA2(IABS(I-ID)+1,Fsigma,T)
      PHIrr=SQR(Fsigma)*
     1 EXP((-4-IABS(I-ID))/T)/(1.-EXP(-2./T))
      P2=PHIdd+PHIdr+PHIrr
      h=h+P2
100   CONTINUE
      S1=1.+Qdr
      S2=S1
      Ss=2.+A*Qdr
      SD=2.+(4.-A)*Qdr
      SB=Ss+h*SD
      FNs=1.
      FNB=1.-h
      FN1=0.5/(1.+Pdr)+G0*G0
      FN2=0.5/(1.+Pdr)+G1*G1
      V1=10.*ALOG10((SB*FN1)/(FNB*S1))
      V2=10.*ALOG10((SB*FN2)/(FNB*S2))
      Vs=10.*ALOG10((SB*FNs)/(FNB*Ss))
      RETURN
      END
      FUNCTION SIGMA2(I,Fsigma,T)
      SIGMA2=Fsigma*EXP(-I/T)
      RETURN
      END
      FUNCTION SQR(X)
      SQR=X*X
      RETURN
      END
```


Bächler, H., and Vonlanthen, A. (**1995**). ''Audio-Zoom Signalverarbeitung zur besseren Kommunikation im Störschall,'' Phonak Focus **18**, 1–20.

Cochlear, Inc. (**1997**). ''Introducing the Audallion BEAMformer Digital Noise Reduction System,'' Cochlear Clinical Bulletin, April, 1–5.

Cox, H., Zeskind, R. M., and Kooij, T. (**1986**). ''Practical supergain,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-34**, 393–398.

DeBrunner, V. E., and McKinney, E. D. (**1995**). ''Directional adaptive least mean square acoustic array for hearing aid enhancement,'' J. Acoust. Soc. Am. **98**, 437–444.

Dillier, N., Fröhlich, T., Kompis, M., Bögli, H., and Lai, W. L. (**1993**). ''Digital signal processing (DSP) applications for multiband loudness correction digital hearing aids and cochlear implants,'' J. Rehabil. Res. Dev. **24**, 95–109.

Graupe, D., Grosspietsch, J. K., and Basseas, S. P. (**1987**). ''A single-microphone-based self-adaptive filter of noise from speech and its performance evaluation,'' J. Rehabil. Res. Dev. **24**, 119–126.

Gravel, J. S., Fausel, N., Liskow, C., and Chobot, J. (**1999**). ''Children's speech recognition in noise using omni-directional and dual-microphone hearing aid technology,'' Ear Hear. **20**, 1–11.

Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (**1993**). ''Intelligibility-weighted measures of speech-to-interference ratio and speech system performance,'' J. Acoust. Soc. Am. **94**, 3009–3010.

Greenberg, J. E., and Zurek, P. M. (**1992**). ''Evaluation of an adaptive beamforming method for hearing aids,'' J. Acoust. Soc. Am. **91**, 1662–1676.

Griffiths, L. J., and Jim, C. W. (**1982**). ''An alternative approach to linearly constrained adaptive beamforming,'' IEEE Trans. Antennas Propag. **30**, 27–34.

Hamacher, V., Mauer, G., and Döring, W. H. (**1996**). ''Untersuchung eines adaptiven beamforming-systems zur Störunterdrückung für Hörgeschädigte,'' Proceedings of the 22nd Deutsche Jahrestagung für Akustik (DAGA), Bonn, Germany (unpublished).

Kates, J. M. (**1997**). ''Relating change in signal-to-noise ratio to array gain for microphone arrays used in rooms,'' J. Acoust. Soc. Am. **101**, 2388–2390.

Kates, J. M., and Weiss, M. R. (**1996**). ''A comparison of hearing-aid array-processing techniques,'' J. Acoust. Soc. Am. **99**, 3138–3148.

Kochkin, S. (**1993**). ''Consumer satisfaction with hearing instruments in the United States,'' The Marketing Edge, Special issue June 1993, pp. 1–4.

Kochkin, S. (**1996**). ''Customer satisfaction and subjective benefit with high performance hearing aids,'' Hearing Rev. **3**, 16–26.

Kompis, M. (**1998**). ''Improving speech intelligibility with multi-microphone noise reduction systems for hearing aids,'' Curr. Top. Acoust. Res. **2**, 1–16.

Kompis, M., and Dillier, N. (**1991**). ''Noise reduction for hearing aids: Evaluation of the adaptive beamformer approach,'' Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. **13**, 1887–1888.

Kompis, M., and Dillier, N. (**1993**). ''Simulating transfer functions in a reverberant room including source directivity and head shadow effects,'' J. Acoust. Soc. Am. **93**, 2779–2787.

Kompis, M., and Dillier, N. (**1994**). ''Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer,'' J. Acoust. Soc. Am. **96**, 1910–1913.

Kompis, M., and Dillier, N. (**2001**). ''Performance of a two-microphone adaptive beamforming noise reduction scheme for hearing aids. II. Experimental verification of the predictions,'' J. Acoust. Soc. Am. **109**, 1134–1143.

Kompis, M., Dillier, N., Francois, J., Tinembart, J., and Häusler, R. (**1997**). ''New target-signal-detection schemes for multi-microphone noise-reduction systems for hearing aids,'' Proc. Annu. Int. Conf. IEEE Eng. Biol. Soc. **19**, 1990–1993.

Kompis, M., Feuz, P., François, J., and Tinembart, J. (**1999**). ''Multi-microphone digital-signal-processing system for research into noise reduction for hearing aids,'' Innovation Technol. Biol. Medicine **20**, 201–206.

Kompis, M., Oberli, M., and Brugger, U. (**2000**). ''A novel real-time noise reduction system for the assessment of evoked otoacoustic emissions,'' Comput. Biol. Med. **30**, 341–354.

Kornbluh, K. (**1999**). ''Math and science software,'' IEEE Spectr. January, 88–91.

Kuhn, G. F. (**1977**). ''Model for the interaural time difference for the azimuthal plane,'' J. Acoust. Soc. Am. **62**, 157–167.

Lim, J. S., and Oppenheim, A. V. (**1979**). ''Enhancement and bandwidth compression of noisy speech,'' Proc. IEEE **67**, 1586–1604.

Lurquin, P., and Rafhay, S. (**1996**). ''Intelligibility in noise using multimicrophone hearing aids,'' Acta Oto-Laryngol. Belg. **50**, 103–109.

Morse, P. M. (**1983**). *Vibration and Sound* (American Institute of Physics, New York), 2nd (Paperback) printing, Chap. 27, pp. 311–326.

Peterson, P. M., Durlach, N. I., Rabinowitz, W. M., and Zurek, P. M. (**1987**). ''Multimicrophone adaptive beamforming for interference reduction in hearing aids,'' J. Rehabil. Res. Dev. **24**, 103–110.

Peterson, P. M., Wie, S. M., Rabinowitz, W. M., and Zurek, P. M. (**1990**). ''Robustness of an adaptive beamforming method for hearing aids,'' Acta Oto-Laryngol., Suppl. **469**, 85–90.

Schwarz, L. (**1943**). ''Zur theorie der Beugung einer ebenen Schallwelle an der Kugel,'' Akust. Z. **8**, 91–117.

Soede, W., Berkhout, A. J., and Bilsen, F. A. (**1993**). ''Development of a new hearing instrument based on array technology,'' J. Acoust. Soc. Am. **94**, 785–798.

Stadler, R. W., and Rabinowitz, W. M. (**1993**). ''On the potential of fixed arrays for hearing aids,'' J. Acoust. Soc. Am. **94**, 1332–1342.

Valente, M. (**1998**). ''The bright promise of microphone technology,'' Hear. J. **51**, 10–15.

