

# USING STATISTICAL ROOM ACOUSTICS FOR COMPUTING THE SPATIALLY AVERAGED PERFORMANCE OF THE MULTICHANNEL WIENER FILTER BASED NOISE REDUCTION

Toby Christian Lawin-Ore, Simon Doclo

University of Oldenburg, Institute of Physics, Signal Processing Group, Oldenburg, Germany  
 {toby.chris.lawin.ore, simon.doclo}@uni-oldenburg.de

## ABSTRACT

The theoretical 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 microphones. To compute this theoretical performance, either time-consuming measurements of the ATFs and the noise field or simulation of the ATFs, e.g. using the image model, need to be performed. Recently, an analytical expression for the spatially averaged output SNR of the MWF has been proposed, requiring only the room properties and the source-microphone distances to be known. In this paper, we derive similar analytical expressions for other performance measures, namely for the spatially averaged mean square error (MSE), and speech distortion at the reference microphone.

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

## 1. INTRODUCTION

For every speech enhancement algorithm it is of significant interest to be able to compute its theoretical performance, e.g. output SNR, for different acoustical scenarios (microphone configuration, source position, noise field). This e.g. enables to compare the performance of different microphone configurations, such that the microphone configuration yielding the best performance can be selected.

In [1] a theoretical performance analysis of the multi-channel Wiener filter (MWF) has been presented. It has been shown that the output SNR, the mean square error (MSE), the noise reduction and the speech distortion of the MWF can be computed when the noise correlation matrix and the ATFs between the desired source and the microphones are known. Hence, for every source-microphones configuration, the theoretical performance of the MWF can be computed using measured or simulated noise correlation matrices and ATFs. However, if we want to compare the performance for a large number of source-microphones configurations (and assuming that an estimated or simulated noise correlation matrix is available), then either a large number of ATFs need to be measured, which could be very time-consuming, or the performance of the MWF can be *numerically* simulated, by simulating the ATFs e.g. using the image model [2] or room acoustics software.

On the other hand, the statistical properties of the ATFs can be described using *statistical room acoustics* (SRA), which has also been used to derive analytical expressions for performance measures. In [4] the robustness of an equalization technique has been

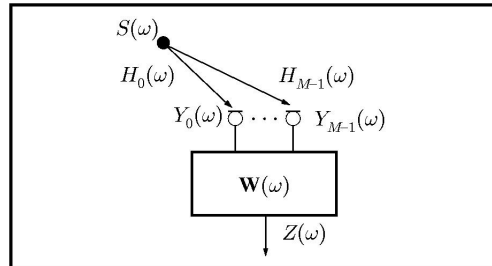


Fig. 1. Configuration of an array with  $M$  microphones.

analyzed using SRA, and in [5] a method to predict the SNR improvement of a delay-and-sum beamformer with two microphones using the statistical properties of ATFs has been presented.

Recently, an analytical expression for the spatially averaged output SNR of the MWF has been derived by incorporating statistical properties of the ATFs in the theoretical formula of the output SNR [6]. Simulation results have shown that the spatially averaged output SNR computed analytically using the statistical properties of ATFs is similar to the results obtained using simulated ATFs, providing an easy and fast way to compute the spatially averaged output SNR of the MWF. In this paper, we extend the computation of the theoretical performance of the MWF using SRA to the MSE, the noise reduction and the speech distortion at the reference microphone. Simulation results show the good estimation accuracy of all performance measures computed using SRA except for the noise reduction.

## 2. SIGNAL MODEL AND CONFIGURATION

Figure 1 depicts the configuration of  $M$  microphones located at positions  $\mathbf{p}_m = [x_m \ y_m \ z_m]^T$ ,  $m = 0 \dots M-1$ , and a single speech source  $S(\omega)$  located at position  $\mathbf{p}_s = [x_s \ y_s \ z_s]^T$ . The complete microphone array configuration can be described by the  $3 \times M$ -matrix  $\mathbf{P}_{mic} = [\mathbf{p}_0 \dots \mathbf{p}_{M-1}]$ . We define the relative distance  $\mathbf{d}$  between the speech source and the microphones as

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

The  $m$ th microphone signal  $Y_m(\omega)$  can be described in the frequency domain as

$$\begin{aligned} Y_m(\omega) &= H_m(\omega)S(\omega) + V_m(\omega), \quad m = 0 \dots M-1 \\ &= X_m(\omega) + V_m(\omega), \end{aligned} \quad (2)$$

