Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Power output regularization in the active reproduction of sound fields in rooms

Nick Stefanakis, Sotiris Dalianis, Tilemachos Karatzas and George Cambourakis

National Technical University of Athens, School of Electrical and Computer Engineering, Heroon Polytechniou 9, 157 73 Athens, Greece
nstefan@mobile.ntua.gr

In this paper we address the problem of using a multi-channel active control system in order to reproduce a harmonic sound field in a large part of the volume of a reverberant room. The problems associated with the calculation of the inverse system matrix are confronted by introducing a term that is proportional to the sound power-output of the system in the cost function that is obtained by the multiple point method. Simulation results show that this technique results to a better conditioning of the system matrix at low frequencies, comparing to other traditional regularization techniques. Moreover, it is shown that this method can be employed to increase the spatial robustness of the control sensor array inside the listening room.

1 Introduction

With the recent advances in digital signal processing, active control of sound has found new applications in the sound engineering problems of the small rooms. By specifying a set of desired signals at a number of receiving positions, active control can be applicable to tasks such as sound equalization and sound field reproduction [1-3]. These tasks are mainly based on the conventional multiple point technique which minimizes (in a least-squares sense) a cost function that expresses the difference between the desired complex sound pressure and the sound pressure that is actually reproduced at a number of sampling points inside the room. The optimum source strengths are calculated by inverting the system matrix that contains the acoustic impedance from the loudspeakers to the receiving positions inside the room. In practise, it is crucial to use a regularization technique in order to avoid the problems associated with the inversion of a badly conditioned or an under-determined system matrix [4].

It is well known that the properties of the system matrix, such as the maximum eigenvalue and the condition number, are straightforwardly related to the room natural dynamics. At the low frequencies, the frequency response of the acoustic path is heavily depended on frequency and on the position of the loudspeaker and the microphone relative to the boundary of the room. In this paper, we propose a technique which can be used in order to overcome the instabilities that characterize the performance of a sound field reproduction system at the low modal density region of a lightly damped rectangular room. This method demands the knowledge of the real part of the acoustic impedance from each source to the other inside the room.

In [5], it was shown how the knowledge of these acoustic impedances can be used in order to achieve power output regularization. This regularization technique was examined in the case of equalization in a rectangular room. Equalization was performed with the generation of a plane wave travelling along the axis of the room. It was shown that power output regularization can lead to increment of the spatial robustness of equalization. This increment was realized as a smoother decay of the equalization performance away from the control sensors. In this paper, we consider a sound field reproduction system inside the rectangular room. Plane waves travelling at all possible directions from 0^o to 360^o have to be generated now. Apart from the effect in the spatial robustness, we present further benefits. In particular, it can be seen that power output regularization can be used in order to compensate for the non-flat system dynamics and can lead to a modified system matrix with important benefits compared to the original one. These benefits are shown in terms of increment of a convergence rate of a modified adaptive algorithm used for the adaptation of the source strengths in the frequency domain.

2 Control model

Suppose that it is desired to control the sound field in a spatial region inside an enclosure that is surrounded by L reproduction sources. The pressure in this spatial region is sampled by M monitor sensors placed at $\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_M\}$, which provide a measure of the performance of reproduction at the entire listening space. The pressure at the monitoring sensors subject to the L source excitations can be written as [1]

$$\mathbf{p}_M = \mathbf{Z}_M \mathbf{q} \tag{1}$$

where $\mathbf{p}$ is a column vector with the M complex sound pressures at the monitor sensors [Pa], $\mathbf{q}$ is a column vector with the complex strengths of the L sources [m^3/sec], and $\mathbf{Z}_M$ is an MxL matrix that carries the acoustic transfer functions from each source to the each field point at $\mathbf{r}_m$. From those M monitor sensors, we choose a small group of N control sensors covering a small spatial region that is centred inside the listening area at $\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N\}$. It is assumed here that this compact control sensor array represents a more feasible sound reproduction system that occupies less space and requires less equipment and input channels. The compact system is thus informed about the performance of reproduction in the controlled region by the difference between the desired sound pressure and the actual reproduced sound pressure at the control sensors as,

$$\mathbf{e} = \mathbf{p}_d - \mathbf{Z}\mathbf{q}\,, \tag{2}$$

where $\mathbf{p}_d$ is the vector with the desired sound pressures at the N control sensors, and $\mathbf{Z}$ is the compact system matrix carrying the acoustic impedances from the L sources to the N control sensors.