Valente, M., Fabry, D. A., and Potts, L. G. (**1995**). ''Recognition of speech in noise with hearing aids using dual microphones,'' J. Am. Acad. Audiol. **6**, 440–449.

Van Compernolle, D. (**1990**). ''Hearing aids using binaural processing principles,'' Acta Oto-Laryngol., Suppl. **469**, 76–84.

Vanden Berghe, J., and Wouters, J. (**1998**). ''An adaptive noise canceller for hearing aids using two nearby microphones,'' J. Acoust. Soc. Am. **103**, 3621–3626.

van Hoesel, R. J. M., and Clark, G. M. (**1995**). ''Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients,'' J. Acoust. Soc. Am. **97**, 2498–2503.

Welker, D. P., Greenberg, J. E., Desloge, J. G., and Zurek, P. M. (**1997**). ''Microphone-array hearing aids with binaural output. II. A two-microphone adaptive system,'' IEEE Trans. Speech Audio Process. **5**, 543–551.

Whitmal, N. A., Rutledge, J. C., and Cohen, J. (**1996**). ''Reducing correlated noise in digital hearing aids,'' IEEE Eng. Med. Biol. Mag. **15**, 88–96.

Widrow, B., Glover, J. R., McColl, J. M., Kaunitz, J. M., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, J. R., and Goodlin, R. C. (**1975**). ''Adaptive noise cancelling: Principles and applications,'' Proc. IEEE **63**, 1692–1716.

Widrow, B., and Stearns, S. D. (**1985**). *Adaptive Signal Processing* (Prentice–Hall, Englewood Cliffs, NJ).

Wouters, J., Litière, L., and van Wieringen, A. (**1999**). ''Speech intelligibility in noisy environments with one- and two-microphone hearing aids,'' Audiology **38**, 91–98.

Zwicker, E., and Zollner, M. (**1984**). *Elektroakustik* (Springer, Heidelberg).

Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. II. Experimental verification of the predictions

Martin Kompis[a)]
University Clinic of ENT, Head and Neck Surgery, Inselspital, 3010 Bern, Switzerland

Norbert Dillier
ENT-Department, University Hospital of Zurich, 8091 Zurich, Switzerland

(Received 5 November 1999; revised 24 April 2000; revised 25 September 2000; accepted 14 November 2000)

A method to predict the amount of noise reduction which can be achieved using a two-microphone adaptive beamforming noise reduction system for hearing aids [J. Acoust. Soc. Am. **109**, 1123 (2001)] is verified experimentally. 34 experiments are performed in real environments and 58 in simulated environments and the results are compared to the predictions. In all experiments, one noise source and one target signal source are present. Starting from a setting in a moderately reverberant room (reverberation time 0.42 s, volume 34 m^3, distance between listener and either sound source 1 m, length of the adaptive filter 25 ms), eight different parameters of the acoustical environment and three different design parameters of the adaptive beamformer were systematically varied. For those experiments, in which the direct-to-reverberant ratios of the noise signal is +3 dB or less, the difference between the predicted and the measured improvement in signal-to-noise ratio (SNR) is -0.21 ± 0.59 dB for real environments and -0.25 ± 0.51 dB for simulated environments (average±standard deviation). At higher direct-to-reverberant ratios, SNR improvement is systematically underestimated by up to 5.34 dB. The parameters with the greatest influence on the performance of the adaptive beamformer have been found to be the direct-to-reverberant ratio of the noise source, the reverberation time of the acoustic environment, and the length of the adaptive filter. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1338558]

PACS numbers: 43.66.Ts, 43.60.Lq, 43.60.Gk [RVS]

I. INTRODUCTION

Poor speech recognition in noisy environments is a major source of dissatisfaction for numerous users of cochlear implants and conventional hearing aids (Kochkin, 1993; Kiefer *et al.*, 1996). One promising approach to solve this problem is the two-microphone Griffiths–Jim beamformer or adaptive beamformer (Griffiths and Jim, 1982; Peterson *et al.*, 1987), where the signals of two microphones mounted close to the user's ears are postprocessed by an adaptive noise reduction scheme (Widrow *et al.*, 1975). A schematic representation is shown in the lower part of Fig. 1. A detailed description of the adaptive beamformer can be found elsewhere (Peterson *et al.*, 1987; Greenberg and Zurek, 1992; Kompis and Dillier, 2001) and is not repeated here. Numerous experiments have already been performed with this method, showing a wide range of signal-to-noise-ratio (SNR) improvements of 0 to 30 dB (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; van Hoesel and Clark, 1995; Hamacher *et al.*, 1996). Comparison of these data is difficult because of the different experimental settings and a lack of theoretical background to estimate the contribution of each of these differences on the results. In the companion paper (Kompis and Dillier, 2001), a theoretical framework has been presented, which allows the prediction of the noise reduction that can be expected from an adaptive beamforming noise reduction system in different acoustic settings. However, this framework by itself is of limited value only for two reasons. First, the predictions have not been validated by comparisons to experimental results. Second, as the prediction is a complex function of 11 input parameters, it is still relatively difficult to gain a concept of the complex behavior of the beamformer without systematic variations of all parameters. It is the aim of this investigation to start to close both of these gaps.

II. METHODS

The computer program presented in the companion paper estimates the SNR improvement which can be expected to be reached by an adaptive beamformer as a function of 11 acoustic and design parameters. It is not possible to test all parameter combinations with a reasonable number of different values for each of the 11 parameters. Instead, a different approach was chosen, in which a realistic experimental setting was defined first, from which each parameter is varied separately toward both greater and smaller values. In this way, the number of experiments was reduced to a manageable 92. Throughout this text, the experimental setting from which all parameters are varied will be called central setting.

A. Central setting

The definition of the central setting includes the room, the noise, and the signal source, and a set of design parameters of the adaptive beamformer. Figure 1 shows a sche-

[a)]Electronic mail: martin.kompis@insel.ch

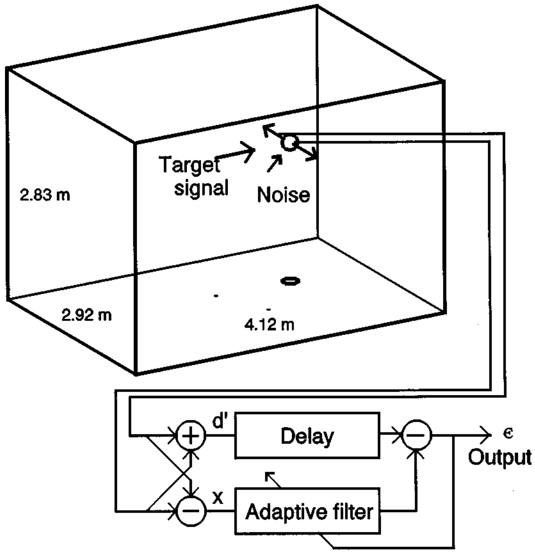

FIG. 1. Schematic drawing of the experimental setup at the central setting (top) and the adaptive beamformer (bottom). The starting points of the arrows in the upper portion of the diagram denote the locations of the loudspeakers and microphones.

matic representation. The parameters of the central setting have been chosen to represent a realistic situation, in which variations of all relevant parameters toward both higher and lower values appear to be reasonable. To define the properties of a suitable room, the dimensions and reverberation times of 18 different rooms (4 offices, 11 living or bedrooms, 1 bath, 2 kitchens) were measured. In this limited sample, the average volume was 34.1 m^3 and the average reverberation time (measured in octave bands with center frequencies of 125–4000 Hz) was 0.41 s and almost frequency independent. These average values may differ, e.g., in a different cultural context. One of these 18 rooms, a shoebox-shaped room with a volume of 34.0 m^3 and an almost frequency independent reverberation time of 0.42 s, was available for experiments for a limited time and was used for the central setting.