This work was supported by the Research Unit FOR 1732 "Individualized Hearing Acoustics", funded by the German Research Foundation (DFG).

where  $H_m(\omega)$  represents the ATF between the speech source  $S(\omega)$  and the  $m$ th microphone, and  $X_m(\omega)$  and  $V_m(\omega)$  represent the speech and the noise component in the  $m$ th microphone signal. We define the  $M$ -dimensional stacked signal vector  $\mathbf{Y}(\omega)$  as

$$\mathbf{Y}(\omega) = \begin{bmatrix} Y_0(\omega) \\ \vdots \\ Y_{M-1}(\omega) \end{bmatrix}, \quad (3)$$

which can be written as  $\mathbf{Y}(\omega) = \mathbf{X}(\omega) + \mathbf{V}(\omega)$ , where the vectors  $\mathbf{X}(\omega)$  and  $\mathbf{V}(\omega)$  are defined similarly as  $\mathbf{Y}(\omega)$ . The output signal  $Z(\omega)$  is obtained by filtering and summing the microphone signals, i.e.,

$$Z(\omega) = \mathbf{W}^H(\omega)\mathbf{X}(\omega) + \mathbf{W}^H(\omega)\mathbf{V}(\omega) = Z_x(\omega) + Z_v(\omega), \quad (4)$$

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.

### 3. MULTI-CHANNEL WIENER FILTERING

The concept of multi-channel Wiener filtering (MWF) is based on estimating the speech component  $X_{m_0}$  of the  $m_0$ th microphone, arbitrarily selected as the reference microphone. The MWF produces a minimum-mean-square error (MMSE) estimate by minimizing the MSE cost function [7]

$$\xi(\mathbf{W}(\omega)) = \mathcal{E}\left\{\left|X_{m_0}(\omega) - \mathbf{W}^H(\omega)\mathbf{Y}(\omega)\right|^2\right\}, \quad (5)$$

where  $\mathcal{E}\{\cdot\}$  denotes the expected value operator. The solution of this minimization problem is given by

$$\mathbf{W}_{m_0}(\omega) = \Phi_y^{-1}(\omega)\Phi_x(\omega)\mathbf{e}_{m_0}, \quad (6)$$

with  $\Phi_y(\omega) = \mathcal{E}\{\mathbf{Y}(\omega)\mathbf{Y}^H(\omega)\}$ ,  $\Phi_x(\omega) = \mathcal{E}\{\mathbf{X}(\omega)\mathbf{X}^H(\omega)\}$  the noisy and clean speech correlation matrix, and  $\mathbf{e}_{m_0}$  an  $M$ -dimensional vector with the  $m_0$ th element equal to 1 and all other elements equal to 0, selecting the column that corresponds to the reference microphone.

Assuming that the speech and the noise components are uncorrelated, the correlation matrix  $\Phi_y(\omega)$  can be expressed as

$$\Phi_y(\omega) = \Phi_x(\omega) + \Phi_v(\omega), \quad (7)$$

where  $\Phi_v(\omega)$  represents the noise correlation matrix, i.e.,  $\Phi_v(\omega) = \mathcal{E}\{\mathbf{V}(\omega)\mathbf{V}^H(\omega)\}$ . Using a robust VAD, the correlation matrix  $\Phi_y(\omega)$  can be estimated during speech + noise periods, while the correlation matrix  $\Phi_x(\omega)$  can be estimated during speech pauses.

For conciseness the frequency-domain variable  $\omega$  will be omitted where possible in the remainder of this paper.

### 4. THEORETICAL PERFORMANCE OF MWF

If we make the following two assumptions:

1. a single desired speech source is assumed, for which the speech correlation matrix  $\Phi_x = \phi_s \mathbf{H}\mathbf{H}^H$  is a rank-one matrix, where  $\phi_s$  represents the power spectral density (PSD) of the source  $S$ , i.e.  $\phi_s = \mathcal{E}\{|S|^2\}$  and  $\mathbf{H} = [H_0 \cdots H_{M-1}]^T$  is the stacked vector of the ATFs,
2. a homogeneous<sup>1</sup> noise field is considered, i.e.  $\Phi_v = \phi_v \Gamma_v$ , where  $\phi_v = \Phi_v(m, m)$ ,  $\forall m$  denotes the noise PSD and  $\Gamma_v$  the noise coherence matrix,

<sup>1</sup>The assumption of a homogeneous noise field always holds for a diffuse noise field or when the microphones are closely spaced.

