KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

| | |
|---|---|
| **Citation/Reference** | Antonello N., De Sena E., Moonen M., Naylor P. A., van Waterschoot T. (2016)<br><br>**Sound field control in a reverberant room using the Finite Difference Time Domain method**<br>in AES 60th International Conference, Leuven, Belgium, Feb. 2016, accepted for publication. |
| **Archived version** | Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher |
| **Published version** | |
| **Journal homepage** | http://www.aes.org/ |
| **Author contact** | niccolo.antonello@esat.kuleuven.be<br>*+ 32 (0)16 321855* |
| **IR** | |

*(article begins on next page)*

# Sound field control in a reverberant room using the Finite Difference Time Domain method

Niccolò Antonello[1], Enzo De Sena[1], Marc Moonen[1], Patrick A. Naylor[2], and Toon van Waterschoot[1,3]

[1]*KU Leuven, Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, 3001 Leuven, Belgium*

[2]*Imperial College London, Dept. of Electrical Engineering, SW7 2AZ, London, United Kingdom*

[3]*KU Leuven, Dept. of Electrical Engineering (ESAT), ETC, AdvISe Lab, Kleinhoefstraat, 2440 Geel, Belgium*

Correspondence should be addressed to Niccolò Antonello (`niccolo.antonello@esat.kuleuven.be`)

**ABSTRACT**

In this paper a novel approach for sound field reproduction and multi-zone sound field control in a reverberant room is proposed. Sound field reproduction and multi-zone sound field control are jointly formulated as an optimization problem. The optimization problem relies on a physical model of the room acoustics which is computed numerically using the finite difference time domain (FDTD) method. Furthermore a novel regularization of the optimization problem is proposed. A non-convex sparsity-inducing term is used to obtain convenient control source positions and control source signals simultaneously. Finally, simulation results are presented: the proposed method is used to recreate an anechoic sound in a highly reverberant room while keeping part of the room silent.

## 1. INTRODUCTION

Sound field reproduction techniques aim to reproduce a measured or synthesized acoustic wave field in a different acoustic environment. Many attempts to improve this technology have been pursued among which the most popular are ambisonics [1] and wave field synthesis [2] which try to physically reconstruct the sound field. Most of these techniques require large amounts of loudspeakers and assume low reverberation of the room where they are applied or even free field conditions. Other techniques focus on perceptual sound field reproduction [3].

Another field of research in acoustics is *multi-zone sound field control*: here the aim is to be able to play sound in a specific area of the room (bright zone) while keeping other locations silent (dark zone) [4]. Multi-zone sound field control shares many features with sound field reproduction. In particular, multi-zone sound field control pressure matching and sound field reproduction can both be formulated as optimization problems, more specifically as a regularized least squares (LSs) problems [4–6], also known as multi-channel equalization problems [7, 8]. It has been recently shown that the least absolute shrinkage and selection operator (Lasso), which seeks sparse solutions, can perform better than LS in both sound field reproduction [9] and multi-zone sound field control [10]. Here convenient placement of the control source positions can be first obtained solving the Lasso and a second optimization is then performed to obtain the control source signals once the optimal position is known.

In many of these approaches, the room impulse responses (RIRs) are usually unknown and need to be estimated. Having a physical room model can compensate for this lack of information and a numerical method can be used to model the room acoustics and to generate the RIRs. In particular, the finite difference time domain (FDTD) method is increasingly becoming popular among the numerical methods for room acoustics simulations. This is due to the recently improved boundary conditions (BCs) formulation [11].

In this paper a novel approach for sound field reproduction and multi-zone sound field control is proposed. The novel aspects of this paper can be summarized as follows: (i) a non-convex sparsity-inducing term will be

used in order to promote sparsity more efficiently than the Lasso. This will make it possible to optimize for the control source positions and control source signals simultaneously; (ii) the optimization problem is modified including an equality constraint that represents the physical model of the room where the algorithm is applied. This equality constraint is obtained through the FDTD method. This enables to compute the RIRs necessary for the sound field control and reproduction.

The paper is structured as follows: in Sec. 2 a brief review of the FDTD method is presented. In Sec. 3 the optimization algorithm is presented where a novel non-convex sparsity-inducing term is used and the FDTD method is used to compute the RIRs. Finally, in Sec. 4 simulation results will be presented. The proposed algorithm will be tested in a highly reverberant two-dimensional room creating an anechoic impulse while keeping the dark zone silent.