Two loudspeakers (Phillips 22AHS86/16R) were placed at a distance of 1 m from a dummy head equipped with a stereo microphone, both from a Sennheiser MKE 2002 set. The azimuth of the loudspeaker emitting the target signal was 0° (i.e., in front of the dummy head), the noise source was 45° to its right. The index of directionality of the loudspeakers was estimated to be 3.4 for band-limited noise 125 and 5000 Hz in an earlier work (Kompis and Dillier, 1993). As to the adaptive beamformer, a sampling rate of 10 240 Hz, a filter length of 25 ms (256 filter coefficients), and a delay of 50% of the filter length (12.5 ms) in the target signal path (marked d' in Fig. 1) was chosen.

B. Experiments in real and simulated rooms

As far as possible, experiments were performed in the real room described. Some of the experimental parameters, most notably the volume of the room and the reverberation time, cannot be readily varied independently using real rooms. Furthermore, the room which was used for the central setting was available only for a limited time for recordings. For these reasons, 58 of the 92 experiments were performed in simulated rooms, using a simulation method presented earlier (Kompis and Dillier, 1993). This simulation procedure is based on an image method introduced by Allen and Berkley (1979). It simulates the impulse responses between acoustic sources and microphones in shoebox-shaped rooms, taking into account the effects of directional sound sources and the acoustic head-shadow of the listener. The head is modeled as a rigid sphere. For the simulations, a value of 18.6 cm was chosen for the diameter of this sphere, as proposed by Kuhn (1977) and used in an earlier study (Kompis and Dillier, 1993). The index of directionality of 3.4 for the two sound sources was approximated by an opening angle of ±65°. Within this opening angle, the signal is emitted equally into all directions, and no signal is emitted outside this angle. For the simulated version of the central setting, all other simulation parameters (i.e., reverberation time, room dimensions, relative positions of the sound sources and the listener) were the same as the corresponding parameters of the real room. The suitability of the simulation method for the purpose at hand was validated in the first experiment (Sec. III A). For the prediction of the performance of the adaptive beamformer, the same set of input parameters was used for both the real and the simulated central setting. Table I shows a synopsis of these input parameters.

TABLE I. Synopsis of the input parameters used to predict the performance for the central setting.

Parameter	Value
Room size V	34 m^3
Reverberation time T_r	0.42 s
Distance listener to target signal source l_s	1 m
Index of directionality of target signal source γ_s	3.4
Alignment factor of target signal source A	4
Distance listener to noise source l_n	1 m
Index of directionality of noise source γ_n	3.4
Azimuth of noise source α_n	45°
Sampling rate F_s	10 240 Hz
Number of coefficients in adaptive filter N	256
Delay in target signal path Δ	128 samples

As the signals from the real and simulated central setting were used for several experiments with different values of the design parameters of the adaptive beamformer, only 14 different sets of recordings in real environments and 37 different sets of simulated signals were used. Table II shows a synopsis of the experiments and environments used.

C. Signal acquisition and signal processing

Target and noise signals were recorded (or for the experiments in simulated rooms: simulated) separately. According to the paradigm used for the prediction of the improvement of the signal-to-noise ratio (SNR) described previously (Kompis and Dillier, 2001), the signals of both the noise and the target signal source were white noise. White noise was generated on a computer and played back via a digital audio tape (DAT) recorder driving only one of the two loudspeakers at a time. Uncorrelated noise sequences of 3 s duration were used for the target and the noise signals, respectively. Recordings of the microphone signals at the dummy head were digitized at a sampling rate of 10 240 Hz

TABLE II. Synopsis of the number of experiments in real and simulated environments.

Parameter varied	Experiments in real environments	Experiments in simulated environments	Sets of recorded signals	Sets of simulated signals
None (central setting)	1	1	1	1
Azimuth of noise signal source (Fig. 3)	5[a]	⋯	5[b]	⋯
Distance to noise source (Fig. 4)	2[a]	4	2	4
Index of directionality of noise source (Fig. 5)	1[a]	6	1	6
Alignment factor (Fig. 6)	1[a]	3	1	3
Distance to target signal source (Fig. 7)	2[a]	4	2	4
Index of directionality of target signal source (Fig. 8)	1[a]	6	1	6
Reverberation time (Fig. 9)	1[a]	7[c]	1	7
Room size (Fig. 10)	⋯	6	⋯	6
Filter length (Fig. 11)	4[a]	⋯	0[b]	⋯
Delay in target signal path (Fig. 12)	16[a,d]	17[c]	0[b]	0[b]
Sampling rate (Fig. 13)	⋯	4[c]	⋯	0[b]
Total	34	58	14	37

[a]In addition, results from the real central setting are shown in the corresponding figure.
[b]Recorded or simulated input signals used are identical to those at central setting.
[c]In addition, results from the simulated central setting are shown in the corresponding figure.
[d]In addition, results of one experiment already shown in Fig. 11 (filter length 512, delay 50%) is shown in Fig. 12.

into a computer using a custom-built 12 bit stereo analog-to-digital converter and appropriate antialiasing filters. The spectra of the recorded signals were found to rise slightly toward higher frequencies. Although the effect of this spectral feature on the SNR improvement was found to be small, a two-coefficient finite impulse-response filter (coefficients 0.5 and 0.65) was used to equalize the spectra to within 2 dB in the frequency range of 125–4900 Hz. For experiments in simulated environments, white noise was directly filtered by the impulse responses generated by the room simulation program. The spectra of the microphone signals in the simulations were found to be flat to within 2 dB without further conditioning.

To measure the gain in signal-to-noise ratio, first the recording of the noise signal alone was processed by an adaptive beamforming algorithm. The adaptation of the filter was performed using a normalized least-mean squares algorithm (Bellanger, 1987), where the adaptation time constant was chosen to be 0.2 s. After 2 s the filter was assumed to be in an adapted state and the signal variance in the following second was used as a measure of the variance of the noise signal at the output of the beamformer. The adapted filter was temporarily stored and used in a second run, where the recorded or simulated target signal was processed alone, with the adaptation disabled, i.e., maintaining the adapted coefficients from the first run. Again, the variance of the output signal during 1 s was used as a measure of the variance of the target signal at the output of the adaptive beamformer. Using this procedure, a perfect target-signal detection scheme, which prevents any filter adaptation while a target signal is present, is mimicked. Several such schemes have been proposed (Van Compernolle, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; Kompis *et al.*, 1997) and used in experiments, and the theoretical analysis and prediction of SNR improvement (Kompis and Dillier, 2001) is based on the assumption that one of these schemes is employed. For all experiments involving longer filters (>25 ms), filter adaptation was allowed for 4 s instead of only 2 s to compensate for the proportionally longer adaptation time constant.

D. SNR improvement and intelligibility-weighted gain

As the investigated prediction method predicts SNR improvement, this measure is used to represent the experimental results. However, for hearing aid applications, the primary goal is improved speech intelligibility and not improved SNR. Because some frequency bands contribute more to speech intelligibility than others, SNR improvement may correlate poorly with improvement in speech recognition, if substantial differences between SNR improvements in different frequency bands do exist. To estimate this effect on the presented data, all experimental results which were compared to the theoretical predictions were also examined by an intelligibility-weighted measure proposed by Greeberg *et al.* (1993). To calculate this intelligibility-weighted gain, signal-to-noise ratios were calculated in 15 one-third-octave bands with center frequencies between 200 and 5000 Hz and weighted according to their contribution to the articulation index (ANSI, 1969).