then the filter  $\mathbf{W}_{m_0}$  in (6) can be written as [7]

$$\mathbf{W}_{m_0} = \frac{\Gamma_v^{-1}\mathbf{H}}{\frac{\phi_v}{\phi_s} + \rho} H_{m_0}^*, \quad (8)$$

where

$$\rho = \mathbf{H}^H \Gamma_v^{-1} \mathbf{H}, \quad (9)$$

and  $\frac{\phi_s}{\phi_v}$  corresponds to the a-priori input SNR. Note that the output SNR of the MWF is computed as [6]

$$\text{SNR}_{\text{out}} = \frac{\phi_s}{\phi_v} \rho. \quad (10)$$

The minimum MSE achieved by using the MWF is obtained by replacing (8) into (5), i.e.,

$$\xi(\mathbf{W}_{m_0}) = \frac{\phi_v |H_{m_0}|^2}{\left(\frac{\phi_v}{\phi_s} + \rho\right)}. \quad (11)$$

The MWF adds distortion to the estimated speech component. In order to evaluate the amount of speech distortion, we define the speech distortion at the reference microphone as

$$\text{SD}^{m_0} = \frac{\mathcal{E}\{|Z_x|^2\}}{\mathcal{E}\{|X_{m_0}|^2\}} = 1 / \left(1 + \frac{1}{\frac{\phi_s}{\phi_v} \rho}\right)^2. \quad (12)$$

Similarly to the speech distortion, we define the noise reduction factor at the reference microphone as

$$\text{NR}^{m_0} = \frac{\mathcal{E}\{|V_{m_0}|^2\}}{\mathcal{E}\{|Z_v|^2\}} = \rho \left(1 + \frac{1}{\frac{\phi_s}{\phi_v} \rho}\right)^2 / |H_{m_0}|^2. \quad (13)$$

As one can see, if the a-priori input SNR  $\frac{\phi_s}{\phi_v}$ , and the spatial characteristics  $\rho$ , i.e. the ATFs  $\mathbf{H}$  between the source and the microphones, and the spatial characteristics of the noise field described by the noise coherence matrix  $\Gamma_v$  are known, the noise reduction, the speech distortion and the MSE of the MWF can be calculated.

### 5. PERFORMANCE OF MWF USING STATISTICAL PROPERTIES OF ATFS

The theory of statistical room acoustics is based on the assumption that the phase and the amplitude of reverberant plane waves arriving at a point in a room are close to random. The resulting reverberant sound field can then be considered as uniformly distributed in the entire room [3]. This model of the reverberant sound field is valid only if a set of conditions are satisfied (for more details, we refer to [3]).

#### 5.1. Statistical properties of ATFs

Without loss of generality, the vector  $\mathbf{H}$  containing the ATFs between the source located at position  $\mathbf{p}_s$  and the  $M$  microphones located at the positions  $\mathbf{P}_{mic}$  can be decomposed as

$$\mathbf{H}(\boldsymbol{\theta}) = \mathbf{H}_d(\boldsymbol{\theta}) + \mathbf{H}_r(\boldsymbol{\theta}), \quad (14)$$

where  $\boldsymbol{\theta} = [\mathbf{p}_s, \mathbf{P}_{mic}]$  and  $\mathbf{H}_d(\boldsymbol{\theta})$  and  $\mathbf{H}_r(\boldsymbol{\theta})$  are the vectors corresponding to the direct and the reverberant component of the ATFs.

We define the spatial expectation operator  $\mathcal{E}_\theta\{\cdot\}$  as the ensemble average over all realizations of  $\boldsymbol{\theta}$ . Using SRA, the following statistical properties for the ATFs are then given [3]:

A1 For a fixed relative distance  $\mathbf{d}$  between source and microphones, the direct path components are independent of the realization of  $\theta$ , i.e.,

$$\mathcal{E}_\theta\{H_{m,d}(\theta)H_{n,d}^*(\theta)|d_m, d_n\} = \frac{e^{j\frac{\omega}{c}(d_n-d_m)}}{(4\pi)^2 d_m d_n} \quad \forall m, n. \quad (15)$$

A2 The spatially expected correlation between the reverberant components of the ATF of the  $m$ th and the  $n$ th microphone is independent of  $\mathbf{d}$  and is given by

$$\mathcal{E}_\theta\{H_{m,r}(\theta)H_{n,r}^*(\theta)\} = \frac{1-\bar{\alpha}}{\pi\bar{\alpha}A} \frac{\sin\left(\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|\right)}{\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|} \quad \forall m, n, \quad (16)$$

where  $A$  is the total surface of the walls and  $\bar{\alpha}$  is the average absorption coefficient. If the reverberation time  $T_{60}$  is known, the average absorption coefficient can be approximated using Sabine's formula, i.e.  $\bar{\alpha} = \frac{0.161V}{AT_{60}}$ .