## 2. THE FINITE DIFFERENCE TIME DOMAIN

The FDTD method is a numerical method which seeks a solution of a partial differential equation (PDE) through spatiotemporal discretization. It is known that solutions of PDEs with specific BCs and initial conditions (ICs) can be derived analytically only for simple geometries [12]. In the context of room acoustics, the wave equation models the acoustic wave phenomena [12]

$$\begin{aligned} &\text{PDE} \quad \triangle p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = s \text{ on } \Omega \times \tau \\ &\text{BCs} \quad \frac{\partial p}{\partial t} = -c\xi \nabla p \cdot \mathbf{n} \text{ on } \partial\Omega \times \tau \\ &\text{ICs} \quad \frac{\partial p}{\partial t} = \hat{p}_0, p = p_0 \text{ on } \Omega, \end{aligned} \tag{1}$$

where $\triangle$ is the Laplacian, $p$ and $s$ are the sound pressure and source distribution, respectively, defined as scalar functions with spatial domain $\Omega \in \mathbb{R}^2$ (for a two-dimensional room) and temporal domain $\tau \in \mathbb{R}$. Here $c$ is the speed of sound, $\xi$ is the specific acoustic impedance which models the acoustic losses occurring at the walls, $\mathbf{n}$ is the normal vector with respect to the boundary surface, $p_0$ is the initial sound pressure and $\hat{p}_0$ the initial sound pressure temporal derivative.

The idea of the FDTD method is to approximate the continuous scalar functions $p$ and $s$ on a uniform grid, i.e. the sound pressure is sampled such that

$$p(x,y,t) \approx p(lX, mX, nT) = p^n_{l,m}, \tag{2}$$

where $X$ is the spatial resolution and $T$ is the temporal resolution. Furthermore the second order derivatives appearing in (1) are approximated using centered finite differences, e.g. for the spatial second-order derivative over $x$ that appears in the Laplacian

$$\frac{\partial^2 p}{\partial x^2} \approx \frac{p^n_{l+1,m} - 2p^n_{l,m} + p^n_{l-1,m}}{X^2}. \tag{3}$$

Using these approximations the wave equation can be converted into a set of linear equations. For explicit FDTD schemes [11] the resulting update equation is given by

$$p^{n+1}_{l,m} = \Sigma P^n_{l,m} - p^{n-1}_{l,m} + s^n_{l,m,i}, \tag{4}$$

where the sound pressure sample at position $(l,m)$ and time $n+1$ is computed using the previous sample $p^{n-1}_{l,m}$ and a weighted sum, $\Sigma P^n_{l,m}$, of the samples of the 8 neighbor samples of $p^n_{l,m}$ and itself. The weights used in $\Sigma P^n_{l,m}$ determines the explicit scheme used to approximate the Laplacian. Spatial and temporal resolutions are bounded by a stability constraint

$$\frac{cT}{X} \leq \lambda_c \tag{5}$$

where $\lambda_c$ is the Courant number [11]. The ratio between $T$ and $X$ is usually chosen at the stability limit where numerical errors are minimized [11]. In order to take into account the BCs, the update equations must be modified at the boundaries by enforcing the approximation of the BCs in (1), which is performed again using finite differences [11].

Typically the set of linear equations given by the FDTD method is solved iteratively. However for easier readability, it is possible to write this system in vectorized notation:

$$\mathbf{Bp} = \mathbf{s} \tag{6}$$

where $\mathbf{B}$ is a sparse $N_x N_y (N_t+2) \times N_x N_y (N_t+2)$ matrix [13] and $\mathbf{p}$ and $\mathbf{s}$ are $N_x N_y (N_t+2)$ vectors containing the $p^n_{l,m}$ and $s^n_{l,m}$ for all $l = 0 \ldots N_x$, $m = 0 \ldots N_y$ and $n = -2, -1, 0 \ldots N_t$, where $N_x$ and $N_y$ are the number of divisions of the uniform grid in the $x$ and $y$ directions respectively and $N_t$ is the number of time samples. The $N_t + 2$ is due to the two ICs which here will be assumed to be zero. Notice that the bandwidth of the RIRs generated with the FDTD method is usually limited to a low frequency range due to the heavy computational load and the numerical errors of the FDTD method, which both increase with frequency.

