

NON-CUBOID ITERATIVE ROOM OPTIMIZER

Rinaldi Petrolli¹ Peter D’Antonio¹ John Storyk¹
Jonathan Andrew Hargreaves² Timo Betcke³

¹ REDI Acoustics, Highland, NY, USA

² Acoustics Research Group, University of Salford, UK

³ Department of Mathematics, University College London, UK

rinaldi.petrolli@rediacoustics.com, peter.dantonio@rediacoustics.com

ABSTRACT

In small room acoustics, the range from the first resonant frequency up to the Schroeder frequency is dominated by modal resonances and the Speaker-Boundary Interference Response (SBIR). Both are very sensitive to the positioning of sources, receivers and room geometry. The source locations determine which modes are excited, and the listener locations determine which modes are heard. In past years, various iterative optimization programs emerged to separately determine the optimal room ratios, sources and listening positions of perfectly reflective cuboid rooms, through the use of the image-source model. Despite its fast computation times, this approach does not account for scattering, phase change at the boundary and cannot be extended to non-cuboid rooms.

The present work presents a solution to those issues by using the Boundary Element Method (BEM) to compute the frequency response at low-frequencies, considering the effects of the boundary’s complex admittance and all acoustical elements inside the room. With BEM as its engine, a Room Optimization Genetic Algorithm (ROGA) was developed to optimize source and receiver positions simultaneously with the room geometry, aiming to present the best possible acoustic environment given imposed restraints.

To control the room’s temporal decay, low-frequency acoustic treatments were added to the BEM model. By using Transfer Matrix Models, the acoustical behavior of different multilayered treatments can be modeled and inserted into the BEM simulation to evaluate the change in the room’s acoustic field and in the frequency response at the receiving positions. 3D waterfall plots illustrate the temporal decay following optimization. Examples will be presented.

Keywords: low-frequency - optimization - boundary element method - genetic algorithm

1. INTRODUCTION

Critical listening environments are crucial to allow content creators and audio engineers to hear the real sound of their work without the room’s influence. The finished content will travel through several different reproduction environments with different responses, so making sure that the ini-

tial material is not colored is the only way to ensure that it will translate adequately to different systems. To achieve that, one has to create a neutral critical listening environment that minimizes the effects of the room on the reproduced sound. In other words, if you cannot get the room out of the mix, you cannot take the mix out of the room.

We must consider all the influences that the room imposes on the sound arriving at the listener. For low-frequencies, the location of the listener, sources and room boundaries are of particular concern given the wavelength size. Correctly placing the mix position in a control room can be the difference between hearing a specific frequency too much or not hearing it at all. This is due to the minimum and maximum pressure nodes that are spread out across the room, which are directly related to the room’s shape and dimensions and where the source and listener are located.

Previous work done by Cox and D’Antonio, namely RPG’s Room Optimizer [1] and Room Sizer [2], addressed this issue. Both programs used the image-source model to calculate the impulse response of a cuboid room. [1] minimized the standard deviation of the weighted modal response and SBIR to determine the positions of speakers and listeners. [2] minimized the standard deviation of the modal response to determine the optimal dimensional ratios of the room. They have fast computing times, but are limited to cuboid-shaped rooms, due to the image-source model geometry limitations.

Cuboid approximations, with equal volume, of irregular shaped rooms are a common practice among acousticians due to its practicality. It allows for modal frequency response predictions and analytical reverberation time estimations without the need of a 3D model and complicated geometries. However, that does not mean it is precise; it just means it is incredibly convenient. To evaluate the accuracy of that approach, an experiment was conducted to verify whether the improvements done in the cuboid model would be valid in the actual irregular shaped room. For that, 16 different non-cuboid control rooms were modeled with their actual geometry, using BEM, and their cuboid equivalent models were optimized, using [1] and [2]. After the optimization, the recommended changes were used in a BEM model to see if the solutions were correct. In some cases, the improvement was correct; in others, it was not. No trend that determines if the improvement would be

present be in the 3D BEM model was found.

To solve this problem for both cuboid and non-cuboid rooms, we developed a wave based program to simultaneously optimize the room geometry and the positions of the speakers and listeners, which will be described in this paper.

2. THEORY AND IMPLEMENTATION

2.1 Room modes

The prediction of room modes in cuboid room is widely covered in the acoustics literature [3]. In summary, it is the solution of the wave equation for a rectangular room with perfectly reflecting walls, which yields Eq. 1, where $f_{n_x n_y n_z}$ are the natural frequencies, L_x , L_y , L_z are the dimensions of the rectangular room, n_x , n_y , n_z are the modal numbers and c_0 is the speed of sound.

$$f_{n_x n_y n_z} = \frac{c_0}{2} \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2} \quad (1)$$

This has been often used to optimize the dimensions of rooms to avoid degenerated modes where multiple natural frequencies fall into a narrow frequency band, but optimizing dimensional ratios is not a necessary and sufficient condition for room design. Even though it will describe the resonant frequency of the modes, it says nothing about their bandwidth and pressure. Moreover, that is all considering a cuboid room, and when we consider complex-shaped rooms one cannot rely on this prediction to find even the resonant frequencies of the modes. More complex, irregular geometries of variable boundary impedance demand the application of discretized models and computational procedures such as the finite element and boundary element methods [4].

