AVERAGE OUTPUT SNR OF THE MULTICHANNEL WIENER FILTER USING STATISTICAL ROOM ACOUSTICS

Toby Christian Lawin-Ore, Simon Doclo

University of Oldenburg, Department of Medical Physics and Acoustics - Signal Processing Group
{toby.chris.lawin.ore, simon.doclo}@uni-oldenburg.de

ABSTRACT

The performance of the multi-channel Wiener filter (MWF), which is often used for noise reduction in speech enhancement applications, depends on the noise field and on the acoustic transfer functions (ATFs) between the desired source and the microphone array. Recently, using statistical room acoustics an analytical expression for the spatially averaged output SNR, given the relative distance between the source and the microphone array, has been derived for the MWF in a diffuse noise field, requiring only the room properties to be known. In this paper, we show that this analytical expression can be extended to compute the average output SNR of the MWF for a specific microphone configuration, enabling to compare the performance of different microphone configurations, e.g. in an acoustic sensor network. Simulation results show that the average output SNR obtained using the statistical properties of ATFs is similar to the average output SNR obtained using simulated ATFs, therefore providing an efficient way to compare different microphone configurations.

***Index Terms*—** Multi-channel Wiener filter, statistical room acoustics

1. INTRODUCTION

The performance of most multi-microphone noise reduction algorithms depends on the acoustical scenario, i.e. the microphone configuration (number and positions of the microphones), the position of the desired source and the noise field. Therefore, one is often interested in computing theoretical performance measures, in order to be able to compare the performance for different acoustical scenarios. Moreover, since the position of the desired source is not always known a-priori, the *average performance* for a specific microphone array, computed by averaging the performance over all possible source positions in the room, can be a valuable performance measure to compare different microphone array configurations.

In [1], it has been shown that the theoretical performance of the multi-channel Wiener filter (MWF) depends on the noise correlation matrix and on the ATFs between the desired source and the microphone array. Therefore, if an estimated or simulated noise correlation matrix is available, the average performance of the MWF can be computed by using *measured* ATFs (which could be very time-consuming) or by using *numerically simulated* ATFs, e.g. using the image method [2] or room acoustics software.

On the other hand, the statistical properties of the ATFs can be described using *statistical room acoustics* (SRA) [3]. Recently,

This work was supported in part by the Research Unit FOR 1732 "Individualized Hearing Acoustics" and the Cluster of Excellence 1077 "Hearing4All", funded by the German Research Foundation (DFG).

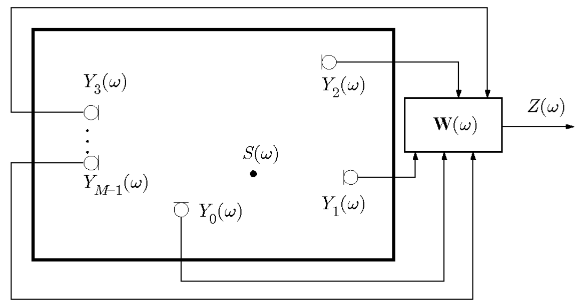

Figure 1: Configuration of a sensor network with M microphones.

assuming that the noise field is homogeneous and known, an analytical expression for the spatially averaged performance of the MWF, given the relative distance between the desired source and the microphone array, has been derived by incorporating the statistical properties of the ATFs into the theoretical formulas of several performance measures [4]. Simulation results have shown that the spatially averaged performance measures of the MWF, computed analytically using the statistical properties of ATFs, are similar to the results obtained using simulated ATFs. However, it should be realized that the analytical expressions derived in [4] do not allow to compute the average performance for a *specific* microphone array, since only the relative distance between the source and the microphone array is given (i.e. neither the position of the source nor the position of the microphone array is fixed).

In this paper, we derive an (approximate) analytical expression for the average output SNR of the MWF for a specific microphone configuration. The proposed expression can be used to easily compare the average performance of different microphone configurations in a specific room, without having to measure or numerically simulate the ATFs.