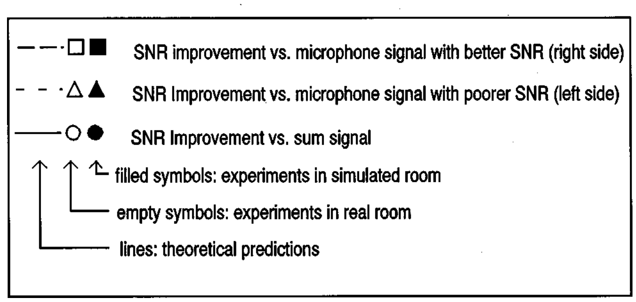

FIG. 2. Legend for the symbols and lines used in Figs. 3–14.

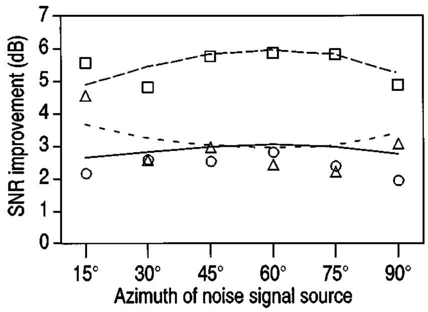

FIG. 3. Influence of the azimuth of the noise signal source. See Fig. 2 for a legend of the symbols used.

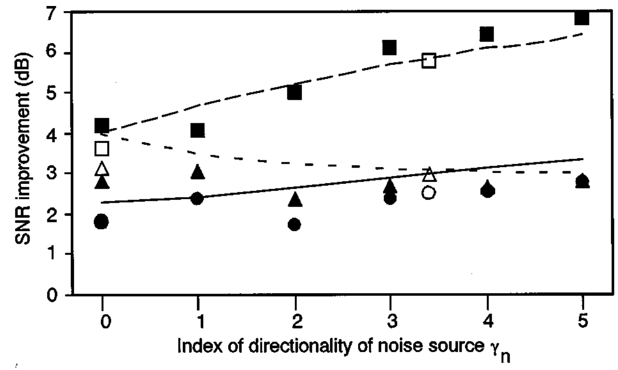

FIG. 5. Influence of the directionality of the noise source. See Fig. 2 for a legend of the symbols used.

E. Representation of the results

The computer program (Kompis and Dillier, 2001) implementing the prediction of the SNR improvement of the adaptive beamformer calculates three different numbers: the SNR improvement versus the microphone signal with the more favorable SNR, the SNR improvement versus the microphone signal with the poorer SNR, and the improvement versus the sum of both microphone signals. The latter corresponds to a simple two-microphone beamformer with fixed postprocessing.

To allow direct comparison, the results of the experiments are similarly calculated as improvements versus each microphone signal and the sum of both microphone signals. Therefore, six sets of data are shown in the figures of Sec. III. All predicted improvements are connected by different lines, and all results from experiments are shown as individual symbols. Results from experiments in real environments are shown using open symbols; results from experiments in simulated environments are shown using closed symbols. Figure 2 shows a legend for all lines and symbols used in Figs. 3–14. SNR improvements at the output of the adaptive beamformer, compared to the SNR at the left microphone (opposite from the noise source and therefore more favorable SNR; triangles in figures) will be lower than SNR improvement versus the right microphone facing the noise source (poorer SNR; squares in figures).

III. RESULTS

A. Results at central setting

At central setting, the predicted SNR improvement was compared to the results from both the experiments in the real room and in the simulated environment. Table III shows the results. All three sets of SNR improvements, i.e., the prediction and the two sets of experimental results, are within 0.5 dB within each other with absolute values ranging from 2.50 to 5.97 dB. None of the data sets exhibit systematically higher or lower values for all SNR improvements when compared to the other two sets.

B. Effect of the acoustic parameters of the noise signal source

The three parameters characterizing the noise signal source are its azimuth α_n (with $\alpha_n = 0$ defined as the forward direction of the listener), the distance between noise source

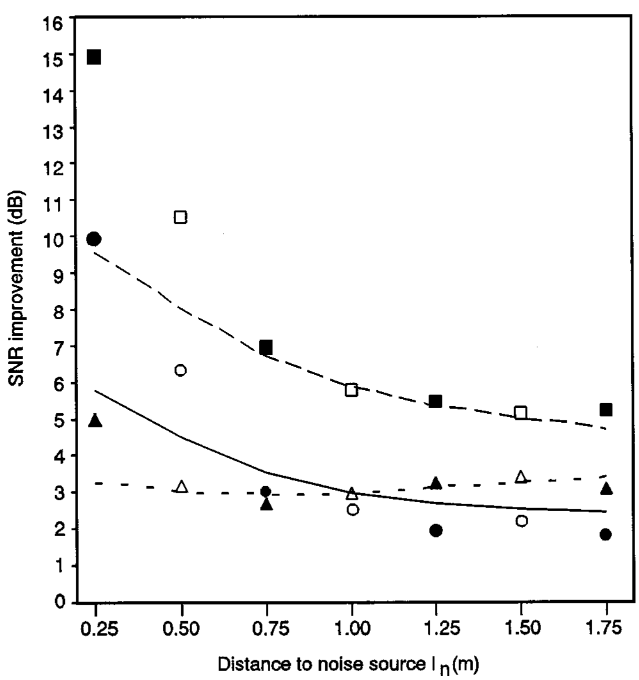

FIG. 4. Influence of the distance between noise source and listener. See Fig. 2 for a legend of the symbols used.

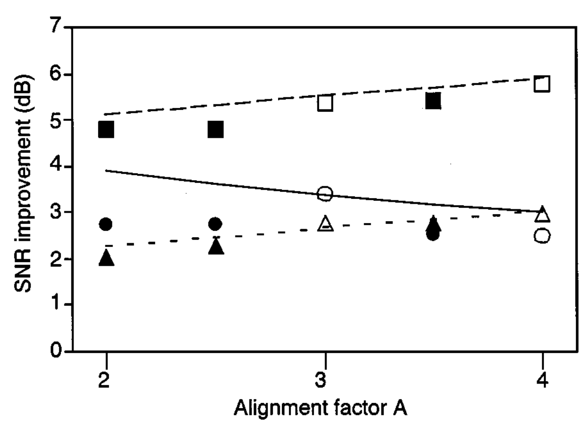

FIG. 6. Influence of the alignment of the target signal source. See the text for the definition of the alignment factor A. See Fig. 2 for a legend of the symbols used.

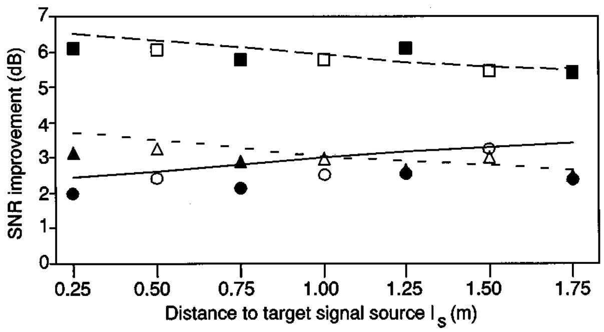

FIG. 7. Influence of the distance between target signal source and listener. See Fig. 2 for a legend of the symbols used.