2.2 Speaker-Boundary Interference Response

When a speaker is positioned inside a room, in addition to the room modes, there will also be peaks and dips that are caused by the interference from the speaker location in relation to the room boundaries. This effect, known as the SBIR [1], or the Alisson effect [5], presents itself across the entire frequency range. However, its effect is much more present in the lower frequencies. Its behavior can be seen on Figure 1, where the peaks and dips are related to the distances from the speaker to the room boundaries. The example shown in Figure 1 is for a speaker placed 1.2 m from one, two, and three walls. With each added equidistant boundary the SBIR becomes more evident, increasing the notch and deepening the dip. It also shows how its frequency is dependent on the distance from the boundaries by placing the speaker equidistant from the three boundaries by 0.3 m.

In order to find its effect on the modal response of a room, one can window the room impulse response to extract the SBIR, as described in [1]. The windowing of the impulse response can be seen in Figure 2. It consists of

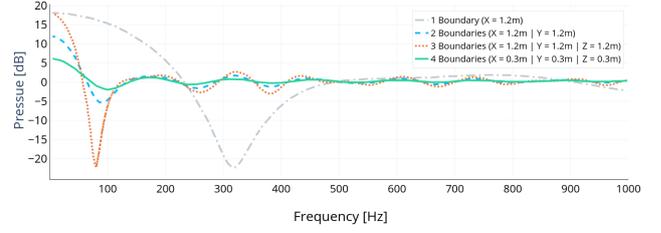


Figure 1: Averaged SBIR for several boundary configurations. Adapted from [1].

a half cosine squared window that is flat up to 32 ms and then follows the cosine squared function out to 64 ms.

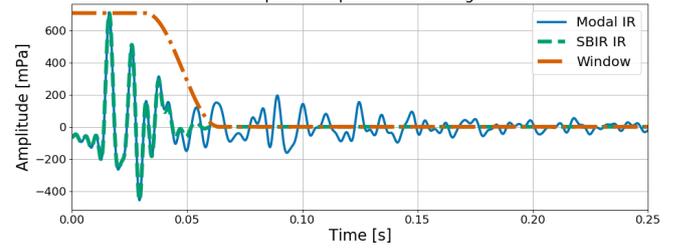


Figure 2: Windowing of the room impulse response to extract the SBIR.

The Fast Fourier Transform of this windowed response will represent the SBIR, as shown in Figure 1, which can also be achieved through the convolution in the frequency domain from the window spectra with the evaluated spectra at the receiving position. This is helpful to reduce the computation time when the modal decay information is not desired.

2.3 Solution using the Boundary Element Method

It will be assumed that the air in the room, is linear, homogeneous and isotropic, with wave speed $c_0 = 343$ m/s and density $\rho_0 = 1.21$ kg/m³. The objective is to compute the scattered acoustic pressure P_s and total acoustic pressure P_t that occurs when the room is driven by a time-harmonic source at some frequency f in Hertz. The scattered sound-field P_s in the room satisfies the time-harmonic wave equation

$$\nabla^2 P_s + k^2 P_s = 0, \quad (2)$$

where $k = 2\pi f/c_0$ is the wavenumber in radians per meter. The acoustic properties of the room boundaries are described by an admittance boundary condition, which relates boundary pressure to U_{in} , the boundary-normal particle velocity into the material, by $U_{in}(\mathbf{x}) = (Y_n(\mathbf{x})P(\mathbf{x})) / (\rho_0 c_0)$, where \mathbf{x} is the position in a 3D Cartesian space. Here Y_n is the frequency dependent normalized specific acoustic admittance, which may vary spatially, and the complex pressure amplitude field $P(\mathbf{x})$ satisfies Eq. 2.

The Boundary Element Method will be used to solve this problem. This has four stages:

- (i) The room boundary is discretized into boundary elements, on which pressure and its derivatives are interpolated.
- (ii) The incident pressure field $P_i(\mathbf{x})$ arriving from the sound sources is computed at the room boundary.
- (iii) The scattered pressure field $P_s(\mathbf{x})$ caused by this incident field $P_i(\mathbf{x})$ is solved for, again at the room boundary. This includes the effect of all the materials present in the room and arbitrarily high-order reflections.
- (iv) The total pressure field $P_t(\mathbf{x}_r)$ is computed at the desired set of receiver points \mathbf{x}_r inside the room.

P_t is defined by $P_t = P_s + P_i$. The use of a separate term P_i to describe the pressure field arriving from the sources makes this a ‘scattering problem’. Physically, P_s is the difference caused by the source being contained in the room compared to it radiating anechoically, therefore, it contains all the modal effects we are interested in.