2.1 The regularization technique

The proposed control approach suggests the use of a cost function defined as

$$J^{(\lambda)} = \mathbf{e}^H \mathbf{e} + \lambda \mathbf{q}^H \mathbf{W} \mathbf{q}\,. \tag{3}$$

Here λ is a real positive scalar that weights the contribution of the penalty term in the cost function, the quantity $\mathbf{q}^H\mathbf{W}\mathbf{q}$ expresses the total sound power emitted by the reproduction sources [W], and $\mathbf{W}$ is a symmetric and positive definite matrix with W_{ij} representing the half of the real part of the transfer function from source i to j [6]. For distributed sources, each element of the matrix $\mathbf{W}$ can be calculated with proper integration of the transfer function on the surface of each source. The second term at the ride side of Eq. (3) is always a non negative quantity (the total power output of the system) which means that matrix $\mathbf{W}$ is positive definite. For such a matrix the Cholesky decomposition is always possible and it can be written as

$$\mathbf{W} = \mathbf{L}^H \mathbf{L}\,. \tag{4}$$

Eq. (3) can thus be written in the general from of Tikhonov regularization [4] by noting that $\mathbf{q}^H\mathbf{W}\mathbf{q} = (\mathbf{L}\mathbf{q})^H(\mathbf{L}\mathbf{q})$. The cost function defined in Eq. (3) implies the addition of an λ-weighted sound power-output penalty term instead of the source effort penalty term used in standard regularization as [4]

$$J^{(\mu)} = \mathbf{e}^H\mathbf{e} + \mu\mathbf{q}^H\mathbf{q}\,. \tag{5}$$

Under the condition that $\lambda\mathbf{W} + \mathbf{Z}^H\mathbf{Z}$ is positive definite and invertible, the optimal vector that minimizes $J^{(\lambda)}$ can be found by

$$\mathbf{q}_o^{(\lambda)} = (\lambda\mathbf{W} + \mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{p}_d\,. \tag{6}$$

The optimum source strengths derived here should be compared with those obtained by standard regularization,

$$\mathbf{q}_o^{(\mu)} = (\mu\mathbf{I} + \mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{p}_d\,. \tag{7}$$

It can be seen that the identity matrix **I** has been replaced by the fully populated matrix **W** which changes dynamically with frequency.

The achieved quality of reproduction for each of the two optimum source strengths is measured over the entire listening space with the use of the *M* monitor sensors. Similar to Eq. (2), the error at the monitor sensors can be expresses as

$$\mathbf{e}_M = \mathbf{p}_{d,M} - \mathbf{Z}_M\mathbf{q}\,, \tag{8}$$

where $\mathbf{p}_{d,M}$ is now the vector with the desired complex pressures at all the monitor sensors. The quality of the actual reproduced sound field over the monitor sensors is quantified by the global reproduction error, which is defined as

$$E_{LS}^{(M)} = (\mathbf{e}_M^H\mathbf{e}_M)^{1/2} = \left(\frac{(\mathbf{p}_{d,M} - \mathbf{Z}_M\mathbf{q}_o)^H(\mathbf{p}_{d,M} - \mathbf{Z}_M\mathbf{q}_o)}{\mathbf{p}_{d,M}{}^H\mathbf{p}_{d,M}}\right)^{1/2}. \tag{9}$$

Here $\mathbf{q}_o$ is the optimum source strength vector which is derived either from Eq. (5) or (6), for power-output penalty and effort penalty regularization respectively. For the case of a plane wave, a value of $E_{LS}^{(M)}$ below 0.5 denotes that the deviations between the reproduced sound pressure and the desired one are within ± 6 dB, and this value can be used as a criterion for a good global sound reproduction result.