A3 The direct and the reverberant components of the ATFs are uncorrelated, i.e.,

$$\mathcal{E}_\theta\{H_{m,d}(\theta)H_{n,r}^*(\theta)|d_m, d_n\} = 0, \quad \forall m, n. \quad (17)$$

## 5.2. Spatially averaged performance measures of MWF

The objective of this subsection is to derive analytical expressions for the MSE, the noise reduction and the speech distortion of the MWF using the statistical properties of the ATFs.

Using (9) and (14), the spatial characteristics  $\rho$  for each realization  $\theta$  can be written as

$$\rho(\theta) = \sum_{m=1}^M \sum_{n=1}^M \check{\gamma}_{mn} (H_{d,m}^*(\theta)H_{d,n}(\theta) + H_{d,m}^*(\theta)H_{r,n}(\theta) + H_{r,m}^*(\theta)H_{d,n}(\theta) + H_{r,m}^*(\theta)H_{r,n}(\theta)), \quad (18)$$

where  $\check{\gamma}_{mn}$  is the  $(m, n)$ -element of the matrix  $\Gamma_v^{-1}$ . Using (15), (16) and (17), the spatially averaged value of  $\rho$  given  $\mathbf{d}$  (relative distance between source and microphones) is then equal to [6]

$$\mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\} = \sum_{m=1}^M \sum_{n=1}^M \check{\gamma}_{mn} \left( \frac{e^{j\frac{\omega}{c}(d_n-d_m)}}{(4\pi)^2 d_m d_n} + \frac{1-\bar{\alpha}}{\pi\bar{\alpha}A} \frac{\sin\left(\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|\right)}{\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|} \right). \quad (19)$$

This value depends only on the relative distance between the source and the microphones and on the room properties ( $A, \bar{\alpha}$ ). While in [6], an analytical expression for the spatially averaged output SNR has been derived using (19) and without any approximation, an approximation is required in order to derive similar expressions for the spatially averaged MSE, noise reduction and speech distortion at the reference microphone.

Using (11), the spatially averaged MSE given  $\mathbf{d}$  can be expressed as

$$\mathcal{E}_\theta\{\xi(\mathbf{W}_{m_0}(\theta))|\mathbf{d}\} = \mathcal{E}_\theta\left\{ \frac{\phi_v |H_{m_0}|^2}{\left(\frac{\phi_v}{\phi_s} + \rho\right)} \middle| \mathbf{d} \right\}. \quad (20)$$

To compute this expectation of a function of two random variables  $\rho$  and  $|H_{m_0}|^2$ , we will use an approximation based on the Taylor expansion. In general, let us consider the random variables  $X$  and  $Y$  with  $\mu_x = \mathcal{E}\{X\}$  and  $\mu_y = \mathcal{E}\{Y\}$ . The Taylor expansion of a differentiable function  $f(x, y)$  around  $(\mu_x, \mu_y)$  is given by

$$f(x, y) = f(\mu) + f'_x(\mu)(x - \mu_x) + f'_y(\mu)(y - \mu_y) + \hat{f}(x, y), \quad (21)$$

where  $\hat{f}(x, y)$  represents a function of the higher-order partial derivatives of  $f(x, y)$ . Assuming that all partial derivatives, except the first-order partial derivatives, can be neglected at  $(\mu_x, \mu_y)^2$ , then  $f(x, y)$  can be approximated by the first-order Taylor expansion, i.e.,  $f(x, y) \approx f(\mu) + f'_x(\mu)(x - \mu_x) + f'_y(\mu)(y - \mu_y)$ . Taking the expectation of both sides of the approximated Taylor expansion yields  $\mathcal{E}\{f(x, y)\} \approx f(\mu_x, \mu_y)$ . However, it should be noted that this approximation does not hold for all functions, i.e. if the higher-order derivatives can not be neglected at the expansion point.