A suitable description of this problem is through a boundary integral equation. Using Green’s formula we can describe the scattered sound field in the room as

$$P_s(\mathbf{x}) = \int_{\Gamma} G(\mathbf{x}, \mathbf{y}) \frac{\partial P_s}{\partial n}(\mathbf{y}) d\Gamma_{\mathbf{y}} - \int_{\Gamma} \frac{\partial G}{\partial n_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) P_s(\mathbf{y}) d\Gamma_{\mathbf{y}} \quad (3)$$

where $G(\mathbf{x}, \mathbf{y}) = e^{ik|\mathbf{x}-\mathbf{y}|}/4\pi|\mathbf{x}-\mathbf{y}|$ is the free-space acoustic Green’s function and \mathbf{y} is a point on the room boundary Γ . By taking the limit of Eq. 3, as \mathbf{x} approaches the boundary, we arrive at the relationship

$$\frac{1}{2}P_s(\mathbf{x}) + \mathcal{D}\{P_s\}(\mathbf{x}) - \mathcal{S}\left\{\frac{\partial P_s}{\partial n}\right\}(\mathbf{x}) = 0. \quad (4)$$

The boundary integral operators \mathcal{S} and \mathcal{D} in [6] are respectively termed the single-layer and double-layer potential operators. They are defined for some argument boundary field $\phi(\mathbf{y})$ as:

$$\mathcal{S}(\phi)(\mathbf{x}) = \int_{\Gamma} \phi(\mathbf{y}) G(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}} \quad (5)$$

$$\mathcal{D}(\phi)(\mathbf{x}) = \int_{\Gamma} \phi(\mathbf{y}) \frac{\partial G}{\partial n_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) d\Gamma_{\mathbf{y}} \quad (6)$$

The notation $\partial/\partial n_{\mathbf{y}}$ is shorthand for $\hat{\mathbf{n}}_{\mathbf{y}} \cdot \nabla_{\mathbf{y}}$, where $\hat{\mathbf{n}}_{\mathbf{y}}$ is a unit vector pointing normal to Γ and out of the room and the subscript \mathbf{y} means ‘with respect to or evaluated at point \mathbf{y} ’. Physically, the single-layer potential can be interpreted as being the pressure field at \mathbf{x} caused by a fluid flow with density $\phi(\mathbf{y})$ emerging through the boundary. The double-layer potential can be interpreted as being the pressure field at \mathbf{x} caused by a pressure differential with distribution $\phi(\mathbf{y})$ across the boundary. The terms $G(\mathbf{x}, \mathbf{y})$ and $\partial G/(\partial n_{\mathbf{y}})(\mathbf{x}, \mathbf{y})$ have monopole and dipole radiation characteristics respectively.

By incorporating the impedance boundary condition in Eq. 4 we arrive at the final boundary integral equation

$$\left[\frac{1}{2}\mathcal{I} + \mathcal{D} - ikSY_n\right]\{P_s\}(\mathbf{x}) = -\mathcal{S}\{f\}, \quad (7)$$

with $f = \partial P_i/(\partial n)(\mathbf{x}) - ikY_n(\mathbf{x})P_i(\mathbf{x})$ and $\mathcal{I}\{\phi\}(\mathbf{x}) = \phi(\mathbf{x})$ is an identity operator. This may be solved by inserting the discretized version of P_s on the boundary mesh and then solving the resulting matrix equation numerically. Once P_s is known on the room boundary Γ , P_t can be computed for a point \mathbf{x}_r located inside the room by the Helmholtz-Kirchhoff Boundary Integral Equation (HKBIE), which can be written:

$$P_t(\mathbf{x}_r) = \mathcal{S}\left\{\frac{\partial P_s}{\partial n_{\mathbf{y}}}\right\}(\mathbf{x}_r) - \mathcal{D}\{P_s(\mathbf{x}_r)\} + P_i(\mathbf{x}_r). \quad (8)$$

A common choice for representing loudspeakers in BEM is to assume they radiate omnidirectionally like a monopole. In this case P_i is given by $P_i(\mathbf{x}) = -i\omega\rho_0 Q G(\mathbf{x}, \mathbf{x}_s)$, where Q is the volume flow rate of the loudspeaker in m^3/s and \mathbf{x}_s is its location. Alternatively, more sophisticated statements that capture directivity can be used [7]. Neither of these schemes is, however, suitable for loudspeakers embedded in room boundaries, as is common in recording control rooms and high-end listening environments. Mathematically the problem is that $G(\mathbf{x}, \mathbf{x}_s)$ becomes infinite as \mathbf{x} , which is on the boundary, approaches \mathbf{x}_s .

This issue can be circumvented by directly including the loudspeaker cone vibration into the BEM model. If the driving velocity amplitude on the boundary is given by $U_l(\mathbf{y})$, then P_i is given by:

$$P_i(\mathbf{x}) = -i\omega\rho_0 \mathcal{S}\{U_l\}(\mathbf{x}). \quad (9)$$

In practice, U_l will be uniform over each loudspeaker cone and zero elsewhere. U_l could be found by measurement of cone vibration or from standard methods based on Thiele-Small parameters [8].