2. NOTATION AND SIGNAL MODEL

Consider the acoustical scenario depicted in Figure 1 with a desired source located at position $\mathbf{p}_s = [x_s \;\; y_s \;\; z_s]^T$ and M microphones located at positions $\mathbf{p}_m = [x_m \;\; y_m \;\; z_m]^T, m = 0 \ldots M-1$. The complete microphone array is described by the $3 \times M$-matrix $\mathbf{P}_{mic} = [\mathbf{p}_0 \cdots \mathbf{p}_{M-1}]$, where the distance $r_{mn} = \|\mathbf{p}_m - \mathbf{p}_n\|$ between the mth and the nth microphones in the array is assumed to be fixed (but not the microphone positions). Since the desired source and the microphone array can be located at any position in the room, we consider $\mathbf{p}_s$ and $\mathbf{P}_{mic}$ as stochastic variables[1]. We

[1]Note that a specific $\mathbf{P}^j_{mic}$ will be considered only in Section 4.2.

978-1-4799-0972-8/13/$31.00 ©2013IEEE

define the stochastic variable $\mathbf{P} = [\mathbf{P}_{mic}, \mathbf{p}_s]$ as the combination of the positions of the microphone array and the desired source and we define the relative distance between the desired source and the microphone array as

$$\mathbf{d} = \begin{bmatrix} d_0 \\ \vdots \\ d_{M-1} \end{bmatrix} = \begin{bmatrix} \|\mathbf{p}_0 - \mathbf{p}_s\| \\ \vdots \\ \|\mathbf{p}_{M-1} - \mathbf{p}_s\| \end{bmatrix}, \quad (1)$$

which is also a stochastic variable. Furthermore, we define the set of all realizations of $\mathbf{P}$ in the room as

$$\mathbf{Q} = \left\{ \mathbf{P}^{jk} = [\mathbf{P}^j_{mic}, \mathbf{p}^k_s] \quad \forall j, k \right\}, \quad (2)$$

where $\mathbf{P}^j_{mic}$ and $\mathbf{p}^k_s$ represent the jth and kth realization of $\mathbf{P}_{mic}$ and $\mathbf{p}_s$ respectively. We define $\mathbf{Q}^i \subset \mathbf{Q}$ as the subset of realizations with a specific relative distance $\mathbf{d}^i$ between the desired source and the microphone array, i.e.,

$$\mathbf{Q}^i = \left\{ \mathbf{P}^{jk} = [\mathbf{P}^j_{mic}, \mathbf{p}^k_s] \quad \forall j, k | \mathbf{d}^i \right\}. \quad (3)$$

Moreover, we define the spatial expectation operator $\mathcal{E}_{\mathbf{P}|\mathbf{d}^i}\{\cdot\}$ as the ensemble average over all realizations of $\mathbf{P}$ with relative distance $\mathbf{d}^i$(i.e. over the subset $\mathbf{Q}^i$) and the spatial expectation operator $\mathcal{E}_{\mathbf{P}|\mathbf{P}^j_{mic}}\{\cdot\}$ as the ensemble average over all realizations of $\mathbf{P}$ for a specific position $\mathbf{P}^j_{mic}$ of the microphone array.

For any realization of the position of the microphone array and the desired source, the vector $\mathbf{Y}(\omega) = [Y_0(\omega) \cdots Y_{M-1}(\omega)]^T$ consisting of the signals received by the microphones can be described in the frequency domain as

$$\mathbf{Y}(\omega) = \mathbf{H}(\omega)S(\omega) + \mathbf{V}(\omega) = \mathbf{X}(\omega) + \mathbf{V}(\omega), \quad (4)$$

where $\mathbf{H}(\omega) = [H_0(\omega) \cdots H_{M-1}(\omega)]^T$ denotes the stacked vector of the ATFs between the speech source $S(\omega)$ and the microphone array, and $\mathbf{X}(\omega)$ and $\mathbf{V}(\omega)$ represent the speech and the noise component of the microphone signals. The output signal $Z(\omega)$ is obtained by filtering and summing the microphone signals, i.e.,