2.2 The ideal system

The calculation of the source strength vectors $\mathbf{q}_o^{(\lambda)}$ and $\mathbf{q}_o^{(\mu)}$ derived by Eqs. (5) and (6) is based on the information provided only by the control sensors. This means that the *M* monitor sensors don't actually exist for the active control system. They are only used to measure the deviation between the actually reproduced and the desired pressure field inside the listening area. This also means that there is another vector, $\mathbf{q}_o^{(ideal)}$, which produces a sound pressure that fits the desired sound field much better inside the listening area. This vector is calculated by minimizing the global cost function $J^{(ideal)} = \mathbf{e}_M^H\mathbf{e}_M$, resulting to the optimum solution

$$\mathbf{q}_o^{(ideal)} = (\mathbf{Z}_M^H\mathbf{Z}_M)^{-1}\mathbf{Z}_M^H\mathbf{p}_{d,M}\,. \tag{10}$$

For this control strategy the system takes into account the information provided by all the monitor sensors inside the room. The control region thus coincides with the listening area, it in that sense it corresponds to the case of an ideal system. No regularization technique is included in Eq. (10). For the conditions examined, matrix $\mathbf{Z}_M^H\mathbf{Z}_M$ is always positive definite and invertible.

For the global system matrix $\mathbf{Z}_M$ we present an interesting modification. This matrix can be multiplied by $\mathbf{L}^{-1}$, where **L** is derived from Eq. (4), resulting to the modified global system matrix

$$\tilde{\mathbf{Z}}_M = \mathbf{Z}_M\mathbf{L}^{-1}\,. \tag{11}$$

As it will be shown, this matrix has some very interesting properties comparing to the original global system matrix $\mathbf{Z}_{\mathrm{M}}$.

2.3 The adaptive approach

The steepest descent method is employed for the adaptive calculation of the optimum source strengths. For the adaptive approach we take into account the ideal system. The source strengths are adapted at each sample by two different iterative equations

$$\mathbf{q}(n+1) = \mathbf{q}(n) + \alpha\mathbf{Z}_M^H\mathbf{e}_M(n)\,. \tag{12}$$

and

$$\tilde{\mathbf{q}}(n+1) = \tilde{\mathbf{q}}(n) + \tilde{\alpha}\tilde{\mathbf{Z}}_M^H\mathbf{e}_M(n)\,. \tag{13}$$

Here α and $\tilde{\alpha}$ are the convergence coefficients and

$$\mathbf{e}_M(n) = \mathbf{p}_{d,M} - \mathbf{Z}_M\mathbf{q}(n) = \mathbf{p}_{d,M} - \tilde{\mathbf{Z}}_M\tilde{\mathbf{q}}(n)\,. \tag{14}$$

It must be said that none of Eqs. (12) or (13) includes any kind of regularization such as effort regularization or power output regularization. The analytical calculation of $\mathbf{q}_o^{(ideal)}$ based on matrixes $\mathbf{Z}_M$ or $\tilde{\mathbf{Z}}_M$ would lead to the exact same solution. However, it is the convergence rate of each algorithm that differentiates the performance between Eqs. (12) and (13), as it will be seen. Also, it is assumed that the matrixes $\mathbf{Z}_M$ and **W** correspond to the exact transfer function matrixes, and as a consequence, this also holds for matrixes **L** and $\tilde{\mathbf{Z}}_M$. In a more realistic approach, both matrixes should include modeling errors and the analysis should be implemented in terms of the magnitude of the plant error in both matrixes $\mathbf{Z}_M$ and **W**.

3 Simulation results

In the simulations presented in what follows the conventional modal sum of the sound field in a lightly damped rectangular enclosure proposed by Morse [8] is used in the form described by Bullmore et al [9]. Each source is modeled as a square pistons that vibrate with a normal velocity $u_l = \dfrac{q_l}{A}$, where $A = a^2$ is the area of the piston sources. The piston sources are oriented inside the room so that their surfaces are parallel either to the *xz*- or to the *yz*-plane.