The BEM simulations were performed using Bempp-cl [6,9] version 0.2.2. This is an open-source Boundary Element Method library that is invoked from Python scripts, creating a flexible interface that allows the boundary integral operators provided to be assembled in customizable ways. Bempp implements a Galerkin BEM algorithm in 3D and includes an Adaptive-Cross-Approximation (ACA) solver that accelerates matrix assembly and solution. For the solution of the Helmholtz equation, Bempp includes the boundary integral operators above, allowing Eq. 7 to be readily solved and Eq. 8 evaluated.

The meshing of the room boundaries was performed using the open-source tool Gmsh [10], which has a Python API that can be encoded along with the other parts of the algorithm to enable iterative mesh construction. The mesh was refined to 3 elements per wavelength of the maximum frequency of analysis [11]. Continuous linear interpolation functions were used on the elements, so the number of unknowns in the matrix to solve equaled the number of mesh nodes.

The BEM calculation is initially done in the frequency domain and the Inverse Fast Fourier Transform (IFFT) is used to determine the Impulse Response (IR) at the receiving positions. To ensure the BEM results taper smoothly at the lowest and highest frequency, to minimize Gibbs artifacts in the time domain, low and high pass filters are applied to the frequency response before the IFFT. A Tukey window is then applied to ensure smoothness at the point where the BEM solutions are truncated in frequency.

To apply this filter and not affect the frequency range of interest, the crossover frequencies have to be below and above our minimum and maximum frequencies of interest, respectively. To mitigate the computational cost that this incurred, the mesh was not refined further beyond the crossover frequency, meaning that computational cost was fixed, but accuracy reduced with increasing frequency [7]. This is however offset by the reducing contribution of those frequencies to the reconstructed IR, due to their attenuation by the crossover filters. IR reconstruction is needed when the time decay information is of interest.

The resolution of the frequency domain determines the length of the IR, denoted by the relation $\Delta f = 1/T$, where Δf is the frequency spacing and T is the IR time length in seconds. To account for a proper representation in the time domain, the minimal length of the IR has to account for the modal decay times at the frequency range of interest. $\Delta f = 0.5$ Hz yields an IR length of 2 seconds, which is enough for smaller rooms, however, higher resolution is needed to account for the modal decay in larger spaces. For rooms containing acoustical treatments $\Delta f = 0.5$ Hz is often sufficient depending on the room type and size.

2.4 Room Optimization Genetic Algorithm

Evolutionary algorithms have been widely used in science and engineering areas to find optimal solutions for non-linear problems. They are search algorithms based on the mechanism of natural selection and evolution, where the goal is to recreate that process to find the fittest candidate to survive a given problem. In the case of room acoustics, a suitable optimization method is expected to find the best possible positions for speakers, listeners and room boundaries within given optimization constraints. This section will describe the basic Genetic Algorithm (GA) concepts used in this work.

2.4.1 Fitness Metric

The proposed objective function corresponds to the weighted sum of the standard deviation of the modal response and the SBIR. This metric is well suited for finding the flattest frequency response possible. Considering the optimization deals with rooms that don't have acoustical treatments, it does not fully represent the objective curve. To address that, three different penalties are defined to calculate the final fitness value. The objective function of the fitness metric is defined as:

$$f(\vec{x}) = \frac{100}{\sqrt{w_1 \sigma \left(\frac{\sum_1^R MR}{R} \right)^2 + w_2 \sigma \left(\frac{\sum_1^R SBIR}{R} \right)^2}} - \beta, \quad (10)$$

where w_1 and w_2 are weights that sum up to 1, σ is the standard deviation operator, MR is the modal frequency response and $SBIR$ is the Speaker Boundary Interference Response, R is the number of receiving points and \vec{x} is the design vector that contains the room vertices, sources and receivers coordinates. β is the value of all the penalties applied to the fitness metric.

Although the standard deviation is a good metric for recording studios, where coloration from the room is not desired, there are other options that might better suit recreational listening environments, such as the one used by Floody and Venegas [12], which uses the Equal Loudness Curves to achieve an isophonically flat response.

The penalties that are introduced in the fitness metric aim to aid the objective function towards a better response for untreated rooms, where enough acoustic gain is wanted for the lower order modes and no dips in the frequency response are desired. The first penalty reduces the fitness value according to the amount of dips in the frequency response that are below the mean. It is frequency dependent for each dip and has a larger weighting for lower frequencies and larger valleys. The second penalty is related to the distance of the receiver from the center of the room, which will penalize candidates that have the listener close to the room's center. The third penalty is related to the pressure amplitude of the room's first resonance and will penalize candidates that have a first resonance close or below to the mean value of the frequency response. In other words, the higher the amount of dips in the frequency response, the closer the receiver is to the center of the room and the closer the first resonance peak is to the mean value of the curve, the higher the penalty is.