$$Z(\omega) = \mathbf{W}^H(\omega)\mathbf{X}(\omega) + \mathbf{W}^H(\omega)\mathbf{V}(\omega) = Z_x(\omega) + Z_v(\omega), \quad (5)$$

where $\mathbf{W}(\omega) = [W_0(\omega) \cdots W_{M-1}(\omega)]^T$ represents the stacked vector of the filter coefficients, and $Z_x(\omega)$ and $Z_v(\omega)$ correspond to the estimated speech and residual noise component respectively. For conciseness the frequency-domain variable ω will be omitted where possible in the remainder of this paper. The correlation matrices are defined as

$$\mathbf{\Phi}_y = \mathcal{E}\{\mathbf{Y}\mathbf{Y}^H\},\ \mathbf{\Phi}_x = \mathcal{E}\{\mathbf{X}\mathbf{X}^H\},\ \mathbf{\Phi}_v = \mathcal{E}\{\mathbf{V}\mathbf{V}^H\}, \quad (6)$$

where $\mathcal{E}\{\cdot\}$ denotes the expected value operator.

In the remainder of this paper, a homogeneous noise field[2] is assumed, i.e. the noise correlation matrix can be expressed as $\mathbf{\Phi}_v = \phi_v \mathbf{\Gamma}_v$, where ϕ_v, denotes the noise power spectral density (PSD) and $\mathbf{\Gamma}_v$ the noise coherence matrix. Furthermore, a single desired speech source is assumed, for which the speech correlation matrix $\mathbf{\Phi}_x = \phi_s \mathbf{H}\mathbf{H}^H$ is a rank-one matrix, where ϕ_s represents the PSD of the source S, i.e. $\phi_s = \mathcal{E}\{|S|^2\}$.

[2]The assumption of a homogeneous noise field holds e.g. for a diffuse noise field or when the microphones are closely spaced.

3. MULTI-CHANNEL WIENER FILTERING

The multi-channel Wiener filter estimates the speech component X_{m_0} of the m_0th microphone, arbitrarily selected as the reference microphone, by minimizing the mean square error (MSE) cost function [5]

$$\xi(\omega) = \mathcal{E}\{|X_{m_0} - \mathbf{W}^H\mathbf{Y}|^2\}. \quad (7)$$

The solution of this minimization problem is given by

$$\mathbf{W}_{m_0} = \mathbf{\Phi}_y^{-1}\mathbf{\Phi}_x\mathbf{e}_{m_0}, \quad (8)$$

where $\mathbf{e}_{m_0}$ is an M-dimensional vector with the m_0th element equal to 1 and all other elements equal to 0. Assuming that the speech and the noise components are uncorrelated and using the matrix inversion lemma, it can be shown that (8) can be rewritten as [5]

$$\boxed{\mathbf{W}_{m_0} = \frac{\mathbf{\Gamma}_v^{-1}\mathbf{H}}{\frac{\phi_v}{\phi_s} + \rho} H^*_{m_0}, \qquad \rho = \mathbf{H}^H\mathbf{\Gamma}_v^{-1}\mathbf{H}} \quad (9)$$

where $\frac{\phi_s}{\phi_v}$ corresponds to the a-priori input SNR.

The output SNR of the MWF, which is defined as the PSD ratio of the speech and the noise component in the output signal, can be expressed using (9) as [5]

$$\boxed{\mathrm{SNR_{out}} = \frac{\mathcal{E}\{|Z_x|^2\}}{\mathcal{E}\{|Z_v|^2\}} = \frac{\mathbf{W}^H_{m_0}\mathbf{\Phi}_x\mathbf{W}_{m_0}}{\mathbf{W}^H_{m_0}\mathbf{\Phi}_v\mathbf{W}_{m_0}} = \frac{\phi_s}{\phi_v}\rho} \quad (10)$$

As can be seen, the output SNR of the MWF depends on the a-priori input SNR $\frac{\phi_s}{\phi_v}$, the ATFs $\mathbf{H}$ and the spatial characteristics of the noise field described by the noise coherence matrix $\mathbf{\Gamma}_v$.