## 3. THE OPTIMIZATION ALGORITHM

Sound field reproduction can be posed as a multi-channel equalization problem. This is an optimization problem that can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{s}_c} \quad & f = \frac{1}{2}\|\mathbf{F}_p\mathbf{p} - \tilde{\mathbf{p}}\|_2^2 \\ \text{s. t.} \quad & \mathbf{B}\mathbf{p} = \mathbf{F}_s^T\mathbf{s}_c, \end{aligned} \tag{7}$$

where $f$ is the cost function which represents the squared $l_2$-norm of the difference between the vector containing the wanted sound pressure signals $\tilde{\mathbf{p}}$ and the room sound pressure signals $\mathbf{p}$ at some specific positions in the room which will be referred here as *control points*. Both vectors are $N_t \cdot K_p$ long, where $K_p$ is the number of control points. The matrix $\mathbf{F}_p$ is a selection matrix which selects the samples of $\mathbf{p}$ that correspond to the control points. The equality constraint consists of the set of linear equations given by the FDTD method. Here $\mathbf{s}_c$ is a $N_s \cdot K_s$ vector containing the signals of the sought *control sources* at $K_s$ positions having $N_s$ samples each. The expansion matrix $\mathbf{F}_s^T$ expands $\mathbf{s}_c$ onto a vector having the same dimension as $\mathbf{p}$.

The optimization problem (7) is convex and can be solved analytically using LS. If the equality constraint is substituted into the cost function, one can write

$$\min_{\mathbf{s}_c} \frac{1}{2}\|\underbrace{\mathbf{F}_p\mathbf{B}^{-1}\mathbf{F}_s^T}_{\mathbf{G}}\mathbf{s}_c - \tilde{\mathbf{p}}\|_2^2 \tag{8}$$

where $\mathbf{G}$ is a rectangular matrix of dimensions $K_pN_t \times K_sN_s$. The solution of (8) is given by:

$$\mathbf{s}_c^* = -(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\tilde{\mathbf{p}}. \tag{9}$$

The LS problem is usually ill-posed: the matrix $\mathbf{G}^T\mathbf{G}$, which represents the Hessian of the cost function, can be non-invertible. In order to avoid this, a regularization term can be added to the cost function (see Sec. 3.3).

### 3.1. Computation of G

As described in the previous section, $\mathbf{G}$ is given by:

$$\mathbf{G} = \mathbf{F}_p\mathbf{B}^{-1}\mathbf{F}_s^T. \tag{10}$$

$\mathbf{G}$ actually consists of a collection of Toeplitz matrices of the RIRs between each of the control points and control source positions and its dimensions are $K_p \cdot N_t \times K_s \cdot N_s$. The matrices $\mathbf{F}_s^T$ and $\mathbf{F}_p$ contain ones at the indexes corresponding to the spatial and temporal positions of the control source positions and control points respectively.

$\mathbf{G}$ can be computed by solving a number of FDTD simulations. In the simple case where $K_p = K_s = 1$, $\mathbf{G}$ will consist of a single Toeplitz matrix and its first column $\mathbf{g}_0$ can be computed by solving

$$\mathbf{B}\mathbf{h}_s = \mathbf{f}_{s,0}, \tag{11}$$

where $\mathbf{f}_{s,0}$ is the first column of $\mathbf{F}_s^T$ which has only one non-zero element at the index of the control source position and time step $n = 0$. $\mathbf{g}_0$ can then be obtained using the selection matrix

$$\mathbf{g}_0 = \mathbf{F}_p\mathbf{h}_s. \tag{12}$$

The second column of $\mathbf{G}$ will actually consist of the vector $\mathbf{g}_0$ being time shifted by one sample. That is because the second column of $\mathbf{F}_s^T$ has only one non-zero element at the index of the control source position and time step $n = 1$.

Hence if $K_p > 1$ and $K_s = 1$, the FDTD simulation represented by (11) needs to be solved only once and multiple RIRs between control points and control source position can be generated simultaneously using (11). In general, if $K_p > K_s$, $K_s$ FDTD simulations are needed to generate $\mathbf{G}$. On the other hand, if $K_p < K_s$ it is more efficient to swap control points and control source positions. The generated RIRs are in fact equivalent due to the reciprocity principle of the Green's function [12]. In this case only $K_p$ FDTD simulations will be necessary.

### 3.2. Newton-type optimization

Since a non-convex regularization term will be used to regularize the ill-posed optimization problem (8), a closed-form solution cannot be achieved. Instead, an iterative method can be used to obtain a solution. The idea is to start from an initial guess of the control source signals, $\mathbf{s}_c^0$, compute the Newton step [16]

$$\mathbf{d}^i = -(\nabla^2 f(\mathbf{s}_c^i))^{-1}\nabla f(\mathbf{s}_c^i), \tag{13}$$

obtain new control source signals $\mathbf{s}_c^1$, and repeat iteratively this procedure until a local minimum is reached. A line search is also performed at each iteration to ensure sufficient decrease on the cost function [16].