2.4.2 Chromosomal representation

To represent each candidate, the first step is to code all the necessary data into a structure that resembles the way living beings evolve. We do that by coding our parameters into a structure that is a pseudo-chromosomal representation of the real values, in which each parameter will represent one gene in the chromosome, as shown in Figure 3.

$$\boxed{X_1} \boxed{Y_1} \boxed{Z_1} \boxed{X_2} \boxed{Y_2} \boxed{Z_2} \boxed{X_3} \boxed{Y_3} \dots$$

Figure 3: Pseudo-chromosomal representation of the parameters, where each block contains the X, Y and Z coordinates of one of the variables in the model.

2.4.3 Genetic Operations

To start the GA, a suitable initial population has to be created. For room optimization, this has to account for positioning limitations of speakers, listeners and the room vertices. The variability of each of the positions and bound-

aries along with the angle and distance constraints are applied in this step. Next, the evolutionary process is started, which is shown in Figure 5 in a block diagram. The description of the genetic operators are as follows:

1. Selection

The selection of candidates to continue to the next step is based on the fitness of a specific candidate in relation to the total fitness of the population. The biased roulette wheel method was used, where a wheel is spun N times, with N being the population size, and the probability of each candidate being select is based on the ratio of their individual fitness over the total population fitness.

2. Crossover

Two random parents are chosen to breed a new solution based on their current genes. To choose what gene are going to be selected, a coin toss selects the first gene and the following ones until we reach the end of the chromosome. The selected genes generate one offspring and the remaining genes are used to generate the second offspring. This process is depicted in Figure 4.

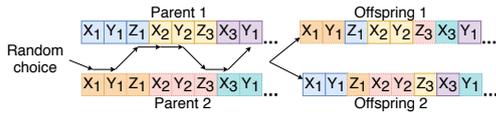


Figure 4: Uniform crossover of two chromosomes. The offspring contain random genes from both parents.

3. Mutation

Applied to the offspring in each generation, mutation is the sporadic random change in one of the genes. This is done by randomly altering one of the values of the chromosome at a random position. It aims to introduce random differences from one generation to the next and keep the population diverse.

4. Elitism

When applying the operators above from one generation to the next, the current best solution can be lost if it is not selected. To avoid that, before the other operators are applied, a small portion of the population is selected to automatically go to the next generation.

With the exception of crossover and mutation, these operations occur once every generation. The rate at which the crossover and mutation operators occur is usually determined by a fixed value, e.g. 0.9 for crossover rate and 0.03 for mutation rate. In order to achieve better convergence for a smaller population size (25 to 200 candidates) than what is usually used in evolutionary algorithms, a Dynamic Decreasing of High Mutation Rate/Increasing of Low Crossover Rate (DHM/ILC) [13] method has been implemented. This method starts with a mutation rate of 1 and crossover rate of 0, which are respectively decreased

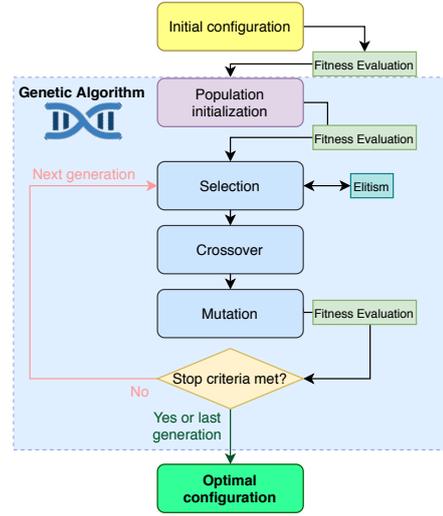


Figure 5: Block diagram of the Room Optimization Genetic Algorithm.

and increased along with the increasing level of generations. The increased mutation at early stages helps to prevent the GA from getting stuck in a local minimum due to a lack of diversity in a small search space. Then the GA is fine tuned to achieve an optimal solution towards the final stages, due to increased crossovers.

2.5 Transfer Matrix Models for treatment designs

The Transfer Matrix Model (TMM) representation, also known as four-pole parameter representation or transmission matrix, enables the calculation of the surface impedance of single and multiple layered treatments by considering continuity in pressure and velocity from one layer to the next, thus enabling the use of surface impedance of one layer as the backing surface of the next. The method works by representing the acoustical properties of the device through its four-pole parameters, which relate pressure and particle velocity on each side of each layer of the device. Consider the single layer acoustic device shown in Figure 6, the transfer matrix of the single layer e_1 can be defined as

$$\begin{bmatrix} p_{in} \\ u_{in} \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} p_{out} \\ u_{out} \end{bmatrix}, \quad (11)$$

where p_{in} , p_{out} and u_{in} , u_{out} are pressure and particle velocity on each side of the layer and A_1 , B_1 , C_1 and D_1 are the four-pole parameters.

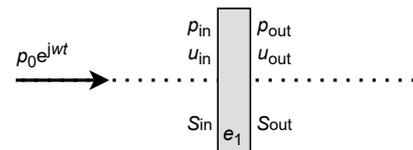


Figure 6: Single layer acoustical device.