4. SPATIALLY AVERAGED OUTPUT SNR OF MWF

In section 4.1, the analytical expression for the spatially averaged output SNR of the MWF, given a specific relative distance $\mathbf{d}^i$ between the desired source and the microphone array, will be briefly reviewed (for more details please refer to [4]). This expression will then be used in section 4.2 to derive an (approximate) analytical expression for the average output SNR of the MWF, given a specific position $\mathbf{P}^j_{mic}$ of the microphone array.

4.1. Spatially averaged output SNR of MWF given $\mathbf{d}^i$

Without loss of generality, for any position $\mathbf{P}$ of the desired source and the microphone array, the ATFs can be decomposed into direct and reverberant components, i.e.,

$$H_m(\mathbf{P}) = H_{m,d}(\mathbf{P}) + H_{m,r}(\mathbf{P}), \quad m = 0 \ldots M-1. \quad (11)$$

In statistical room acoustics [3] it is assumed that the reverberant sound field is diffuse and the direct component of the ATFs can be modeled as the free space Green's function, such that analytical expressions for the spatially averaged correlation between the direct components of the ATFs $\mathcal{E}_{\mathbf{P}|\mathbf{d}^i}\{H_{m,d}(\mathbf{P})H^*_{n,d}(\mathbf{P})\}$ and the reverberant components $\mathcal{E}_{\mathbf{P}|\mathbf{d}^i}\{H_{m,r}(\mathbf{P})H^*_{n,r}(\mathbf{P})\}$ can be derived. Furthermore, it is assumed that, for a given relative distance $\mathbf{d}^i$, the direct and the reverberant components of the ATFs are uncorrelated. By incorporating these statistical properties of the ATFs into the theoretical formula of the output SNR of the MWF in (10), an analytical expression for the spatially averaged output SNR, given a relative distance $\mathbf{d}^i$, i.e.

$$\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i) = \mathcal{E}_{\mathbf{P}|\mathbf{d}^i}\{\mathrm{SNR_{out}}(\mathbf{P})\}, \quad (12)$$

has been derived in [4] as

$$\boxed{\begin{aligned}\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i) =& \frac{\phi_s}{\phi_v}\sum_{m=1}^{M}\sum_{n=1}^{M}\breve{\gamma}_{mn}\Big(\frac{e^{j\frac{\omega}{c}(d_n^i - d_m^i)}}{(4\pi)^2 d_m^i d_n^i} \\ &+ \frac{1-\bar{\alpha}}{\pi\bar{\alpha}A}\frac{\sin\left(\frac{\omega}{c}r_{mn}\right)}{\frac{\omega}{c}r_{mn}}\Big)\end{aligned}} \quad (13)$$

where $\bar{\alpha}$ represents the average absorption coefficient of the room, A the total surface of the walls, c the speed of sound propagation and $\breve{\gamma}_{mn}$ denotes the (m,n)-element of the inverse coherence matrix $\mathbf{\Gamma}_v^{-1}$. If the reverberation time T_{60} is known, the average absorption coefficient can be approximated e.g. using Sabine's formula as $\bar{\alpha} = \frac{0.161V}{AT_{60}}$, where V is the volume of the room. Although simulation results in [4] have shown that the analytically computed spatially averaged output SNR $\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i)$ is close to results obtained through numerical simulations, it is important to realize that $\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i)$ is an expression for the average output SNR of all realizations in $\mathbf{Q}^i$ (i.e. given the relative distance $\mathbf{d}^i$), but is not equal to the output SNR for each realization in this subset, i.e.

$$\mathrm{SNR}_{\mathrm{out}}(\mathbf{P}^{jk}|\mathbf{d}^i) \neq \widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i) \quad \forall j,k. \quad (14)$$

This is due to the fact that for the computation of $\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i)$ neither the position of the microphone array nor the position of the desired source is fixed.