3.1 Conditions for the simulations

A sound field reproduction system using 26 piston sources is considered in a two-dimensional rectangular room with dimensions $Lx = 3.2$, $Ly = 3.6$, and $Lz = 0.2$ m as in Fig. 1. The 26 sources are modeled as square pistons with side length of 0.1 m. The pistons are oriented in such way that their surfaces are parallel to the closest wall at a distance of 0.05 m. The listening area in the room is defined as a rectangular plane centered inside the room with dimensions of 1.6x1.6 m with lower left corner at (0.8, 1, 0.1) m and upper right one at (2.4, 2.6. 0.1) m. The performance of the reproduction system is quantified by 17 x 17 = 289 monitor sensors that totally cover the listening region. The distance between adjacent sensors along the x- and y-axes is always equal to 0.1 m. Also, a compact array of 16 control sensors is placed in the middle of the listening area. The control sensors are represented by the black dots in Fig. 1. This array covers less than 6% of the listening area.

All the modes up to 1200 Hz are used to model the sound field inside the room. The damping factor is set to 0.02 for all the modes, corresponding to a reverberation time of 0.55 s at 100 Hz. The desired pressure field inside the room is defined as a two-dimensional plane wave traveling at direction θ as shown in Fig.1.

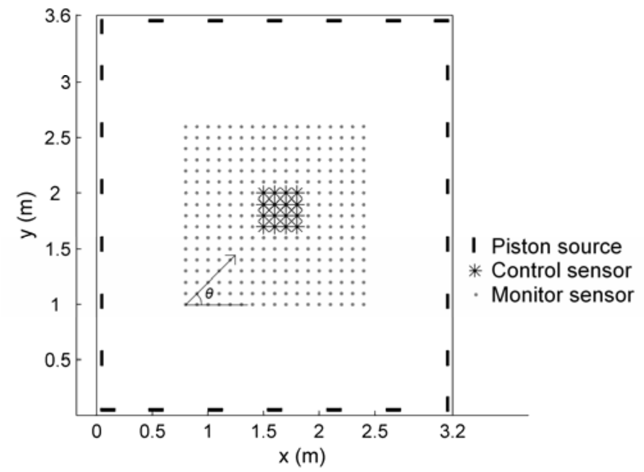

Fig.1 Geometry of the room and arrangement of the sound reproduction system.

3.2 Global reproduction performance

The results that follow intend to distinguish the two regularization methods in terms of their global reproduction performance, as defined in Eq. (8). The performance of each regularization method, as well as of the ideal system, is shown at two different frequencies at 330 and 400 Hz as a function of the angle of propagation in θ (degrees) in Fig. 2(a) and 2(b) respectively.

The figures illustrate that power output regularization is characterized by a uniform reproduction performance at all angles of propagation, while effort regularization has led to serious degradation at angles of propagation that coincide with the directions of the axial modes at 0^o, 90^o, 180^o, and 270^o. In [5] it has been shown that the reproduction error at the exact control sensor positions is trivial for both techniques at these angles of propagation but the classical regularization technique leads to unexcused activation of standing wave patterns outside the control region. Effort regularization has the tendency to activate the x- and y-axial modes simultaneously, irrespective of the direction of propagation. The axial modes of the room are the strongest mean for the reproduction of a given pressure level inside the room and effort regularization takes advantage of them in order to reproduce the desired pressures at the control region with the minimum effort from the sources. Unfortunately, the x-axial modes have a very destructive effect in the generation of a plane wave propagating along the y-axis and for the same reason, the y-axial modes are very destructive for the case of propagation along the x-axis, resulting to the strong peaks in the global reproduction performance of effort regularization seen in Fig. 2. Interestingly, the global reproduction performance is improved at the diagonal angles of propagation, which can be explained by the fact that both the x- and y- axial modes, as well as their combinations, are desired at these angles for the generation of the traveling wave.