An efficient computation of the gradient of the cost function is fundamental to ensure fast convergence. The gradient of (8) can be written as

$$\nabla f = \mathbf{G}^T(\mathbf{p}^i - \tilde{\mathbf{p}}), \tag{14}$$

where $\mathbf{p}^i$ is the room sound pressure at the control points for a given $\mathbf{s}_c^i$. As stated in the previous section, $\mathbf{G}$ consists of a collection of Toeplitz matrices. These actually resemble linear convolution operations, and $\nabla f$ can then be written alternatively as

$$\nabla f = \sum_{k=1}^{K_s}\sum_{m=1}^{K_p} \mathbf{F}_{N_s}\left(\hat{\mathbf{h}}_{k,m} * (\mathbf{p}_m^i - \tilde{\mathbf{p}}_m)\right), \tag{15}$$

where $*$ stands for linear convolution, $\hat{\mathbf{h}}_{k,m}$ is the time reversed RIRs between the control point position $m$ and control source point position $k$, $\mathbf{p}_m^i$ is the room sound pressure at the $i$th Newton step and $\tilde{\mathbf{p}}_m$ is the wanted sound pressure, both at the control point $m$. $\mathbf{F}_{N_s}$ is a selection matrix that takes only the first $N_s$ samples of the linear convolution. When the size of $\mathbf{G}$ is large, expression (15) can be more efficient than computing (14) since FFT can be used to compute the linear convolution. Finally the Hessian $\nabla^2 f$ can be computed using

$$\nabla^2 f = \mathbf{G}^T\mathbf{G}. \tag{16}$$

While the gradient has to be evaluated at each Newton step due to its dependence of $\mathbf{p}^i$ in expression (15), the Hessian has to be computed only once. The matrix inversion performed in (13) represents the computational bottleneck of this method.

### 3.3. Regularization

In most application cases, the inverse problem in (7) is ill-posed. In particular, if $N_s \cdot K_s > N_t \cdot K_p$, i.e. if the problem is underdetermined, there are many combinations of control source signals that can recreate $\tilde{\mathbf{p}}$ at the control points. One approach is to look for a unique solution by adding a regularization term $\phi(\mathbf{s}_c)$ to the cost function $f$. The new regularized cost function will be called $f_r$. A simple choice is to use Tikhonov regularization

$$\phi_T(\mathbf{s}_c) = \frac{\lambda_T}{2}\|\mathbf{s}_c\|_2^2, \tag{17}$$

where $\lambda_T$ is a scalar which weights the importance of this term inside the cost function. Here the regularization term will force the problem to look for a solution which penalizes large values in the control source signals. Unfortunately, Tikhonov regularization will also evenly distribute the energy through all the control source signals. Therefore if the number of active control sources is to be minimized, a different regularization should be used. One could in fact try to minimize the number of active control source positions by actually adding this number to the cost function. Such a regularization would be achieved using the following regularization term:

$$\phi_0(\mathbf{s}_c) = \lambda_0 \sum_{k=1}^{K_s} \|\|\mathbf{F}_k\mathbf{s}_c\|_2\|_0, \tag{18}$$

where $\mathbf{F}_k$ selects the samples of the $k$th control source signal whose energy is evaluated by the $l_2$-norm. The $l_0$-norm will give 1 if the control source is active and a 0 if it is not.

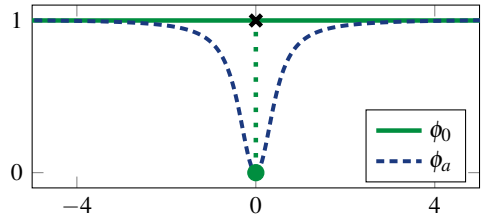


**Fig. 1:** $l_0$-norm of a scalar variable and its relaxation using (1) with $\gamma = 10$ and $\lambda_a = 0.12$.

This term is a *group sparsity-inducing regularization* since it forces the solution to be spatially sparse. However, solving an optimization problem that involves the $l_0$-norm is very difficult: its discontinuity at 0, which can be seen in Fig. 1, makes Newton-type optimization not feasible and yields an NP-hard combinatorial problem. In general, a relaxation of the $l_0$-norm is used to avoid this problem. A typical choice is the $l_1$-norm, which leads to the Lasso. Here, however, the following relaxation will be used:

$$\phi_a(\mathbf{s}_c) = \sum_{k=1}^{K_s} \frac{2\lambda_a}{\gamma\sqrt{3}}\left(\operatorname{atan}\left(\frac{1+\gamma\|\mathbf{F}_k\mathbf{s}_c\|_2^2}{\sqrt{3}}\right) - \frac{\pi}{6}\right), \tag{19}$$

where $\lambda_a$ is the weighting scalar and $\gamma$ is a function parameter. As it can be seen in Fig. 1, this relaxation approximates well the behavior of the $l_0$-norm and has the advantage that it is continuous and differentiable at 0. Notice that for $\gamma \to 0$ the $\phi_a$ will tend to Tikhonov regularization. Instead, for high values of $\gamma$ and a specific $\lambda_a$ this function will approach the $l_0$-norm.