Although the first-order Taylor expansion might not be a good approximation for all functions, we still propose to use it for computing the spatially averaged MSE, i.e.,

$$\mathcal{E}_\theta\{\xi(\mathbf{W}_{m_0}(\theta))|\mathbf{d}\} \approx \frac{\phi_v \mathcal{E}_\theta\{|H_{m_0}(\theta)|^2|\mathbf{d}\}}{\frac{\phi_v}{\phi_s} + \mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\}}, \quad (22)$$

where  $\mathcal{E}_\theta\{|H_{m_0}(\theta)|^2|\mathbf{d}\}$  and  $\mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\}$  are given by (15) and (19) respectively. Similarly using (12) and its first-order Taylor expansion, the spatially averaged speech distortion can be approximated by

$$\mathcal{E}_\theta\{\text{SD}^{m_0}(\theta)|\mathbf{d}\} \approx 1 / \left( 1 + \frac{1}{\frac{\phi_s}{\phi_v} \mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\}} \right)^2. \quad (23)$$

Finally, using (13) and its first-order Taylor expansion, the noise reduction can be approximated by

$$\mathcal{E}_\theta\{\text{NR}^{m_0}(\theta)|\mathbf{d}\} \approx \frac{\frac{\phi_s^2}{\phi_v} \frac{1}{\mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\}} + \mathcal{E}_\theta\{\rho(\theta)|\mathbf{d}\} + \frac{2\phi_v}{\phi_s}}{\mathcal{E}_\theta\{|H_{m_0}(\theta)|^2|\mathbf{d}\}}. \quad (24)$$

Although an analytical expression has been derived for the spatially averaged noise reduction, simulation results (cf. section6) show that this expression does not yield a good approximation.

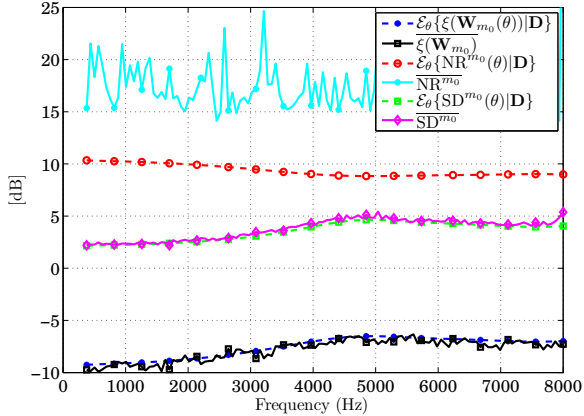
## 6. SIMULATION RESULTS

### 6.1. Experimental setup

In order to validate the theoretical results derived in the previous section, we consider the acoustical scenario depicted in Figure 1, which consists of a linear microphone array with  $M = 3$  microphones. The distance between the microphones is set to 4 cm. The desired source is located at endfire of the microphone array such that  $\mathbf{d} = [1.36 \ 1.40 \ 1.44]^T$ . In a room with dimensions 7 m  $\times$  5 m  $\times$  3.5 m and reverberation time  $T_{60} = 0.25$  s (resulting in an average absorption coefficient  $\bar{\alpha} = 0.51$ ), different realizations of  $\theta$  given the constant relative distance  $\mathbf{d}$  have been generated by rotating and translating the source-microphones configuration, and considering only the realizations of  $\theta$  that are located in the interior of the room and half a wavelength away from the walls, satisfying the conditions in [3]. For each realization, impulse responses have been simulated using the image model [2], and the corresponding ATFs have been calculated. The length of the simulated impulse responses is 4096 samples and the sampling frequency  $f_s = 16000$  Hz. Diffuse noise has been used and the noise coherence matrix was theoretically computed using

$$\gamma_{mn}(\omega) = \frac{\sin\left(\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|\right)}{\frac{\omega}{c}\|\mathbf{p}_m-\mathbf{p}_n\|}. \quad (25)$$

<sup>2</sup>Although this can not be proven, since no analytical expressions for the higher order Taylor terms are available, simulation results for multiple acoustical scenarios have shown that this assumption can be made.



**Fig. 2.** Average performance measures of MWF calculated using a Monte Carlo simulation with 10000 realizations and spatially averaged performance measures computed using SRA.