A multilayered system can be represented by adding the four-pole parameters of the other layers, which for a double layer device results in the following matrix

$$T_G = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} = \begin{bmatrix} A_G & B_G \\ C_G & D_G \end{bmatrix}, \quad (12)$$

in which A_2, B_2, C_2 and D_2 are the four-pole parameters of the second layer. The multiplication of these matrices can be defined as the global transfer matrix T_G with A_G, B_G, C_G and D_G as the global four-pole parameters. With the global transfer matrix, the input impedance at the top layer can be calculated as [14]:

$$Z_{in} = S_{in} \frac{B_G + A_G \widetilde{Z}_{out}/S_{out}}{D_G + C_G \widetilde{Z}_{out}/S_{out}}, \quad (13)$$

where S_{in} and S_{out} are the surface areas of the first and last layers and \widetilde{Z}_{out} is the radiation impedance at the last layer. If the last layer has a rigid backing ($u_{out} = 0$ and $\widetilde{Z}_{out} \rightarrow \infty$) the input impedance can be defined as $Z_{in} = S_{in} \frac{A_G}{C_G}$.

In room acoustics, it is common to find locally and non-locally reactive acoustic treatments, *i.e.* devices that are independent and dependent of the angle of incidence θ_i , respectively. To account for the extended reaction of non-locally reactive acoustic treatment, the field admittance Y_{field} can be calculated [11]

$$Y_{field} = \frac{\int_0^{78} Y(\theta) \sin(\theta) d\theta}{\int_0^{78} \sin(\theta) d\theta}, \quad (14)$$

where $Y(\theta)$ is the angle dependent admittance, Z_{field} is the field impedance ($Z_{field} = 1/Y_{field}$) and θ is the incidence angle.

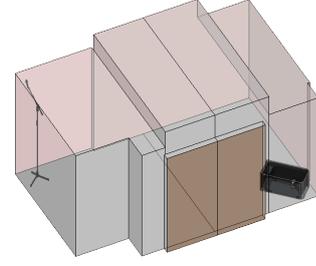
3. RESULTS

3.1 Model validation

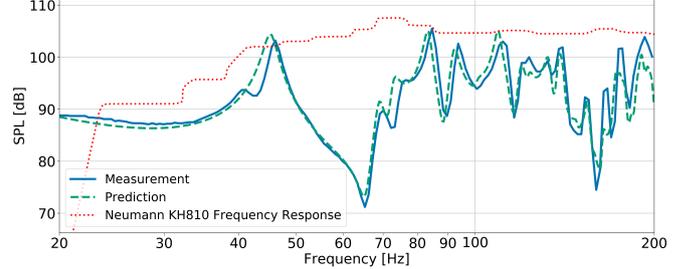
To assess the accuracy of our predictions, a test room was developed. The room contained two CMU walls and a concrete floor. The other walls and ceiling were modified to contain a gypsum-plywood-gypsum sandwich construction which reflects the isolation and modal damping in untreated recording control rooms.

A measurement was made to compare the real world modal response with the one predicted in our simulation environment, in which an omnidirectional microphone was placed in one of the room's corners and a sub-woofer was placed in the opposite diagonal corner. This represents the worst case scenario, in which all the modes are being excited by the speaker and all the modes are being heard by the receiver. An Earthworks M30 omnidirectional microphone and a Neumann KH810 [15] sub-woofer were used. The speaker was chosen due to its relatively flat frequency response in the 20 to 200 Hz region. A 3D model of the measurement setup can be seen in Figure 7a.

Some uncertainties remains, like the complex impedance of the door, which in the model was considered to have the same admittance as the walls, and some minor differences between the different walls due to the pre-existing construction. However, the agreement



(a) 3D model of the measurement setup.



(b) Comparison between the measured frequency response and the frequency response obtained from the BEM simulation. Measured response in solid blue line, simulated in dashed green and sub-woofer response in dotted red.

Figure 7: Validation measurement setup and results.

between the measurement and prediction is good and proves the accuracy of the algorithm to predict the modal range of a small complex shaped room with the use of a monopole source. The best match between measurement and simulation was found to be with the normalized admittance values of the walls to be of 0.015.

3.2 Control Room Optimization Example

The optimization will start with a cuboid room with the ratios 1:1.55:1.85, as proposed by Cox and D'Antonio in [1], with an initial room volume of 75 m³. It will consider a symmetric pair of sub-woofers and a single listening position. Figure 8 depicts the acoustical center [16] of the speakers in dark red squares and the receiving points in green circles, which are used to create an average frequency response at the listening position. It also shows the room vertices in purple circles around the room boundaries. The origin of the coordinates system is at the lower center of the rear wall, with X being the room width, Y the length and Z the height.

The constrained search ranges for vertices and positions of speakers and listeners are shown in Table 1. The initial positions and cuboid geometry can be seen in Figures 8a and 8b. The genetic algorithm was set to have a population of 150 candidates and run for 300 generations. The optimized positions and room geometry are shown in Figures 8c and 8d. The frequency range was set from 20 to 150 Hz in order to account for both the Schroeder frequency and the maximum crossover frequency usually found in sub-woofers.