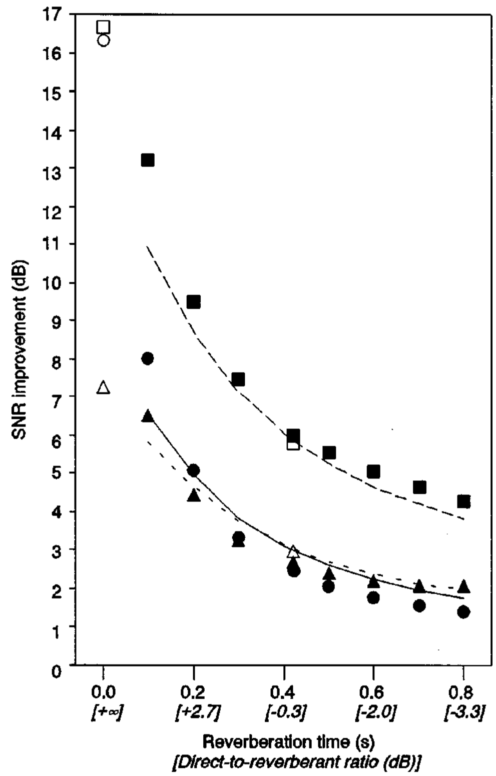

FIG. 9. Influence of the reverberation time. See Fig. 2 for a legend of the symbols used.

and listener l_n, and the index of directionality γ_n of the noise source. The last two factors influence the direct-to-reverberant $P_{\mathrm{d/r}}$ ratio of the noise signal. The index of directionality is defined as the ratio between the signal intensity emitted in the direction of the listener to the intensity of a hypothetical omnidirectional source with the same total acoustic output power (DeBrunner and McKinney, 1995).

Figure 3 shows the results of the experiments at 6 different angles of incidence between 15° and 90°. Because of the symmetry of the setting, results can be extrapolated for all angles in the horizontal plane, except for the front (0°) and the rear (180°), where the adaptive beamformer assumes the position of a target- and not of a noise-signal source. The largest difference between prediction and experimental results is 0.89 dB, with more than half of the experimental results lying within 0.5 dB of the predictions.

Figure 4 shows the SNR improvement as a function of the distance between listener and noise source. For distances of 0.75 m ($P_{\mathrm{d/r}}=2.0$ dB) and more, predictions and experimental results are reasonably in accordance. At distances of 0.5 m ($P_{\mathrm{d/r}}=5.5$ dB) and less, however, predictions considerably underestimate the SNR improvement which can actually be achieved using the adaptive beamformer. At $l_n = 0.25$ m, the difference is as large as 5.34 dB.

Figure 5 shows the SNR improvement as a function of the index of directionality γ_n of the noise source. For the experiments in the real acoustic environment, $\gamma_n=3.4$ corresponds to the loudspeaker facing the dummy head (central setting), whereas $\gamma_n=0$ was approximated by turning the loudspeaker away from the dummy head. For the simulations, $\gamma_n=1$, 2, 3, 4, and 5 was approximated similar to the central setting with opening angles of ±180°, ±90°, ±70°, ±60°, ±53°, respectively. For $\gamma_n=0$, an opening angle of ±90° facing *away* from the dummy head was used.

There is a reasonable agreement between the predicted SNR improvement and the results of the actual measurements, with an average error of 0.51 dB (range 0.02–1.14 dB). SNR improvement increases with γ_n when compared to the sum signal and to the microphone signal with the less favorable SNR, but decreases slightly when compared to the microphone signal with the higher SNR, presumably due to the increased direct-to-reverberant ratio of the noise signal.

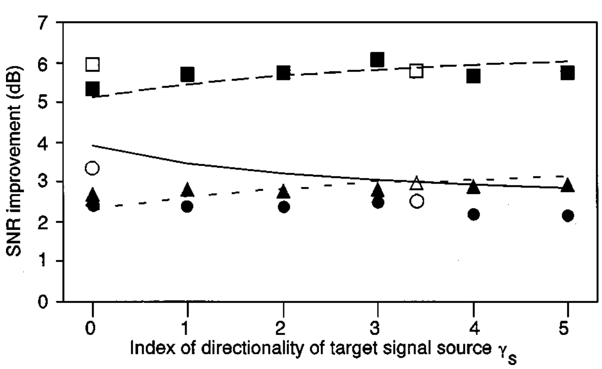

FIG. 8. Influence of the directionality of the target signal source. See Fig. 2 for a legend of the symbols used.

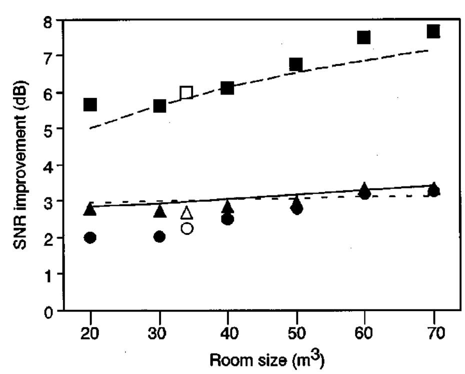

FIG. 10. Influence of the room size. See Fig. 2 for a legend of the symbols used.

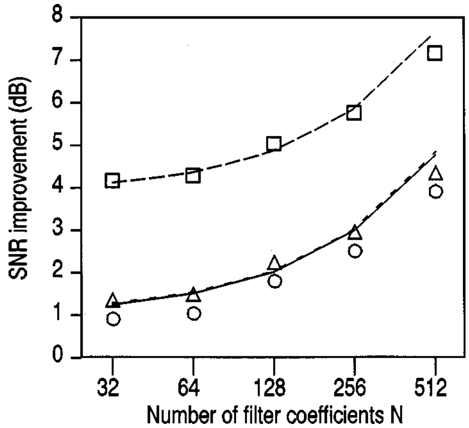

FIG. 11. Influence of the length of the adaptive filter. See Fig. 2 for a legend of the symbols used.

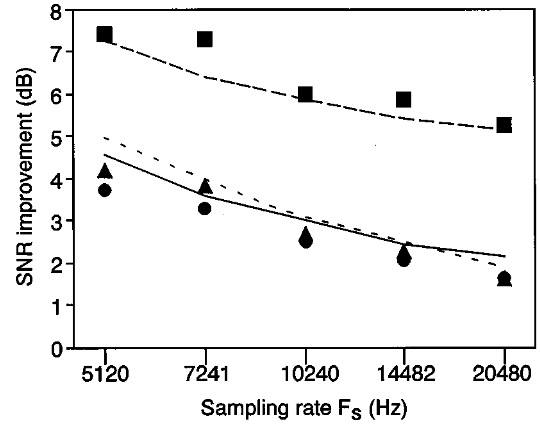

FIG. 13. Influence of the sampling rate. See Fig. 2 for a legend of the symbols used.

C. Effect of the acoustic parameters of the target signal source

The three parameters describing the target signal source are the alignment factor A, the distance between dummy head and target signal source l_s, and the index of directionality of the target signal source γ_s.

The alignment factor A is defined as the ratio between the variance of the nonreverberant portion of a white noise signal after summation of the microphone signals (signal d' in Fig. 1) and the sum of the variances of the two individual microphone signals [cf. Kompis and Dillier (2001) for a detailed discussion]. For perfect alignment, i.e., if there is no delay between the nonreverberant part of the target at the two microphones, A is 4. For a head-sized spacing between microphones and a sampling rate of 10 240 Hz, A drops to 2 (no alignment) for azimuths of approximately 8° and more. The alignment factor can be directly measured in anechoic (real or simulated) environments. For the given setting, the alignment factor was found to be 4 at an azimuth α_s of 0°, 3.5 at $\alpha_s = 3°$, 3 at $\alpha_s = 5°$, 2.5 at $\alpha_s = 6°$, and 2 at $\alpha_s = 8°$.

Figure 6 shows the comparison between predictions and experimental results. The values are in reasonable agreement, i.e., within 0.5 dB when taking either one microphone signal as a reference. For $A=2$ and $A=2.5$, the agreement between predicted and measured SNR improvement versus the sum signal differ by 0.88 and 1.15 dB, respectively. The reason for this difference is not completely clear, but most probably a result of the relatively simple model of the direct and reverberant signal parts used to predict the SNR improvement (Kompis and Dillier, 2001).