For a total number of realizations  $N$ , the average performance measures are numerically computed as

$$\overline{\text{PM}}(N) = \frac{1}{N} \sum_{i=1}^N \text{PM}(\tilde{\theta}_i), \quad (26)$$

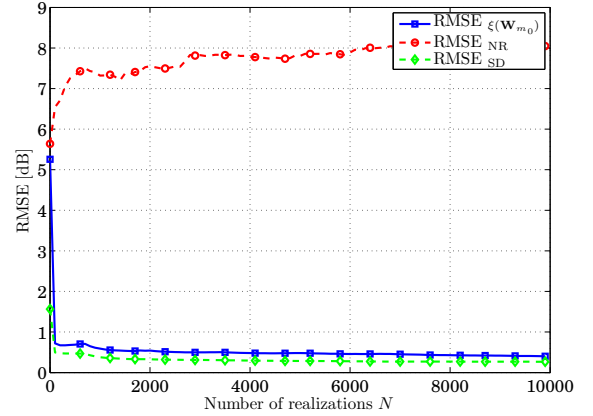
where  $\tilde{\theta}_i$  corresponds to a single realization of  $\theta$  and PM represents either  $\text{NR}^{m_0}$ ,  $\text{SD}^{m_0}$  or  $\xi(\mathbf{W}_{m_0})$ . Without loss of generality, the a-priori input SNR is assumed to be frequency flat.

## 6.2. Results

Figure 2 shows  $\overline{\text{NR}}^{m_0}$ ,  $\overline{\text{SD}}^{m_0}$  and  $\overline{\xi}(\mathbf{W}_{m_0})$  computed numerically using simulated ATFs by means of a Monte Carlo simulation with  $N = 10000$  realizations, together with the spatially averaged performance measures  $\mathcal{E}_\theta\{\xi(\mathbf{W}_{m_0}(\theta)|\mathbf{d})\}$ ,  $\mathcal{E}_\theta\{\text{NR}^{m_0}(\theta)|\mathbf{d}\}$  and  $\mathcal{E}_\theta\{\text{SD}^{m_0}(\theta)|\mathbf{d}\}$  calculated analytically using (24), (23) and (22). As one can see, the spatially averaged speech distortion and minimum MSE of the MWF computed using SRA are very close to the average measures obtained using simulated ATFs. Therefore, if the source and microphone positions 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 express the average speech distortion and minimum MSE of the MWF. Unfortunately, the spatially averaged noise reduction computed using SRA does not correspond well to the averaged noise reduction computed using simulated ATFs. The difference between both values can be explained by the fact that the higher-order partial derivatives in (21) can not be neglected at the expansion point, leading to a poor approximation. In order to improve the approximation of the spatially averaged noise reduction using SRA, the second-order Taylor expansion could be used, however, requiring the statistical properties of the variance of  $\rho(\theta)$  and  $|H_{m_0}(\theta)|^2$ .

Figure 3 depicts the root mean square error (RMSE) between the spatially averaged performance measures computed using SRA and the average performance measures computed using simulated ATFs as a function of the number of realizations  $N$ . The RMSE for each performance measure is calculated as

$$\text{RMSE}_{\text{PM}}(N) = \sqrt{\sum_{\omega} |\mathcal{E}_\theta\{\text{PM}(\theta)|\mathbf{d}\} - \overline{\text{PM}}(N)|^2}. \quad (27)$$



**Fig. 3.** Root mean square error between simulated and analytical results obtained using statistical room acoustics.

As expected, the larger the number of realizations, the smaller the RMSE for all performance measures except for the noise reduction which does not converge. For a very large number of realizations, the RMSE of the speech distortion and the minimum MSE converge to nearly zero, showing the good estimation accuracy of both spatially averaged measures using SRA.

## 7. CONCLUSION

In this paper we have derived analytical expressions for the spatially averaged MSE, noise reduction and speech distortion of the MWF using statistical room acoustics. The proposed analytical expressions depend on the room dimensions, source-microphones distances and reverberation time. Simulation results have shown that the speech distortion and the minimum MSE computed using the statistical properties of ATFs are similar to the results obtained using simulated ATFs, providing an easy and fast way to compute both spatially averaged performance measures of the MWF. For the noise performance measure, a further analysis is required.

## 8. 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] E. A. P. Habets, "Room impulse response (RIR) generator," [http://home.tiscali.nl/ehabets/rir\\_generator.html](http://home.tiscali.nl/ehabets/rir_generator.html), Oct. 2008.
- [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] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment. Robustness results," *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 3, pp. 311–319, Aug. 2002.
- [5] M. Kompis, and N. Dillier, "Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. I. Prediction of the signal-to-noise-ratio improvement," *Journal of the Acoustical Society of America*, vol. 109, no. 3, pp. 1123–1133, Mar. 2001.
- [6] T. C. Lawin-Ore, and S. Doclo, "Using statistical room acoustics for analysing the output SNR of the MWF in acoustic sensor networks," *Proc. of EUSIPCO 2012*, Bucharest, Romania, Aug. 2012.
- [7] 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.