4.2. Spatially averaged output SNR of MWF given $\mathbf{P}_{mic}^j$

As a more useful performance measure, we would actually like to derive an analytical expression for the spatially averaged output SNR of the MWF for a specific position $\mathbf{P}_{mic}^j$ of the microphone array, i.e.

$$\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^j) = \mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}^j}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}. \quad (15)$$

However, note that it is not straightforward to derive this analytical expression similarly as in Section 4.1, since no analytical expressions for the spatially averaged correlations $\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}^j}\{H_{m,d}(\mathbf{P})H_{n,d}^*(\mathbf{P})\}$, $\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}^j}\{H_{m,r}(\mathbf{P})H_{n,r}^*(\mathbf{P})\}$ and $\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}^j}\{H_{m,d}(\mathbf{P})H_{n,r}^*(\mathbf{P})\}$ can be computed using statistical room acoustics. Nevertheless, we will show that an approximate analytical expression for $\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^j)$ can be derived using $\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i)$ in (13).

The average output SNR of the MWF in (15) can be written as

$$\begin{aligned}\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}^j}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\} &= \mathcal{E}_{\mathbf{p}_s}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^j, \mathbf{p}_s)\} \quad (16)\\ &= \int \mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^j, \mathbf{p}_s) f_{\mathbf{p}_s} d\mathbf{p}_s,\end{aligned}$$

where $f_{\mathbf{p}_s}$ is the probability density function of $\mathbf{p}_s$. For the derivation, we assume that the positions of the desired source $\mathbf{p}_s$ are uniformly distributed inside a large sphere centered around the microphone array. Although this assumption is generally not fulfilled in practice, the simulations in Section 5 show that the derived expression can also be (approximately) used in e.g. rectangular rooms. Now consider two different orientations $\mathbf{P}_{mic}^1$ and $\mathbf{P}_{mic}^2$ of the microphone array. For any source position $\mathbf{p}_s^{1k}$, there always exists a corresponding source position $\mathbf{p}_s^{2l}$ such that for homogeneous noise fields, the output SNRs of the MWF, for both combinations of the orientations of the microphone array and the desired source, are equal, i.e.,

$$\mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^1, \mathbf{p}_s^{1k}) = \mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^2, \mathbf{p}_s^{2l}). \quad (17)$$

Therefore, the spatially averaged output SNRs for both orientations of the microphone array are also equal, i.e.,

$$\begin{aligned}\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^1) &= \int \mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^1, \mathbf{p}_s) f_{\mathbf{p}_s} d\mathbf{p}_s \quad (18)\\ &= \int \mathrm{SNR}_{\mathrm{out}}(\mathbf{P}_{mic}^2, \mathbf{p}_s) f_{\mathbf{p}_s} d\mathbf{p}_s \\ &= \widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^2).\end{aligned}$$

Assuming that all realizations of $\mathbf{P}_{mic}$ can be considered as different orientations of the microphone array, the spatially averaged output SNR is equal for all realizations, such that

$$\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^j) = \mathcal{E}_{\mathbf{P}_{mic}}\{\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}\}, \forall j. \quad (19)$$

In addition, using the law of total expectation [6], i.e.

$$\begin{aligned}\mathcal{E}_{\mathbf{P}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\} &= \mathcal{E}_{\mathbf{P}_{mic}}\{\mathcal{E}_{\mathbf{P}|\mathbf{P}_{mic}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}\} \quad (20)\\ &= \mathcal{E}_{\mathbf{d}}\{\mathcal{E}_{\mathbf{P}|\mathbf{d}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}\},\end{aligned}$$