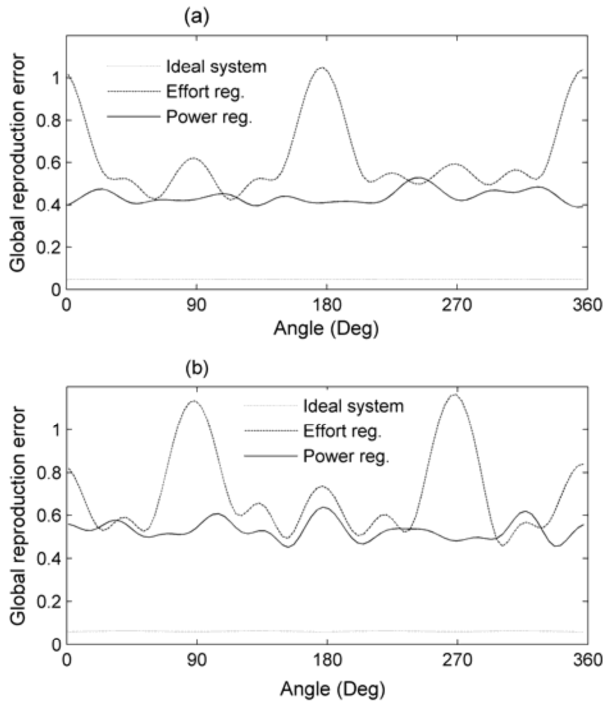

Fig. 2 Global reproduction performance as a function of the angle of propagation of the plane wave at **a.** 330 Hz and **b.** 400 Hz.

Power output regularization avoids the generation of such “unwanted” modes because in the opposite case, this would result to increment of the power output of the system. At the same time, the power output penalty term in the cost function forces some of the sources to behave like acoustical sinks (negative power), which is necessary in order to activate the mechanisms of power absorption which are required for the generation of the traveling wave [3, 5]. In that sense, it can be said that the proposed technique makes a more selective excitation of the room modes and as a result, avoids the severe reduction of the spatial robustness that is observed from the use of the traditional regularization technique at the axial angles of propagation.

3.3 System equalization

The variation in the dynamic response of the each path, which is a consequence of the room natural dynamics, defines the algebraic properties of the matrixes $\mathbf{Z}_M$ and $\mathbf{W}$, such as the maximum singular value and the condition number at each frequency.

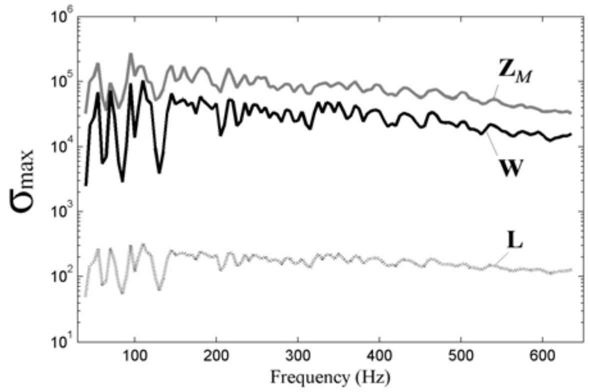

Fig. 3 Variation of the maximum singular value of matrixes $\mathbf{Z}_M$, $\mathbf{W}$ and $\mathbf{L}$ with frequency.

In Fig. 3 we plot the maximum singular values of matrixes $\mathbf{Z}_M$, $\mathbf{W}$, and $\mathbf{L}$ as a function of the frequency. One can see that the peaks and dips in the maximum singular value of matrixes $\mathbf{W}$ and $\mathbf{L}$ coincide exactly with the peaks and dips of matrix $\mathbf{Z}_M$. This coincidence is a consequence of the room natural dynamics. One can observe that both expressions $\mathbf{q}^H\mathbf{Z}_M^H\mathbf{Z}_M\mathbf{q}$ and $\mathbf{q}^H\mathbf{W}\mathbf{q}$ represent two relevant quantities: the average potential energy over the listening area [10] and the total acoustic power output. The condition number of the original and the modified system matrix is plotted as a function of the frequency in Fig. 4. It can be seen that post-filtering with $\mathbf{L}^{-1}$ has completely smoothed the frequency variation in the condition number of the modified matrix. From a point of view, the above action has equalized the non-flat system dynamics and has thus achieved system equalization [11]. In the next section, the advantages of this post-filtering process are shown in terms of increment of the convergence rate in the adaptive approach.