Figure 7 shows the dependence of the SNR improvement as a function of the distance l_s between listener and the target sound source. Again, the majority of all measured values lie within 0.5 dB of the predictions and display the same tendencies (i.e., SNR improvements decreasing with the distance when taking the microphone signals as a reference, but

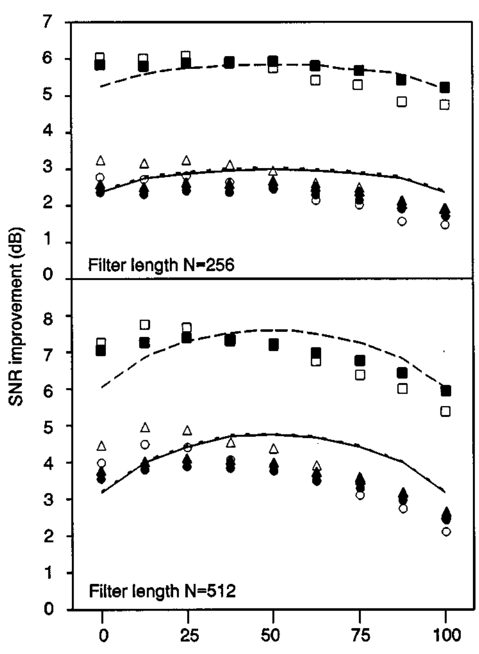

FIG. 12. Influence of the delay in the target signal path. See Fig. 2 for a legend of the symbols used.

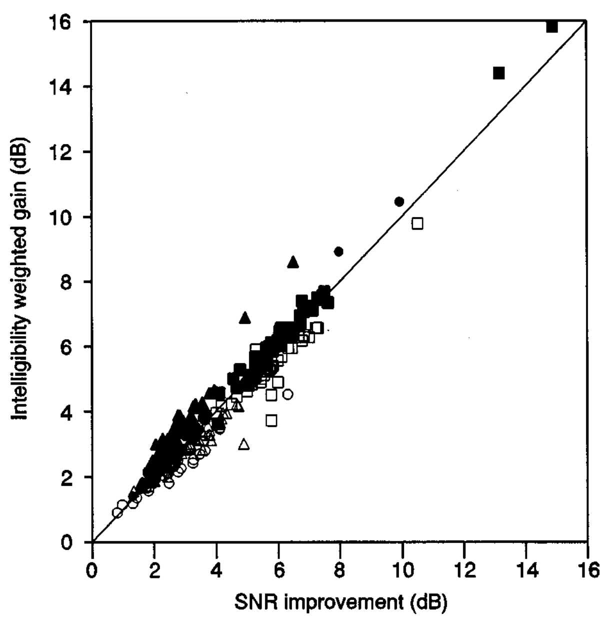

FIG. 14. SNR improvement vs Intelligibility-weighted gain for the experiments in real and simulated environments. See Fig. 2 for a legend of the symbols used.

TABLE III. SNR Improvement at central setting.

	Improvement vs microphone with better SNR (dB)	Improvement vs microphone with poorer SNR (dB)	Improvement vs sum of microphone signals (dB)
Model prediction	3.03	5.85	2.99
Experiment in real room	2.97	5.77	2.52
Experiment in simulated room	2.69	5.97	2.50

increasing when taking the sum signal as a reference). For the given range of distances between 0.25 and 1.75 m, SNR improvements change on the order of magnitude of 1 dB.

Figure 8 shows the dependence of the SNR improvement as a function of the index of directionality of the target signal source γ_s. To obtain the different values for γ_s between 0 and 5, exactly the same procedures were used as for the noise source in Sec. III C. Again, there is a reasonable agreement between prediction and measurement for most data points, with the greatest differences lying toward small γ_s and SNR improvements versus the sum of the microphone signals. For the given range of values $\gamma_s = 0-5$, SNR improvements change on the order of magnitude of 1 dB.

D. Room size and reverberation time

The two acoustical parameters of the room considered in the theoretical analysis of the performance of the adaptive beamformer are volume V of the room and reverberation time T_r. Reverberation time is the time required for a reverberant signal to decay by 60 dB. To vary these two parameters independently, experiments were performed predominately in simulated rooms. Apart from the central setting, only one experiment was performed in a real, anechoic environment.

Figure 9 shows the dependence of the SNR improvement as a function of reverberation time. SNR improvement increases rapidly at short reverberation times. Theoretical predictions and results from the experiments are in reasonable agreement for reverberation times of approximately 0.2 s and above. For shorter reverberation times, the predictions systematically underestimate SNR improvements by up to 2.28 dB at $T_r = 0.1$ s. At $T_r = 0$ s (anechoic environment) the prediction cannot be calculated, as one of the underlying assumptions, the existence of a reverberant signal portion, is violated. The direct-to-reverberant ratio of the noise source is approximately 5.7 dB at $T_r = 0.1$ s, 2.7 dB at $T_r = 0.2$ s, and 1.0 dB at $T_r = 0.3$ s.

Figure 10 shows SNR improvement by the adaptive beamformer as a function of room size. For these experiments, rooms with volumes of 20–70 m^3, in steps of 10 m^3 were simulated. The reverberation time T_r of all these rooms was 0.42 s. To keep T_r constant, the absorption coefficients of the simulated rooms were higher for the larger rooms. Note that for the same absorption coefficients for all rooms, reverberation time would have increased with room size, as everyday experience suggests. As the direct-to-reverberant ratio increases with room size (-2.8 dB at $V = 20\,m^3$, $+2.7$ dB at $V = 70\,m^3$), SNR improvement, especially when compared to the microphone signal with the lower signal-to-noise ratio, increases by approximately 2 dB.

E. Design parameters of the adaptive beamformer

In the theoretical analysis (Kompis and Dillier, 2001), the influence of three design parameters of the adaptive beamformer is considered: the number of filter coefficients in the adaptive filter N, the delay in the sum signal path, and the sampling rate F_s of the system. In practical situations, there will be additional design parameters which influence the performance of a given beamformer, such as the adaptation time constant, the resolution of the analog-to-digital converters, and the performance of any target-signal-detection/adaptation inhibition scheme to prevent filter adaptation in the presence of target signal and therefore target signal cancellation. However, in the theoretical analysis and as a consequence in this study, these additional factors are assumed to be ideal, i.e., filter adaptation is perfect and occurred in the presence of the noise signal only, and all implementation issues are considered to be negligible.

Figure 11 shows SNR improvement as a function of the number N of coefficients in the adaptive filter. In the literature in similar noise reduction algorithms filter lengths of up to 40 ms (Peterson *et al.*, 1987) have been reported. At a sampling rate of 10 240 Hz, this corresponds to a range of $N = 410$ coefficients. In this study, filter lengths between $N = 32(3.125\,\text{ms})$ and $N = 512(50\,\text{ms})$ have been studied. It can be seen that the amount of SNR improvement increases substantially with filter length, especially for values of N of 128 and more. With short filters, e.g., $N = 32$ or $N = 64$, the adaptive beamformer practically routes the microphone signal with the more favorable SNR to the output, but provides only relatively little (approximately 1 dB) SNR improvement above that. In Fig. 11 this is shown by the SNR improvement versus the microphone with the lower SNR improving by more than 4 dB, but only by just above 1 dB when compared to the other microphone signal. For long filters ($N = 512$), the SNR is improved by more than 4 dB, even when compared to the microphone signal with the *more* favorable SNR. The agreement between experimental results and the predictions is reasonable for the entire range of $N = 32-512$, and is poorest for the longest filter.