It was recently shown that a similar relaxation can enhance sparsity better than the $l_1$-norm relaxation [14,15]. The usage of a non-convex regularization such as (19) can keep the overall problem still convex for a certain set of values of $\gamma$ [14, 15]. Nevertheless, in this work non-convex optimization will be used, leaving its convex formulation to future work.

Tikhonov regularization will still be used to limit control source signals being too large. The gradient and Hessian of the regularized cost function $f_r$ will become

$$\nabla f_r(\mathbf{s}_c^i) = \nabla f + \lambda_T \mathbf{s}_c^i + \nabla \phi_a(\mathbf{s}_c^i), \tag{20}$$

and

$$\nabla^2 f_r(\mathbf{s}_c^i) = \nabla^2 f + \lambda_T \mathbf{I} + \nabla^2 \phi_a(\mathbf{s}_c^i), \tag{21}$$

where $\mathbf{I}$ is the identity matrix, and $\nabla\phi_a(\mathbf{s}_c^i)$ and $\nabla^2\phi_a(\mathbf{s}_c^i)$ are the gradient and Hessian of (19) respectively. These can be computed analytically and are shown in equations (22) and (23) respectively.

When the optimization problem is solved, the group sparsity term will indicate which control sources are active out of the candidate control source positions. The active control sources are then used to generate the wanted sound field. Notice that compared with the Lasso approach, the proposed algorithm gives control source positions and control source signals simultaneously. In the case of Lasso, active control sources are usually biased towards smaller values [14]. This bias can be removed by applying once more the optimization algorithm using LS or Lasso with the control source positions obtained from the first solution [9, 10]. Thanks to the atan regularization, sparsity is induced more effectively and non-active control sources signals are unbiased making it possible to optimize simultaneously for control source positions and control source signals.

### 3.4. Dark Zone Regularization

A dark zone can be enforced in a similar fashion as the regularization performed in the previous section. A dark zone inducing term can be added to the cost function

$$\phi_{dz}(\mathbf{p}) = \frac{\lambda_{dz}}{2} \|\mathbf{F}_{dz}\mathbf{p}\|_2^2, \tag{24}$$

which forces the sound pressure being small at $K_{dz}$ specific positions, referred here as *dark zone points*. Here $\lambda_{dz}$ weights the cost of this term with respect to the other ones appearing in the cost function. The room sound pressure samples at the dark zone points are selected by $\mathbf{F}_{dz}$. If the equality constraint of (7) is substituted into (24) the gradient and Hessian of this term can be computed as in Sec. 3.1 and 3.2. The gradient of (24) will be

$$\nabla\phi_{dz}(\mathbf{p}^i) = \lambda_{dz}\mathbf{G_{dz}}^T\mathbf{p}^i, \tag{25}$$

where $\mathbf{G_{dz}}^T$ is now the collection of Toeplitz matrices of the RIRs between the control source positions and dark zone points. The gradient can be computed more efficiently using linear convolution

$$\nabla\phi_{dz}(\mathbf{p}^i) = \lambda_{dz}\sum_{k=1}^{K_s}\sum_{m=1}^{K_{dz}} \mathbf{F}_{N_s}\left(\hat{\mathbf{h}}_{k,m} * \mathbf{p}_m^i\right). \tag{26}$$

The Hessian will be given by

$$\nabla^2\phi_{dz} = \lambda_{dz}\mathbf{G_{dz}}^T\mathbf{G_{dz}}. \tag{27}$$

By adding (26) and (27) to (20) and (21) respectively, and using these results into (13) the iterative procedure can be used to find the spatially sparse control sources which recreate $\tilde{\mathbf{p}}$ at the control points while keeping silent the dark zone points.

## 4. SIMULATION RESULTS

The performance of the algorithm described in Sec. 3 is evaluated in a two-dimensional highly reverberant room. A low spatial resolution of $X = 28$ cm is used, resulting in a sampling frequency of $F_s = 1426$ Hz. A sketch of the room is shown in Fig. 2: here the $K_p = 4$ control points, the $K_{dz} = 9$ dark zone points and $K_s = 22$ candidate control source positions are shown. The room's reverberation time is $T_{60} = 2.8$ seconds.