It is important to note that the objective of this optimization is to find the best architecturally possible acoustic performance, which must incorporate some usual aspects of

Axis	X [m]	Y [m]	Z [m]
Vertex	± 0.5	± 0.5	± 0.5
Sources	± 3.0	± 3.0	± 3.0
Receiver	± 0.0	± 2.0	± 0.0

Table 1: Search range for room vertices, sources and receiver.

control room design. With that in mind, the ceiling was constrained to have an upward slope toward the back of the room, the receiving position was constrained to be at least 2 m from the front and rear walls and centered in the X axis, and the source position was constrained to be at least 0.7 m from any of the walls. The distance between source and receiver was also set to a minimum of 1.5 m. All three penalties previously described were applied. This example is the best case scenario where the layout and geometry of the room are free to move.

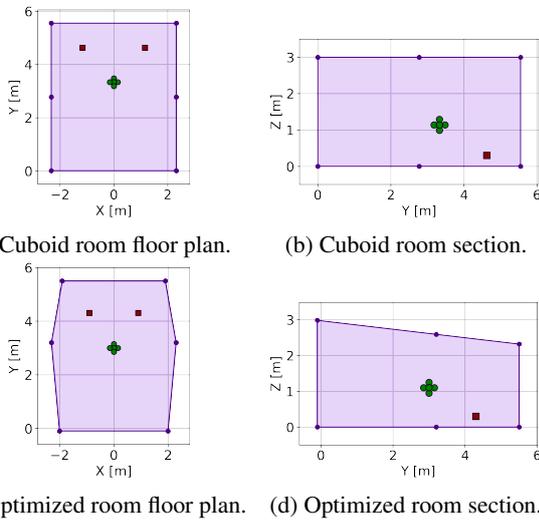


Figure 8: Cuboid room with ratios 1:1.55:1.85 and optimized room floor plans and sections.

Figure 9 shows the frequency responses for the cuboid room in dashed green and for the optimized geometry and positions in dotted orange. The standard deviation after the optimization decreased by 1.86 dB, but more importantly, all the dips (with the exception of the antiresonance before the first axial mode) were minimized. At this stage of the design, this is the desired response in order to apply acoustic treatments to reduce the modal peaks and control the temporal decay in the room, which will be addressed next.

3.3 Treatment design

When dealing with low-frequencies, the common practice used by acousticians is to place low-frequency treatments around the room corners. This is due to the fact that low-frequency treatments are pressure absorbers (membrane and Helmholtz resonators) and have the highest efficiency where pressure is maximum. This is a simple assumption for the lower order modes, even in a non-cuboid room. However, once the shape of the room starts to deviate from

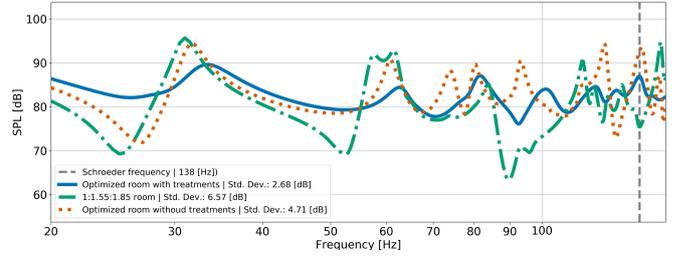
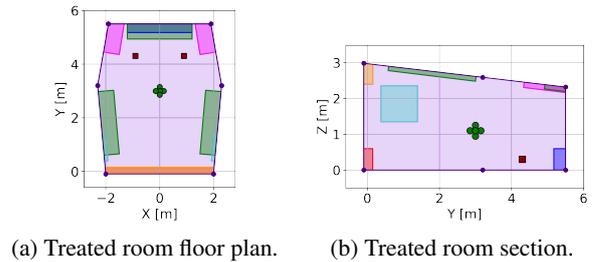


Figure 9: Optimization results. Cuboid room in dashed green, optimized room in dotted orange and optimized room with treatments in solid blue.

a cuboid shape and we start to go up in frequency, that becomes very hard to estimate accurately.

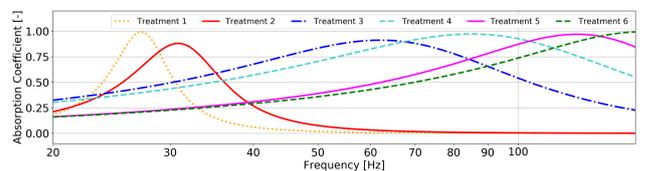
To address that, the pressure distribution in the room is used to define treatment placement. With the critical frequencies that need treatment shown in the frequency response, one can analyze the pressure distribution inside the room and place each treatment at a high pressure zone for each target frequency.

From the frequency response shown in Figure 9, the critical frequencies were selected and 6 different acoustic treatments were designed using the TMM. The treatment layout and their performance are shown in Figure 10. Treatment 1 and 2 are membrane resonators and treatments 3 to 6 are perforated Helmholtz resonators.



(a) Treated room floor plan. (b) Treated room section.

(c) 3D mesh with treatments.



(d) Treatments absorption coefficients.

Figure 10: Treatment layout and absorption coefficients.