The theoretical analysis predicts that influence of the delay in the target signal path d' on the amount of noise reduction depends on the length of the adaptive filter N. For this reason, experiments with delays between 0% and 100% of the filter lengths and two different filter lengths ($N = 256$ and $N = 512$ coefficients) were performed for both the real and the simulated central setting. Figure 12 shows the results. It can be seen that for the shorter filter, a variation of only 0.6 dB in SNR improvement is predicted for the entire range of delays between 0 and 100% of the filter length. For the longer filter, this effect is predicted to be larger (1.59 dB). In both cases, the maximal noise reduction is predicted at a delay of 50% of the filter length. The experimental results show an amount of noise reduction and—to a certain degree—a shape of the curves which are similar to the predicted ones. However, there is one major difference: The maximal noise reduction is reached at delays between 12.5% and 25% of the filter length and not at 50%, as predicted. This holds for both the real and the simulated environment and for both filter lengths. For the shorter filter, there is a second, only slightly smaller maximum at a delay of 0% in both environments. In this respect, the theoretical prediction clearly fails. The reasons for and implications of this failure will be discussed in Sec. IV E.

The last design parameter to be considered is the sampling rate F_s in Fig. 13. In real applications, the range of possible values is small, as sampling rates below approximately 7000 Hz will reduce speech recognition unacceptably, and the computational load rises rapidly with higher sampling rates. For the range of $F_s = 5120\text{–}20\,480\,\text{Hz}$, the SNR improvement drops on the order of magnitude of 2 dB. This effect can be explained by the effectively shorter filter (12.5 ms at $F_s = 20\,480\,\text{Hz}$, compared to 50 ms at $F_s = 5120\,\text{Hz}$) for the same number of coefficients $N = 256$, which was kept constant.

F. SNR improvement and intelligibility-weighted gain

Figure 14 shows the comparison between SNR improvement (as shown in Figs. 3–13) and intelligibility-weighted gain G_i (Greenberg *et al.*, 1993) for all experimental results which were compared to the theoretical predictions. Although differences up to 2.11 dB do exit, for the majority of all data points SNR improvement and intelligibility-weighted gain G_i are within 0.5 dB. On the average, SNRs are slightly higher than G_i for experiments in real environment (average difference +0.37 dB), while SNRs are slightly lower for the experiments in simulated environments (average difference −0.34 dB).

IV. DISCUSSION

A. Agreement between predictions and measurements

Agreement between experimental results and predictions appear to be reasonable for low direct-to-reverberant ratios of the noise signal $P_{\text{d/r}}$, but considerably poorer for high direct-to-reverberant ratios (cf. Figs. 4 and 9). When the experimental results are compared for all 88 experiments with a $P_{\text{d/r}} < +3$ dB, an average difference of −0.23 dB and a standard deviation of 0.54 dB can be observed. For the 32 experiments in real rooms, the mean difference is −0.21 dB (std. dev. 0.59 dB) and for the 56 experiments in simulated rooms it is −0.25 dB (std. dev. 0.51 dB). As to the 4 experiments with a $P_{\text{d/r}}$ above +3 dB, one comparison with the predicted values is not possible, as the predictions fail at infinite $P_{\text{d/r}}$ (anechoic environment, Fig. 9), and for the other 3 cases differences up to 5.34 dB (Fig. 4) can be observed. From the assumptions of the underlying theoretical analysis (Kompis and Dillier, 2001), it can be expected that agreement between experimental results and predictions is reasonable for low direct-to-reverberant ratios of the noise signal $P_{\text{d/r}}$, but poor for high direct-to-reverberant ratios. From the results shown in Figs. 4 and 9 it can be concluded that the prediction is reasonable for situations with direct-to-reverberant ratios of the noise source of up to approximately +3 dB, while noise suppression is underestimated for higher $P_{\text{d/r}}$. In contrast, the influence of direct-to-reverberant ratio of the *target* signal source appears to be small.

If $P_{\text{d/r}}$ is less than +3 dB, predictions give, on average, a slightly (0.23 dB) higher noise suppression than the experimental results. As the delay in the target-signal path (between d' and d in Fig. 1) is kept at a suboptimal 50% of the filter length (cf. Fig. 12) for the majority of the experiments, it can be expected that experimental results are slightly poorer than the predictions. The standard deviation of the differences between predicted and measured values of approximately 0.5 dB is comparable to the small variations in results, if, e.g., the entire experimental apparatus is shifted by a few centimeters in any direction, as verified by informal tests (Kompis and Dillier, 1993). As seen in Table III, and confirmed by the data presented in Figs. 3–13, results of the experiments in real and simulated rooms are in reasonable agreement, thus supporting the assumption that the chosen room-simulation algorithm (Kompis and Dillier, 1993) is suitable for these experiments involving the adaptive beamformer. One difference between the results of the experiments in real and simulated environments is the tendency to overestimate intelligibility-weighted gain G_i by using SNR improvements for real rooms, and underestimate G_i for simulated environments. This relatively small difference can be attributed to the small differences in the spectra of the simulated and recorded signals.

B. Influence of the noise signal source

From the three parameters defining the noise source, the azimuth has the smallest effect on the amount of noise reduction of the adaptive beamformer. Experimental results and predictions are in reasonable agreement.

Noise reduction is greatly increased with smaller distances between noise source and listener (Fig. 4). The agreement between experimental results and predicted noise reduction is reasonable for distances of 0.75 m or more, but poor for smaller distances. From the assumptions used to derive the predictions (Kompis and Dillier, 2001) it is known that the direct-to-reverberant ratio of the noise signal may not become too large for the predictions to remain valid. The estimated direct-to-reverberant ratio is 2.0 dB at 75 cm and 5.5 dB at 50 cm. According to the presented data, the transition between reasonable prediction and substantial underestimation of the SNR improvement takes place in this range. Although this does limit the usefulness of the prediction method in environments with no or very little reverberation, it is not a serious limitation for most normal rooms with realistic amounts of reverberation. To reach a direct-to-reverberant ratio of +3 dB or more in the room used for the central setting, an omnidirectional noise source must be less than 36 cm away from the listener.

For the investigated range of the index of directionality of the noise source ($\gamma_n = 1\text{–}5$), noise reduction changes only moderately (order of magnitude 1–2 dB), depending on the reference signal (left microphone, right microphone, or sum of microphone signals) to which SNR improvement is compared.

C. Influence of the target signal source

The influence of the target signal source on the performance of the adaptive beamformer is small. For the entire range of the alignment factor $A=2$ to $A=4$, the measured and predicted SNR improvement changes by less than 1 dB. SNR improvements differ by the same order of magnitude for the range of values considered for the distance to the signal sound source ($l_s=0.25$–1.75 m) as well as the index of directionality ($\gamma_s=0$ to 5).

D. Influence of room size and reverberation

While room size has only a limited effect on the performance of the adaptive beamformer at a fixed reverberation time (Fig. 10), noise reduction drops rapidly with increasing reverberation times T_r (Fig. 9). This phenomenon has been reported previously by several researchers (Peterson *et al.*, 1987; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; van Hoesel and Clark, 1995).

Experimental results and predictions are in reasonable agreement for the range of room sizes considered in this study and for reverberation times of 0.2 s and more, i.e., corresponding to direct-to-reverberant rations $P_{\mathrm{d/r}}$ of the noise source of +2.7 dB or less. As noted earlier, predictions systematically underestimate the noise reduction for lower T_r and—consequently—higher $P_{\mathrm{d/r}}$.

E. Influence of the design parameters of the adaptive beamformer

From the three design parameters considered, the sampling rate (Fig. 13) has the smallest range of reasonable values and at the same time a relatively small impact on the performance. The length of the adaptive filter significantly influences the SNR improvement of the adaptive beamformer, as has been noted by several researchers (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993). From our data, we conclude that short filters (e.g., $N=32$) improve SNR to only little above the SNR of the microphone with the more favorable SNR. Only a longer filter in the range of 128–512 coefficients provides substantial additional gains in SNR of 2–4 dB.