Typically, a sound field measured in a different room is imposed at the control points. Here instead, an impulse will be enforced simultaneously at all the control positions. This rather unnatural sound field is chosen in order to show that, despite the high reverberation of the room, an anechoic sound can be reproduced. Moreover it is possible to test the algorithm performance for different frequency ranges. Fig. 3 shows two different *mexican hat* wavelets, which are the impulses that will be used in $\tilde{\mathbf{p}}$. The parameter $\sigma$ controls the bandwidth of the wavelet as Fig. 3 shows. A broad frequency range is achieved using $\sigma_h = 15 \cdot 10^{-4}$ while a low frequency range is obtained using $\sigma_l = 10^{-2}$.

Notice that for all the results presented here the parameter $\gamma$ is set to $5 \cdot 10^3$ and $\lambda_T = 10^{-5}$. The parameter

$$\nabla\phi_a(\mathbf{s}_c) = \sum_{k=1}^{K_s} \frac{4\lambda_a}{3} \frac{1}{\left(\frac{1+\gamma\|\mathbf{F}_k\mathbf{s}_c\|_2^2}{\sqrt{3}}\right)^2 + 1} \mathbf{F}_k^T\mathbf{F}_k\mathbf{s}_c \tag{22}$$

$$\nabla^2\phi_a(\mathbf{s}_c) \approx \sum_{k=1}^{K_s} \frac{4\lambda_a}{3} \left[ \frac{1}{\left(\frac{1+\gamma\|\mathbf{F}_k\mathbf{s}_c\|_2^2}{\sqrt{3}}\right)^2 + 1} \mathbf{F}_k^T\mathbf{F}_k - \frac{4\gamma}{\sqrt{3}} \frac{\left(\frac{1+\gamma\|\mathbf{F}_k\mathbf{s}_c\|_2^2}{\sqrt{3}}\right)}{\left[\left(\frac{1+\gamma\|\mathbf{F}_k\mathbf{s}_c\|_2^2}{\sqrt{3}}\right)^2 + 1\right]^2} \mathbf{F}_k^T\mathbf{F}_k\mathbf{s}_c^2 \right] \tag{23}$$

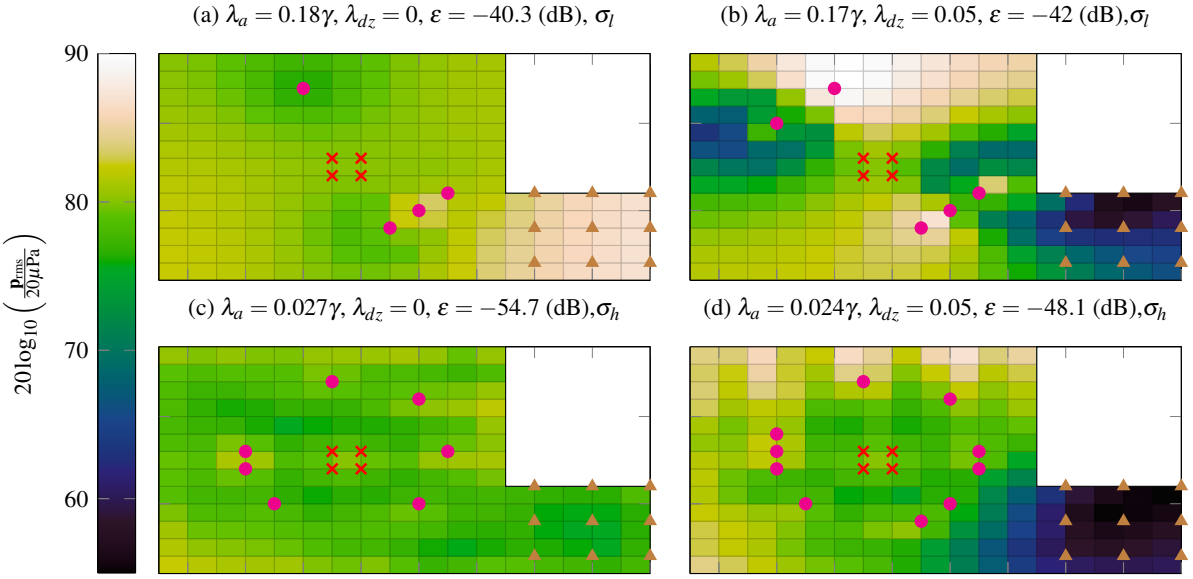


**Fig. 4:** Surface plot of the SPL integrated over time at each position of the room. (a) low frequency wavelet, (b) low frequency wavelet with dark zone, (c) high frequency wavelet and (d) high frequency wavelet with dark zone. Normalization was applied in order to have 80 dB SPL at the control points. Video animations can be seen in [17].