The solid blue curve in Figure 9 shows the frequency response of the optimized room with the application of the treatments, and the reduction in the Q of the treated frequency peaks is noticeable. In order to fully evaluate the effectiveness of the treatment application, one can evalu-

ate the temporal decay of the room impulse response. Figure 11 shows a waterfall plot of the room impulse response before and after the application of treatments. The reduction in decay time is evident, roughly going from 0.9 s without treatments to 0.3 s with treatments.

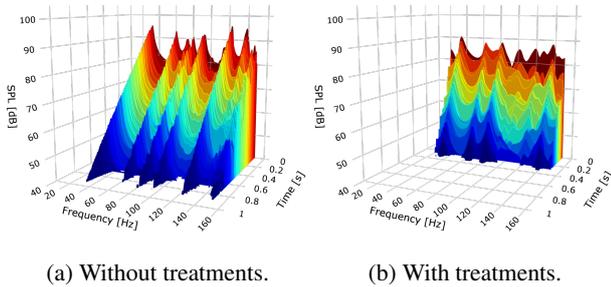


Figure 11: Waterfall plot of the optimized room without acoustic treatments to the left and with acoustic treatments to the right.

4. CONCLUSION

This paper presents a new iterative method for optimizing the acoustical performance of a room given the architectural constraints. It relies on a BEM formulation to predict the room response and uses a genetic algorithm to perform the optimization. The proposed fitness metric optimizes the standard deviation of both the modal response and the SBIR, along with penalties for frequency dips, receiver positioning and emphasis of first order axial modes. Through the use of Transfer Matrix Models to predict the complex surface impedance of acoustic treatments, the optimized design can be further improved to control the modal decay at low-frequencies.

The model was validated experimentally and the agreement is excellent, even with the considerations of frequency independent wall admittance and of the unaccounted influence of a door in the test laboratory. A case study was carried out in order to verify both the effectiveness of the optimization algorithm and of the proposed treatment design. The results are satisfactory and prove that room ratios are not a sufficient metric for a good low-frequency response. It also highlights the need to simultaneously optimize the room geometry and the positions of the speakers and listeners to find the best possible acoustic response within the architectural constraints.

5. REFERENCES

[1] P. D’Antonio and T. J. Cox, “Room optimizer: A computer program to optimize the placement of listener, loudspeakers, acoustical surface treatment and room dimensions in critical listening rooms,” in *Audio Engineering Society Convention 103*, Sep 1997.

[2] T. Cox, P. D’Antonio, and M. Avis, “Room sizing and optimization at low frequencies,” *Journal of the Audio Engineering Society. Audio Engineering Society*, vol. 52, 06 2004.

[3] H. Kuttruff, *Room Acoustics*. Taylor & Francis, 2000.

[4] F. Fahy, *Foundations of Engineering Acoustics*. Elsevier Academic Press, 2001.

[5] R. F. Allison, “The influence of room boundaries on loudspeaker power output,” *J. Audio Eng. Soc.*, vol. 22, no. 5, pp. 314–320, 1974.

[6] W. Smigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger, “Solving boundary integral problems with BEM++,” *ACM Trans. Math. Softw.*, vol. 41, 02 2015.

[7] J. A. Hargreaves, L. R. Rendell, and Y. W. Lam, “A framework for auralization of boundary element method simulations including source and receiver directivity,” *The Journal of the Acoustical Society of America*, vol. 145, no. 4, pp. 2625–2637, 2019.

[8] T. M. Leo Beranek, *Acoustics: Sound Fields, Transducers and Vibration*. Academic Press, 2 ed., 2019.

[9] T. Betcke, E. van ’t Wout, and P. G elat, *Computationally Efficient Boundary Element Methods for High-Frequency Helmholtz Problems in Unbounded Domains*, pp. 215–243. Cham: Springer International Publishing, 2017.

[10] C. Geuzaine and J.-F. Remacle, “Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities,” *International Journal for Numerical Methods in Engineering*, vol. 79, pp. 1309 – 1331, 09 2009.

[11] M. Aretz, *Combined wave and ray based room acoustic simulations of small rooms : [challenges and limitations on the way to realistic simulation results]*. PhD thesis, Berlin, 2012. Zugl.: Aachen, Techn. Hochsch., Diss., 2012.

[12] S. Floody and R. Venegas, “Shape optimization of polygonal rooms based on spatially homogeneous sound field distribution and psychoacoustic criteria at low frequencies,” vol. 31, 10 2009.

[13] A. Hassanat, K. Almohammadi, E. Alkafaween, E. Abunawas, A. Hammouri, and S. Prasath, “Choosing mutation and crossover ratios for genetic algorithms-a review with a new dynamic approach,” p. 390, 12 2019.

[14] G. Scavone, “An acoustic analysis of single-reed woodwind instruments with an emphasis on design and performance issues and digital waveguide modeling techniques,” 1997.

[15] “Neumann KH-810 technical data.” <https://ende.neumann.com/kh-810technical-data> (accessed October 2020).

[16] J. Vanderkooy, “Applications of the acoustic centre,” in *Audio Engineering Society Convention 122*, May 2007.