the spatially averaged output SNR can be computed as

$$\boxed{\begin{aligned}\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^j) =& \ \mathcal{E}_{\mathbf{d}}\{\mathcal{E}_{\mathbf{P}|\mathbf{d}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P})\}\} \\ =& \int \mathcal{E}_{\mathbf{P}|\mathbf{d}}\{\mathrm{SNR}_{\mathrm{out}}(\mathbf{P}) f_{\mathbf{d}} d\mathbf{d}\end{aligned}} \quad (21)$$

with $f_{\mathbf{d}}$ denoting the probability density function of $\mathbf{d}$. Solving this multi-dimensional integral by inserting (13) into (21) is a tedious problem. However, this integral can be approximated by a finite sum (moreover assuming that the relative distances are uniformly distributed) as

$$\boxed{\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}_{mic}^j) \approx \frac{1}{N_d}\sum_{i=1}^{N_d}\widetilde{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{d}^i)} \quad (22)$$

where N_d is the total number of considered relative distances.

5. SIMULATION RESULTS

5.1. Experimental setup

In order to validate the analytical expression derived in the previous section, we now present simulation results. In a room with dimensions 8 m×6 m×5 m and reverberation time $T_{60} = 0.4$ s (resulting in an average absorption coefficient $\bar{\alpha} \approx 0.4$), we consider an acoustical scenario with 3 spatially distributed nodes, where each node consists of four microphones with inter-microphone distance set to 1 cm. We consider 3 different microphone configurations with $N = 1, 2, 3$ nodes respectively, i.e. $M = 4, 8, 12$ microphones. A total number $N_s = 2000$ of positions $\mathbf{p}_s^k$ of the desired source have been uniformly distributed in the room. For each position of the desired source, impulse responses have been simulated using the image method [7], and the corresponding ATFs have been calculated. The length of the simulated impulse responses is 4096 samples and the sampling frequency $f_s = 16$ kHz. A simulated diffuse noise field has been used, for which the elements in the noise coherence matrix are equal to

$$\gamma_{mn}(\omega) = \frac{\sin\frac{\omega}{c}r_{mn}}{\frac{\omega}{c}r_{mn}}. \quad (23)$$

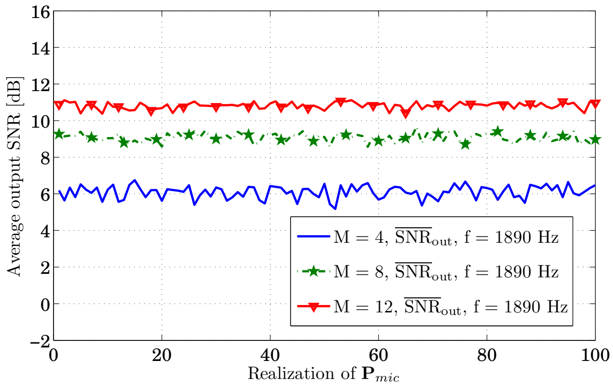

Figure 2: Average output SNR of the MWF for different realizations of $\mathbf{P}_{mic}$ and at frequency $f = 1890$ Hz.

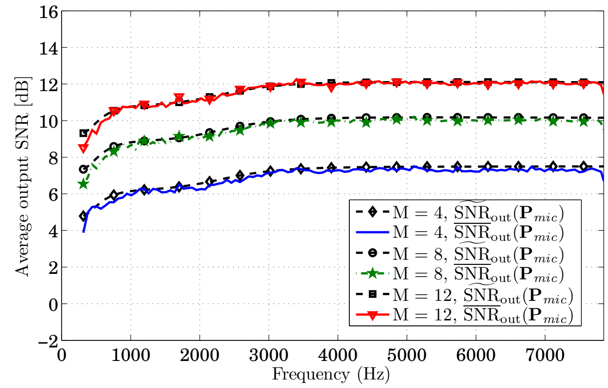

Figure 3: Average output SNR of the MWF for different microphone configurations in a diffuse noise field.