$\gamma$ is chosen empirically to be large enough for $\phi_a$ to be a representative relaxation of the $l_0$-norm while being small enough to ensure a weak non-linearity of $\phi_a$. The simulations are performed for 0.5 seconds ($N_t = 700$) while restricting control sources to be 0.18 seconds long ($N_s = 250$). The Newton-type optimization is initialized with $\mathbf{s}_c^0 = \mathbf{0}$. The optimization is stopped when $\|\nabla f_r\|_2 < 10^{-6}$.

Fig. 4 shows the SPL integrated over time in the room when the different wavelets are imposed at the control points with and without the dark zone regularization. Fig. 4 (a) and (b) shows the results when the low frequency wavelet is imposed. In Fig. 4 (a) only 4 control sources are active. At these frequencies the modal behavior of the sound field is prominent. The SPL is not spatially uniform and this is particularly visible at the bottom right corner where the highest SPL is present. When the dark zone regularization is applied in this area, as Fig. 4 (b) shows, it can be noticed that the SPL is reduced on average by 23 dB in the dark zone points with respect to the previous case. This is achieved by adding only one control source. The error at the control points can be seen in the value of $\varepsilon$ displayed on top of the figures and defined as:

$$\varepsilon = 10\log_{10}\left(\frac{\|\mathbf{F}_p\mathbf{p} - \tilde{\mathbf{p}}\|_2^2}{\|\tilde{\mathbf{p}}\|_2^2}\right). \tag{28}$$

Notice that in both cases by relaxing the weight of the

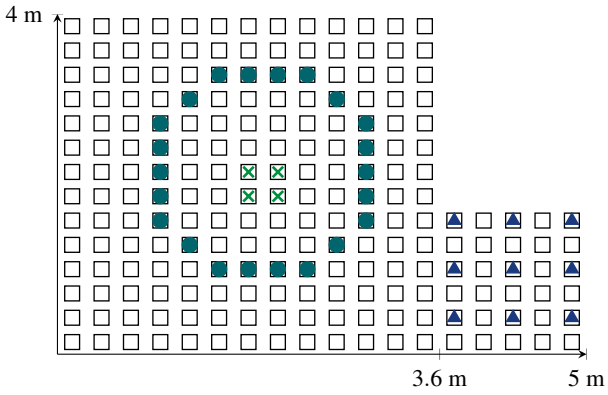


**Fig. 2:** Sketch of the L-shaped room used in the simulations. Control points are shown with crosses, control source positions in circles and dark zone points with triangles.

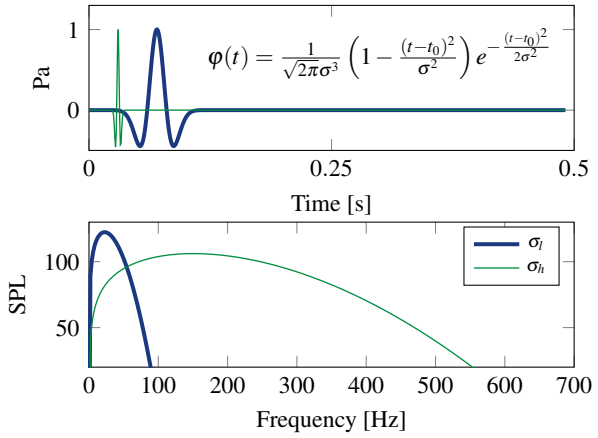


**Fig. 3:** Wavelets imposed at the control points.

sparsity regularization $\lambda_a$, the error would be further reduced but more control sources would be activated.

Fig. 4 (c) and (d) show the results for the case of the high frequency wavelet. With respect to the low frequency case, in order to obtain convergence, the weight of $\lambda_a$ had to be relaxed substantially, leading to a less sparse solution with more active control sources: 7 without dark zone regularization and 10 with dark zone regularization. This leads to an increased accuracy at the control points with the error $\varepsilon$ being smaller compared to the low frequency case. Here, the SPL is reduced on average by 19.5 dB in the dark zone points.

As stated before, the room where the simulations were performed is highly reverberant. These simulations show that an anechoic sound can be reproduced in such an acoustic environment. This is possible because the optimization algorithm actually seeks not only for control sources but also for *control sinks* capable of canceling out the direct sound and reflections created by the other control sources. This is particularly visible in the videos which can be viewed at [17].

## 5. CONCLUSIONS AND FUTURE WORK

A novel method for sound field reproduction and multi-zone sound field control which relies on a physical model of the room acoustics has been proposed. This method has the advantage of being capable of computing all of the RIRs involved in the optimization algorithm by performing multiple FDTD simulations. Using a non-convex sparsity regularization term in the optimization problem, convenient placement of the control sources can be achieved while simultaneously obtaining the control source signals, avoiding a two-step optimization that would be required by the Lasso approach. The simulation results suggest that the number of active control sources increases with the bandwidth of the control point signals and if a dark zone is applied. It was also shown that this algorithm can be used when high reverberation is present in the reproduction room and that anechoic sounds can be reproduced.