The influence of the amount of delay (Fig. 12) is clearly not predicted correctly. For both filter lengths and in both the real and the simulated environment, optimal performance of the beamformer is reached at considerably shorter delays than the predicted 50% of the length of the adaptive filter. This may also explain why in Fig. 11 the agreement between prediction and experimental result is poorest for the longest filter, where the influence of the delay is largest. The reason for the shorter optimal delay is not completely understood. Preliminary results from a small separate investigation suggest a loose relationship between the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise signal and the optimal delay: for small $P_{\mathrm{d/r}}$, the optimum seems to be close to the predicted 50%, whereas for greater $P_{\mathrm{d/r}}$, e.g., above 0 dB, the optimum tends to be often between 12.5% and 25%.

F. Applicability of the results for hearing aid applications

The presented experiments involve several simplifications, which are not necessarily met in real-life situations encountered by potential future users of an adaptive beamformer. These simplifications, which were made necessary by the assumptions on which the theoretical predictions are based, include: (1) a completely adapted filter, (2) filter adaptation in the absence of the target signal, (3) white noise emitted by both the noise and the target signal source, (4) no movement of either listener or either sound source, and (5) the presence of a single noise source only. Because of the usually fast adaptation time constants [order of magnitude: below 0.1 s (Dillier *et al.*, 1993)] and the availability of several target-signal-detection/adaptation-inhibition schemes (Van Compernolle, 1990; Greenberg and Zurek, 1992; van Hoesel and Clark, 1995; Kompis *et al.*, 1997) assumptions (1) and (2) are likely to be reasonably approached in real-life situations. Acoustic signals in relevant everyday situations will probably be composed of predominately low frequency signals such as speech and traffic noise rather than white noise, as assumed here. However, many implementations of adaptive beamformers use pre-emphasis filters just after the microphones, which prewhiten the spectra of these signals. Therefore, the spectras of probable real-life acoustic signals will at least approach that of white noise to a certain degree. SNR improvements appear to be reasonable estimates for the expected improvement in intelligibility in a number of situations as shown by the data in Fig. 14 and confirmed by tests using a portable real-time realization of the adaptive beamformer (Kompis *et al.*, 1999).

In every-day situations, a certain amount of relative movement between the listener and sound sources must be expected. It is difficult to estimate the influence of such movements. However, due to the usually short adaptation time constants this influence may be small. As to the presence of multiple noise sources, a limited number of experiments have already been reported (Peterson *et al.*, 1990). A substantial drop in performance can be expected if the spectra and levels of the sound sources are similar (Greenberg and Zurek, 1992).

V. SUMMARY AND CONCLUSIONS

A method to predict the amount of noise reduction which can be achieved using a two-microphone adaptive beamforming noise reduction system for hearing aids (Kompis and Dillier, 2001) was verified experimentally. 92 experiments were performed in real and simulated environments and the results were compared with the predictions. It was shown that predictions and experimental results agree reasonably, if the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise source is smaller than approximately +3 dB. For higher $P_{\mathrm{d/r}}$, the predictions systematically underestimate the performance of the adaptive beamformer. The parameters with the greatest influence on the performance of the adaptive beamformer were found to be the direct-to-reverberant ratio of the noise source, the reverberation time of the acoustic environment, and the length of the adaptive filter.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Research Foundation, Grant Nos. 4018-10864 and 3238-56352.99, and Ascom Tech Ltd.

Allen J. B., and Beakley, D. A. (**1979**). ''Image method for efficiently simulating small-room acoustics,'' J. Acoust. Soc. Am. **65**, 943–950.

ANSI (**1969**). ANSI S3.5-1969, ''American National Standard Methods for the Calculation of the Articulation Index'' (American National Standards Institute, New York).

Bellanger, M. G. (**1987**). *Adaptive Digital Filters and Signal Analysis* (Marcel Dekker, New York).

DeBrunner, V. E., and McKinney, E. D. (**1995**). ''Directional adaptive least mean square acoustic array for hearing aid enhancement,'' J. Acoust. Soc. Am. **98**, 437–444.

Dillier, N., Fröhlich, T., Kompis, M., Bögli, H., and Lai, W. L. (**1993**). ''Digital signal processing (DSP) applications for multiband loudness correction digital hearing aids and cochlear implants,'' J. Rehabil. Res. Dev. **24**, 95–109.

Greenberg, J. E., and Zurek, P. M. (**1992**). ''Evaluation of an adaptive beamforming method for hearing aids,'' J. Acoust. Soc. Am. **91**, 1662–1676.

Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (**1993**). ''Intelligibility-weighted measures of speech-to-interference ratio and speech system performance,'' J. Acoust. Soc. Am. **94**, 3009–3010.

Griffiths, L. J., and Jim, C. W. (**1982**). ''An alternative approach to linearly constrained adaptive beamforming,'' IEEE Trans. Antennas Propag. **30**, 27–34.

Hamacher, V., Mauer, G., and Döring, W. H. (**1996**). ''Untersuchung eines adaptiven Beamforming-Systems zur Störunterdrückung für Hörgeschädigte,'' Proceedings of the 22nd Deutsche Jahrestagung für Akustik (DAGA), Bonn, Germany (unpublished).

Kiefer, J. *et al.* (**1996**). ''Speech understanding in quiet and in noise with the CIS speech coding strategy (MED-EL Combi-40) compared to the multipeak and spectral peak strategies (Nucleus),'' ORL **58**, 127–135.

Kochkin, S. (**1993**). ''Consumer satisfaction with hearing instruments in the United States,'' The Marketing Edge, Special issue June 1993, 1–4.

Kompis, M., and Dillier, N. (**1993**). ''Simulating transfer functions in a reverberant room including source directivity and head shadow effects,'' J. Acoust. Soc. Am. **93**, 2779–2787.

Kompis, M., and Dillier, N. (**2001**). ''Performance of a two-microphone adaptive beamforming noise reduction scheme for hearing aids. II. Experimental verification of the predictions,'' J. Acoust. Soc. Am. **109**, 1123–1133.

Kompis, M., Dillier, N., Francois, J., Tinembart, J., and Häusler, R. (**1997**). ''New target-signal-detection schemes for multi-microphone noise-reduction systems for hearing aids,'' Proc. Annu. Int. Conf. IEEE Eng. Biol. Soc. **19**, 1990–1993.

Kompis, M., Feuz, P., François, J., and Tinembart, J. (**1999**). ''Multi-microphone digital-signal-processing system for research into noise reduction for hearing aids,'' Innovation Technol. Biol. Med. **20**, 201–206.

Kuhn, G. F. (**1977**). ''Model for the interaural time difference for the azimuthal plane,'' J. Acoust. Soc. Am. **62**, 157–167.

Peterson, P. M., Durlach, N. I., Rabinowitz, W. M., and Zurek, P. M. (**1987**). ''Multimicrophone adaptive beamforming for interference reduction in hearing aids,'' J. Rehabil. Res. Dev. **24**, 103–110.

Peterson, P. M., Wie, S. M., Rabinowitz, W. M., and Zurek, P. M. (**1990**). ''Robustness of an adaptive beamforming method for hearing aids,'' Acta Oto-Laryngol., Suppl. **469**, 85–90.

Van Compernolle, D. (**1990**). ''Hearing aids using binaural processing principles,'' Acta Oto-Laryngol., Suppl. **469**, 76–84.

van Hoesel, R. J. M., and Clark, G. M. (**1995**). ''Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients,'' J. Acoust. Soc. Am. **97**, 2498–2503.

Widrow, B., Glover, J. R., McColl, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, J. R., and Goodlin, R. C. (**1975**). ''Adaptive noise cancelling: Principles and applications,'' Proc. IEEE **63**, 1692–1716.