For each microphone array, the average output SNR of the MWF has been numerically computed as

$$\overline{\mathrm{SNR}}_{\mathrm{out}}(\mathbf{P}^{j}_{mic}) = \frac{1}{N_s}\sum_{k=1}^{N_s}\mathrm{SNR}_{\mathrm{out}}(\mathbf{P}^{j}_{mic}, \mathbf{p}^{k}_{s}). \quad (24)$$

The analytical average output SNRs computed using (22) have been approximated using a number $N_d = 2000$ of relative distances.

5.2. Results

In order to verify the assumption that the average output SNR is equal for all realizations of a specific microphone configuration (used in (19)), the 3 microphone configurations have been placed at 100 different positions in the room. For each position of the microphone configurations, the average output SNRs have been numerically computed using (24). Figure 2 shows the average output SNRs of the MWF at (an arbitrarily chosen) frequency 1890 Hz as a function of the realization. As one can see, the average output SNRs are fairly constant for different positions of the 3 microphone configurations with standard deviations of about 0.35 dB, 0.2 dB and 0.15 dB respectively. Similar results are obtained for other frequencies. The variation of the average output SNRs might be explained by the fact that instead of a spherical room a rectangular room has been used and the microphone configurations have been placed at locations in the middle of the room as well as near the walls.

Figure 3 shows the average output SNRs of the MWF for the 3 considered microphone configurations, computed numerically using simulated ATFs, together with the analytical approximations, calculated using (22). As can be seen, at all frequencies and for all microphone configurations, the analytical approximation derived using SRA is very close to the numerically computed average output SNR. Therefore, if the positions of the microphones and the room characteristics $(A,\bar{\alpha})$ are known and if the noise coherence matrix can be estimated, the statistical properties of ATFs can be used to analytically compute the average output SNR. Furthermore, the results clearly show that different microphone configurations lead to different average output SNRs. Therefore, for a specific room, the analytical expression for the average output SNR can be used to easily compare the performance of different microphone configurations. Moreover, Figure 3 shows the relation between the average output SNR and the number of microphones in an homogeneous noise field, i.e. the larger the number of microphones, the higher the average output SNR.

6. CONCLUSION

In this paper, the spatially averaged output SNR of the MWF given a relative distance between the desired source and the microphone array has been reviewed. It has been shown that, although this spatially averaged output SNR does not correspond to the output SNR of the MWF for specific positions of the microphone array and the desired source, it can be used to derive an analytical expression for the average output SNR of the MWF for a specific microphone configuration. Simulation results have shown that this analytical approximation is similar to the results obtained using simulated ATFs, providing an efficient way to compare the performance of different microphone configurations, e.g. in an acoustic sensor network, without having to measure or numerically simulate the ATFs.

7. REFERENCES

[1] A. Spriet, M. Moonen, and J. Wouters, "Robustness analysis of multi-channel Wiener filtering and Generalized Sidelobe Cancellation for multi-microphone noise reduction in hearing aid applications," *IEEE Trans. on Speech and Audio Processing*, vol. 13, no. 4, pp. 487–503, Jul. 2005.

[2] J. Allen, and D. Berkley, "Image method for efficiently simulating small room acoustics," *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[3] M. R. Schroeder, "Frequency correlation functions of frequency responses in rooms," *Journal of the Acoustical Society of America*, vol. 34, no. 12, pp. 1819–1823, Dec. 1962.

[4] T. C. Lawin-Ore, and S. Doclo, "Using statistical room acoustics for computing the spatially averaged performance of the multichannel Wiener filter based noise reduction, *in Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Aachen, Germany, pp. 145–148, Sep. 2012.

[5] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, *Acoustic beamforming for hearing aid applications*, chapter 9 in "Handbook on Array Processing and Sensor Networks," pp. 269–302, Wiley, 2010.

[6] N. A. Weiss, *A Course in Probability*, pp. 380–383, Addison-Wesley, 2005.

[7] E. A. P. Habets, "Room impulse response (RIR) generator," *Available: http://home.tiscali.nl/ehabets/rir generator.html*, Oct. 2008.