Further research will focus on the extension of the proposed method for its application in a real scenario. The computational cost of the FDTD method and the optimization algorithm should be reduced and a realistic model of the room using the FDTD method should be achieved. Moreover the control sources used here are omnidirectional point sources which are a valid assumption for loudspeakers only at low frequencies. The proposed method will be extended to include directional sound sources.

## 6. ACKNOWLEDGEMENTS

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of the FP7-PEOPLE Marie Curie Initial Training Network "Dereverberation and Reverberation of Audio, Music, and Speech (DREAMS)", funded by the European Commission under Grant Agreement no. 316969, KU Leuven Research Council CoE PFV/10/002 (OPTEC), the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office IUAP P7/19 "Dynamical systems control and optimization" (DYSCO) 2012-2017, KU Leuven Impulsfonds IMP/14/037 and was supported by a Postdoctoral Fellowship (F+/14/045) of the KU Leuven Research Fund. The scientific responsibility is assumed by its authors.

## 7. REFERENCES

[1] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *J. Audio Eng. Soc.*, vol.33, no. 11, pp. 859-871, 1985.

[2] A. J. Berkhout, D. de Vries, P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no.5 pp. 2764-2778, 1993.

[3] E. De Sena, H. Hacihabiboglu, Z. Cvetkovic, "Analysis and design of multichannel systems for perceptual sound field reconstruction," *IEEE Trans. on Audio, Speech and Language Process.* vol. 21 no. 8, pp. 1653-1665, 2013.

[4] T. Betlehem, W. Zhang, M. A. Poletti, T. D. Abhayapala, "Personal Sound Zones: Delivering interface-free audio to multiple listeners," in *IEEE Signal Proc. Magazine*, vol. 32 no. 2, pp. 81-91, 2015.

[5] M. Olik, J. Francombe, P. Coleman, P. J.B. Jackson, M. Olsen, M. Møller, R. Mason, S. Bech, "A comparative performance study of sound zoning methods in a reflective environment," in *Proc. Audio Eng. Society 52nd Int. Conf. Sound Field Control-Engineering and Perception*, 2013.

[6] P. A. Gauthier, A. Berry, W. Woszczyk, "Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain," *J. Audio Eng. Soc.*, vol. 117, no. 2, pp. 662-678, 2005.

[7] M. Miyoshi, Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. on Acoustics, Speech and Signal Process.*, vol. 36, no. 2, pp. 145152, 1988.

[8] M. Kolundzija, C. Faller, M. Vetterli, "Reproducing sound fields using MIMO acoustic channel inversion," *J. Audio Eng. Soc.*, vol.59, no.10, pp. 721-734, 2011.

[9] G. N. Lilis, D. Angelosante, and G. B. Giannakis, "Sound field reproduction using the lasso," *IEEE Trans. on Audio, Speech and Language Process.*, vol.18 no.8, pp. 1902-1912, 2010.

[10] N. Radmanesh, I. S. Burnett, "Generation of isolated wideband sound fields using a combined two-stage lasso-ls algorithm," *IEEE Trans. on Audio, Speech and Language Process.*, vol. 21, no. 2, pp. 378-387, 2013.

[11] K. Kowalczyk, M. van Walstijn, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Trans. on Audio, Speech and Language Process.*, vol. 19, no. 1, pp. 34-46, 2011.

[12] F. Jacobsen, P. M. Juhl, *Fundamentals of General Linear Acoustics*, Wiley, 2013

[13] N. Antonello, T. van Waterschoot, M. Moonen, P. A. Naylor, "Evaluation of a Numerical Method for Identifying Surface Acoustic Impedances in a Reverberant Room," in *Proc. 2015 European Congress and Exposition on Noise Control Engineeringt (EURONOISE 2015)*, Maastricht, The Netherlands, 2015.

[14] I. W. Selesnick, I. Bayram, "Sparse signal estimation by maximally sparse convex optimization," *IEEE Trans. on Signal Process.*, vol. 62, no. 5, pp. 1078-1092, 2014.

[15] P. Chen, I. W. Selesnick, "Group-sparse signal denoising: non-convex regularization, convex optimization," *IEEE Trans. on Signal Process.*, vol. 62, no. 13, pp. 3464-3478, 2014.

[16] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Verlag, 1999.

[17] N. Antonello. (2015). *Sound field control in a reverberant room using the FDTD method Videos* [Online]. Available FTP: ftp://ftp.esat.kuleuven.be/ Directory: SISTA/nantonel/Videos/AES60Leuven/