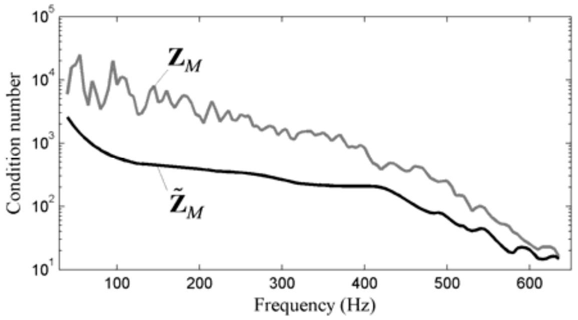

Fig. 4 Variation of the condition number of the original and the modified system matrixes.

3.4 Convergence rate of the adaptive algorithm

The convergence rate of each adaptive algorithm defined in Eqs. (12) and (13) depends on the conditioning of the matrixes $\mathbf{Z}_M$ and $\tilde{\mathbf{Z}}_M$ respectively [12]. From simulations made at various frequencies and angles of propagation it was generally observed that Eq. (13) converges faster than Eq. (12) at all frequencies and especially near the axial angles of propagation at 0°, 90°, 180° and 270°. It was observed that the value of α had to vary dynamically in order to ensure the stability of the algorithm at different frequencies whereas a constant value of $\tilde{a}$ was satisfactory for a wider range of frequencies. This seems to be in agreement with the fact that the values of the convergence coefficients in order to ensure stability are bounded from above as [12]

$$0 < a < \frac{2}{\sigma_{\max}},\ 0 < \tilde{a} < \frac{2}{\tilde{\sigma}_{\max}}, \tag{15}$$

where $\sigma_{\max}$ and $\tilde{\sigma}_{\max}$ is the largest singular value of the original and the modified system matrix respectively.

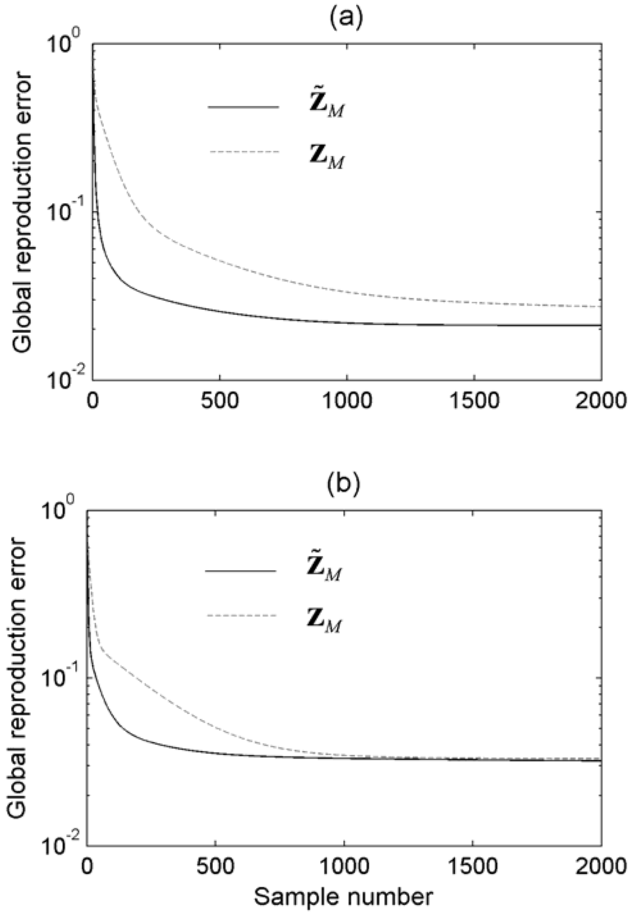

Fig. 5 Global reproduction error as a function of the iteration number of the adaptive algorithm for the generation of a plane wave traveling **a.** at 90° at 160 Hz and **b.** at 0° at 220 Hz.

Two examples can be seen, when generating a plane wave traveling at the angles of 90° and 0°, at the frequencies of 160 Hz and 220 Hz, in Fig. 5(a) and (b) respectively. The convergence coefficient α at each frequency was set to the maximum value that could be used in order to maximize the convergence rate of the conventional algorithm without forcing it to become unstable, whereas $\tilde{a}$ was set to constant value equal to $2 \cdot 10^{-8}$ for both frequencies.

4 Conclusion

The spatial robustness and the convergence rate of the conventional control approaches used in the active reproduction of a sound field in a reverberant rectangular room show severe deterioration at specific frequencies and angles of propagation. These problems are caused by the room resonance modes and by the resulting eigenvalue spread that is observed in the original system matrix that carries the acoustic impedances from the loudspeakers to the receivers inside the room. A promising way to compensate for these problems in the frequency domain has been shown, assuming that the matrix **W** that carries the real part of the acoustic impedances between sources is known. The use of matrix **W** here has been successfully incorporated in the solution in terms of regularization and system equalization. As a result, improvement of the spatial robustness and increment of the convergence rate was achieved.

Although the calculation of **W** is a straightforward task for the simulation model, this information needs further theoretical and experimental investigation in order to be derived in a real problem. Obviously, additional sensing equipment, as for example sound pressure and sound velocity microphones near the diaphragms of the loudspeakers will be required [13, 14]. Also, the investigation must be adapted to the case of broadband signals as well. This should be the next step that will judge the applicability of the proposed technique for the case of a real sound reproduction system.

References

[1] F. Asano and D. C. Swanson, "Sound equalization in enclosures using modal reconstruction," *J. Acous. Soc. Am.* 98, 2062-2069 (1995)

[2] A. O. Santillan, "Spatially extended sound equalization in rectangular rooms," *J. Acoust. Soc. Am.* 110, 1989-1997 (2001)

[3] P. A. Gauthier, A. Berry, and W. Woszczyk, "Sound field reproduction in-room using optimal control techniques," *J. Acoust. Soc. Am.* 117, 662-678 (2004)

[4] P. A. Nelson, "A review of some inverse problems in acoustics," *International J. Acoust. Vib.* 6, 118-134 (2001).

[5] N. Stefanakis, J. Sarris, G. Cambourakis and F. Jacobsen, "Power-output regularization in global sound equalization," *J. Acoust. Soc. Am.* 123, 33-36 (2008)

[6] S. J. Elliot, P. Joseph, P. A. Nelson, and M. E. Johnson, " Power output minimization and power absorption in the active control of sound," *J. Acoust. Soc. Am.* 90, 2501-2512 (1991)

[7] C. Hansen, *Rank deficient and discrete ill-posed problems*, SIAM, Philadelphia (1998).

[8] P. M. Morse, Vibration and Sound, New York, McGraw Hill, second edition, 1948.

[9] A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, "The active minimization of harmonic enclosed sound fields, part II: A computer simulation," *J. Sound Vib.* 117, 15-33 (1987).

[10] P. A. Nelson, A. R. Courtis, S. J. Elliott and A. J. Bullmore, "The active minimization of harmonic enclosed sound fields, Part I: Theory," *J. Sound Vib.* 117, 1-13 (1987).

[11] M. R. Bai and S. J. Elliott, "Preconditioning multichannel adaptive filtering algorithms using EVD- and SVD-based signal prewhitening and system decoupling," *J. Sound. Vib.* 270, 639-655 (2004).

[12] S. J. Elliott, C. C. Boucher and P. S. Nelson, "The behavior of a multiple channel active control system," *IEEE Trans. Signal Process.* 40, 1041-1052 (1992)

[13] G. J. Adams, "Adaptive control of loudspeaker frequency response at low frequencies," presented at the 73rd AES convention, Eindhoven, 1983.

[14] M. E. Johnson and S. J. Elliott, "Measurement of the acoustic power output in the active control of sound," *J. Acous. Soc. Am.* 93, 1453-1459 (1993).